

А.М. Голиков

**КОДИРОВАНИЕ И ШИФРОВАНИЕ ИНФОРМАЦИИ
В СИСТЕМАХ СВЯЗИ**

Учебное пособие

**для специалитета: 11.05.01 - Радиоэлектронные системы и
комплексы (Радиоэлектронные системы передачи информации)**

Курс лекций, компьютерный практикум, задание на самостоятельную
работу

Второе издание дополненное и переработанное

Томск 2017

Министерство образования и науки Российской Федерации
Федеральное государственное бюджетное образовательное
учреждение высшего образования
**Томский государственный университет систем управления и
радиоэлектроники**

А.М. ГОЛИКОВ

**КОДИРОВАНИЕ И ШИФРОВАНИЕ ИНФОРМАЦИИ
ВСИСТЕМАХ СВЯЗИ**

Учебное пособие

для специалитета: 11.05.01 - Радиоэлектронные системы и комплексы

(Радиоэлектронные системы передачи информации)

Курс лекций, компьютерный практикум, задание на самостоятельную работу
Второе издание дополненное и переработанное

Томск 2017

УДК 621.39(075.8)

ББК 32.973(я73)

Г 60

Голиков А.М.

Кодирование и шифрование информации в системах связи. Учебное пособие для специалитета: 11.05.01 - Радиоэлектронные системы и комплексы (Радиоэлектронные системы передачи информации). Курс лекций, компьютерные лабораторные работы, компьютерный практикум, задание на самостоятельную работу. Второе издание дополненное и переработанное / А.М.Голиков. – Томск: Томск. гос. ун-т систем упр. и радиоэлектроники, 2017. – 746 с.: ил. — (Учебная литература для вузов)

Учебное пособие предназначено для направления подготовки специалистов по направлению 11.05.01 - Радиоэлектронные системы и комплексы (Радиоэлектронные системы передачи информации). Представляет учебное пособия для курса "Кодирование и шифрование информации в радиоэлектронных системах передачи информации". Современные учебные курсы редко рассматривают комплексно вопросы модуляции и кодирования, а также их сигнально-кодовые конструкции. Мало учебников и для реализации современных модемов и кодеков. Актуальность пособия велика, так как в современных системах связи и телевидения, а также кабельных сетях применяются все более сложные виды модуляции и кодирования, обеспечивающие высокую помехоустойчивость.

Методология изучения курса состоит в закреплении теоретических знаний на примерах компьютерной реализации модемов и кодеков современных телекоммуникационных систем и индивидуальных заданий на самостоятельную работу.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ.....	7
1. ЦИФРОВЫЕ ВИДЫ МОДУЛЯЦИИ И СИГНАЛЬНОГО КОДИРОВАНИЯ, ИХ СПЕКТРАЛЬНАЯ И ЭНЕРГЕТИЧЕСКАЯ ЭФФЕКТИВНОСТЬ.....	14
1.1. Анализ цифровых методов модуляции.....	14
1.2. Модемы сотовой связи FSK, MSK GMSK и численный анализ вероятности символьной ошибки с использованием ПО LabVIEW.....	30
1.3. Модемы спутниковых систем связи M-QAM, M-PSK и численный анализ вероятности символьной ошибки с использованием ПО LabVIEW.....	47
2. ПРОПУСКНАЯ СПОСОБНОСТЬ КАНАЛА СВЯЗИ. КОДИРОВАНИЕ ИСТОЧНИКА.....	69
2.1. Пропускная способность канала связи. Объем сигнала и емкость канала связи, условия их согласования.....	69
2.2. Исследование кодирования источника дискретных сообщений методами Шеннона-Фано.....	71
2.3. Исследование алгоритмов Лемпеля - Зива.....	78
2.4. Фрактальные методы кодирования изображений.....	87
2.5. Вейвлет преобразования сигналов и изображений с использованием.....	99
3. ПОМЕХОУСТОЙЧИВОЕ КОДИРОВАНИЕ В ТЕЛЕКОММУНИКАЦИОННЫХ СИСТЕМАХ.....	122
3.1. Исследование кодов Хемминга, БЧХ (Боуза-Чоудхури-Хоквенгема), Рида- Соломона на базе MATLAB.....	122
3.2. Циклические избыточные коды CRC (Cyclic redundancy check).....	156
3.3. Сверточные коды. Декодирование сверточных кодов.....	170
3.4. Декодирование сверточных кодов по методу Витерби с использованием ПО MATLAB.....	191
3.5. Турбокодирование. Обобщенная схема турбокодера с параллельным каскади- рованием. Сверточные турбокоды. Декодирование турбокодов. Характеристики помехоустойчивости сверточных турбокодов . Исследование турбокодов с использованием ПО MATLAB.....	211
3.6. Низкоплотностные коды. Классификация LDPC-кодов. Методы построения проверочных матриц. Алгоритмы декодирования низкоплотных кодов. Оценка сложности алгоритмов декодирования на базе MATLAB и LabVIEW.....	225
3.7. Исследование каскадных кодов.....	243

4. СИГНАЛЬНО-КОДОВЫЕ КОНСТРУКЦИИ В ТЕЛЕКОММУНИКАЦИОННЫХ СИСТЕМАХ.....	265
4.1. Сигнально-кодovые конструкции на основе Треллис кодовой модуляции (TCM) и их анализ с использованием MATLAB.....	292
4.2. Исследование сигнально-кодovой конструкции на базе системы с ортогональным частотным мультиплексированием и пространственно-временным кодированием OFDM - MIMO.....	285
5. ОПТИМИЗАЦИЯ МЕТОДОВ ПОМЕХОУСТОЙЧИВОГО КОДИРОВАНИЯ ДЛЯ СИСТЕМ СВЯЗИ (ЗАДАНИЕ НА САМОСТОЯТЕЛЬНУЮ РАБОТУ).....	318
6. КЛАССИЧЕСКИЕ ШИФРЫ	336
6.1. Теория классических шифров ..	375
6.2. Компьютерный практикум для классических шифров	402
6.3. Задания на самостоятельную работу по классическим шифрам.....	426
7. ШИФРОВАНИЕ С СЕКРЕТНЫМ КЛЮЧОМ	454
7.1. Теория шифров с секретным ключом.....	454
7.2. Компьютерный практикум для шифров с секретным ключом....	552
8. ШИФРОВАНИЕ С ОТКРЫТЫМ КЛЮЧОМ	610
8.1. Теория шифров с открытым ключом	610
8.2. Компьютерный практикум для шифров с открытым ключом....	634
9. КРИПТОГРАФИЧЕСКИЕ ПРОТОКОЛЫ В СЕТЯХ ПЕРЕДАЧИ ДАННЫХ	681
9.1. Теория шифров с открытым ключом криптографические протоколы в сетях передачи данных.....	681
9.2. Компьютерный практикум для шифров с открытым ключом....	712
10. ШИФРОВАНИЕ В СОВРЕМЕННЫХ СИСТЕМАХ СВЯЗИ	723
10.1. Безопасность GSM сетей.....	723
10.2. Криптографическая защита беспроводных сетей стандартов LTE.....	726
ЗАКЛЮЧЕНИЕ.....	743
ЛИТЕРАТУРА.....	745

ЧАСТЬ 1. КОДИРОВАНИЕ В РАДИОЭЛЕКТРОННЫХ СИСТЕМАХ ПЕРЕДАЧИ ИНФОРМАЦИИ

ВВЕДЕНИЕ

В зависимости от назначения и условий функционирования системы радиосвязи ее эффективность оценивается на основании тех или иных показателей (критериев), основными из которых являются энергетический и спектральный. Соответственно, важнейшими характеристиками любой системы радиосвязи являются энергетическая и спектральная эффективность, характеризующие, соответственно, энергетические затраты и полосу занимаемых частот, необходимые для передачи сообщений.

К сожалению, одновременное достижение предельных значений этих показателей эффективности оказывается невозможным, так что в каждом конкретном случае построения системы радиосвязи приходится руководствоваться компромисными соображениями при оптимизации характеристик и режимов функционирования системы [1-4].

Эффективные методы модуляции и помехоустойчивого кодирования все шире используются в современных цифровых телекоммуникационных системах (ТКС). Переход в ТКС к ансамблям многопозиционных сигналов увеличивает информационную скорость и обеспечивает передачу больших потоков информации. Современная элементная база позволяет применять в ТКС достаточно сложные методы помехоустойчивого кодирования и обеспечивать тем самым высокую верность передачи информации. Исследованию методов модуляции и кодирования в ТКС посвящена обширная литература, отражающая как достижения теории, так и вопросы ее практических приложений [1-18].

Как известно, теория сигналов и теория кодирования длительное время развивались независимо. В последние годы значительно возрос интерес к новому перспективному направлению, возникшему на стыке этих наук. В работах отечественных и зарубежных авторов интенсивно исследуются возможности ТКС, в которых для передачи информации используются ансамбли многопозиционных сигналов в сочетании с помехоустойчивыми кодами, причем, процедуры модуляции/кодирования (демодуляции/декодирования) осуществляются *совместно*. При рациональном построении такие *сигнально-кодовые конструкции* (СКК) сочетают в себе положительные качества как многопозиционных ансамблей сигналов, так и помехоустойчивых кодов, допускают достаточно простые и реализуемые на практике алгоритмы декодирования и при использовании их в ТКС позволяют существенно продвинуться к теоретическим пределам эффективности. Вопросы синтеза таких систем модуляции/кодирования, анализа их структуры, помехоустойчивости и

эффективности, демодуляции/декодирования составляют основное содержание нового перспективного направления в теории связи - *теории сигнально-кодовых конструкций*.

Один путь базируется на введенном А.Г. Зюко [1, 2] понятии эффективности систем связи. Анализ и сравнение эффективности систем с многопозиционными сигналами и систем с корректирующими кодами приводит в *информационной теории СКК* к идее *их комбинации в единой сигнально-кодовой конструкции*. Такой путь является наглядным, убедительным и излагается без строгих математических выкладок. Этот путь может быть использован для популярного изложения оснований теории СКК [1-4].

Статистическая теория связи предлагает большое количество вариантов построения телекоммуникационных систем. Как из этого множества выбрать вариант, наиболее целесообразный в заданных условиях? По каким критериям следует производить этот выбор? Насколько оправдано применение тех или иных новых систем и как совершенствовать существующие системы? Принципиальное решение этих вопросов в конечном итоге сводится к *оптимизации систем связи по критериям эффективности*.

Типовая структура одноканальной системы передачи дискретной информации (СПДИ) приведена на рис. 1.1. Источник вырабатывает сообщения, которые необходимо передавать по каналу СПДИ. Это могут быть последовательности *дискретных* сообщений (данные, телеграфные сообщения и т.д.) либо *непрерывные* сообщения (речь, телевидение, результаты телеизмерений и др.), преобразованные в *цифровую форму*.

Реальные сообщения содержат *избыточность* и для согласования источника с каналом передачи информации используют *кодер источника*. Совместно с *декодером* они образуют *кодек источника*. Методы кодирования источника изложены в работах [2, 3, 7, 9-11, 13-16, 18].

Основные требования к СПДИ формулируются достаточно просто:

1. ***верность,***
2. ***скорость,***
3. ***своевременность доставки информации от отправителя к получателю.***

В системах без помехоустойчивого кодирования верность и скорость зависят от вида используемых сигналов-переносчиков. Применение помехоустойчивого кодирования позволяет повысить верность передачи, но за счет снижения скорости. С другой стороны, выбором сигналов можно добиться повышения скорости передачи информации, но, зачастую, в ущерб верности передачи. Однако, следует учитывать, что применение корректирующих кодов вносит задержку в передачу цифровых данных, что ухудшает такой показатель, как своевременность доставки информации. В каждом конкретном случае телекоммуникационных систем (ТКС) могут быть сформулированы количественные

выражения этих требований. Это зависит от назначения ТКС и вида передаваемого сообщения. В общем случае показатели верности, скорости и своевременности передачи информации находятся в противоречивых соотношениях.

Кроме отмеченных выше, следует упомянуть о *дополнительных требованиях*, которые определяют применение того или иного метода передачи информации:

Спектральная эффективность. Современные ТКС работают в условиях дефицита частотного спектра, отводимого для передачи информации. Это и обуславливает важность методов формирования *компактных спектров сигналов*, передаваемых по линиям связи с ограниченной полосой частот.

Сложность реализации методов модуляции-кодирования. Теория информации указывает путь повышения помехоустойчивости - использование для передачи *длинных последовательностей сигналов и кодов*, для извлечения информации из которых на приемной стороне приходится применять достаточно сложные алгоритмы обработки. Поэтому *конструктивная теория кодирования* направлена на поиск корректирующих кодов, допускающих *реализуемые алгоритмы декодирования* [4].

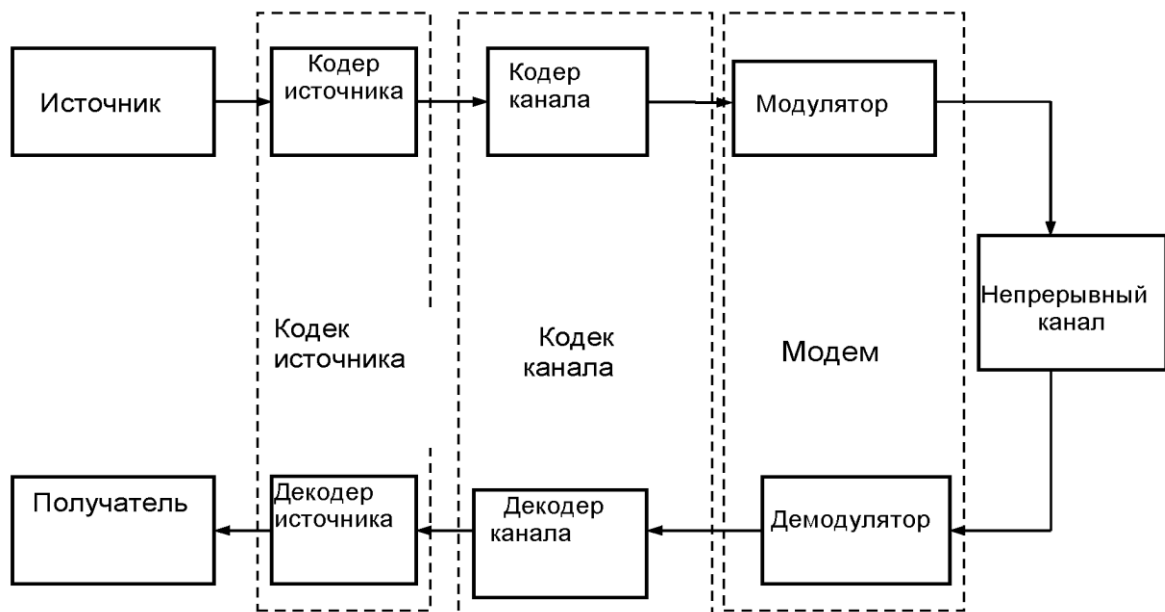


Рис. 1.1. Структурная схема СПДИ

Согласно известным теорем К. Шеннона, в принципе *возможно сколь угодно большое повышение верности передачи информации*, если скорость передачи по каналу $R_{кан}$ не превышает пропускной способности канала C_K . Достигается это применением достаточно длинных *корректирующих кодов* (КК). С этой целью в структуру КК вводится *избыточность*.

Кодек КК (кодер и декодер канала) показан на рис. 1.1. В реальных условиях длина кода ограничена допустимой сложностью устройств кодирования и, прежде всего, декодирования, поэтому эффект от применения корректирующих кодов зависит от параметров кода и ограничений на реализацию кодека канала.

Современная теория предлагает широкий набор корректирующих кодов, различных по структуре, принципам построения и корректирующей способности. *Кодем СКК* включает в себя сочетание кодека/модулятора и, соответственно, демодулятора/декодера.

В общем случае результат работы системы связи определяется количеством и качеством передаваемой информации. Количество оценивается скоростью передачи информации по каналу (бит/с), а качество - величиной ошибки. Согласно теоремам К. Шеннона, ошибка при соответствующем выборе метода передачи (модуляции/кодирования) может быть сделана произвольно малой. В то же время, скорость передачи не может быть выше некоторого информационного ресурса, называемого *пропускной способностью канала С*.

В работе А.Г. Зюко [2] было предложено считать одним из показателей эффективности системы величину средней скорости $R_{кан}$ при коМесто для формулы.торой обеспечивается заданная верность передачи информации. При этом могут быть определены следующие *показатели эффективности*:

- ***Информационная эффективность системы, определяющая степень использования пропускной способности канала относительной величиной***

$$\eta = \frac{R_{кан}}{C} \quad (1.1)$$

В реальных условиях показатель η всегда меньше единицы. Чем ближе η к единице, тем совершеннее система передачи информации. Достижение необходимых скорости и верности передачи сопровождается определенными затратами других важнейших ресурсов: ***мощности сигнала P_c и полосы частот канала F_k*** . Такой подход позволил ввести показатели эффективности использования ресурсов системы, а именно:

- ***Энергетическая эффективность***

$$\beta = \frac{R_{кан}}{P_c / N_0} \quad (1.2)$$

где N_0 - спектральная плотность мощности шума;

- ***Частотная эффективность***

$$\gamma = \frac{R_{\text{кан}}}{F} \quad (1.3)$$

где F - полоса частот канала.

Показатели β и γ имеют смысл удельных скоростей, а обратные величины $\beta' = 1/\beta$ и $\gamma' = 1/\gamma$ определяют *удельные расходы соответствующих ресурсов* на передачу информации с единичной скоростью (1 бит/с).

Для гауссовского канала с полосой F , отношением мощностей сигнала и шума $\rho = \frac{P_c}{P_{ш}}$ и пропускной способностью $C = F \log_2(1 + \rho)$ была установлена

связь между показателями эффективности

$$\eta = \frac{\gamma}{\log_2\left(1 + \frac{\gamma}{\beta}\right)} \quad \text{и} \quad \gamma = \rho\beta$$

Для идеальной системы (η) определена также предельная зависимость

$$\beta = \frac{\gamma}{2^\gamma - 1}. \quad (1.5)$$

Зависимость энергетической эффективности от удельной скорости удобно представить в виде кривой на плоскости $\beta\gamma$ (рис. 1.2). Она отражает наилучший обмен между β и γ в непрерывном канале (НК-предел Шеннона). При этом частотная эффективность γ изменяется в пределах от 0 до ∞ то время как энергетическая эффективность ограничена сверху, поскольку

$$\beta_{\text{max}} = \lim_{\gamma \rightarrow 0} \beta = \lim_{\gamma \rightarrow 0} \frac{\gamma}{2^\gamma - 1} = \frac{1}{\ln 2} \approx 1,443(1,59 \text{ дБ}) \quad (1.6)$$

Аналогичные предельные кривые могут быть построены для других типов каналов [2] (ДНК-дискретно-непрерывный канал, ДСК-дискретный симметричный канал на рис. 1.2).

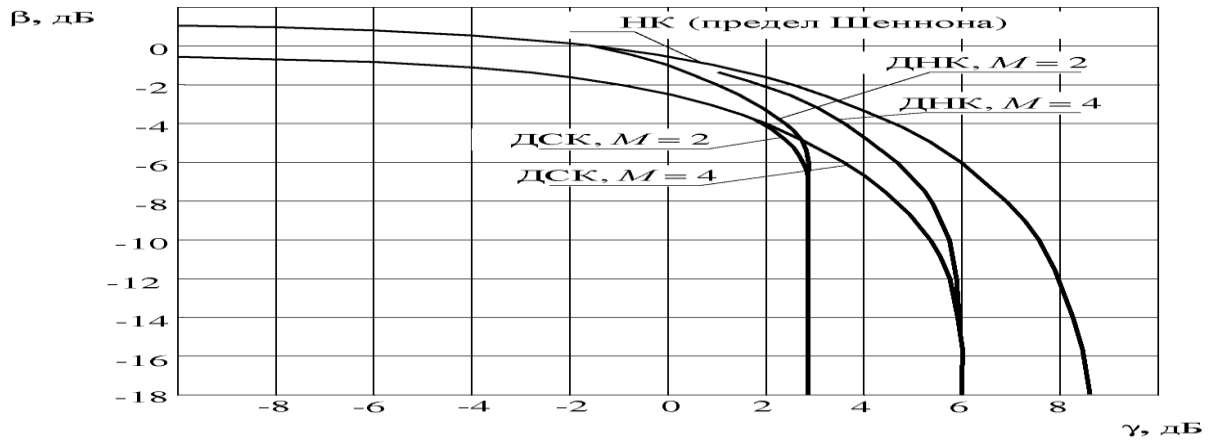


Рис. 1.2. Кривые предельной эффективности систем связи

Полосу частот, занимаемую сигналом конечной длительности не всегда можно определить однозначно. Поэтому далее для анализа и сравнения эффективности будем рассматривать следующую модель процесса передачи информации.

Полагаем, что на интервале T по каналу передается определенное количество независимых отсчетов N_T (по В. А. Котельникову $N_T = 2FT$).

Тогда $D = \frac{N_T}{T}$ число передаваемых отсчетов в единицу времени.

В этой модели могут быть определены следующие показатели:

- **Удельная скорость передачи информации**

$$RN = R_{\text{кан}} / D \text{ (бит / отсчет)}, \quad (1.7)$$

- **Удельная информационная емкость* канала**

$$C_N = \frac{C}{D} \text{ (бит / отсчет)}, \quad (1.8)$$

- **Удельная средняя энергия сигнала, затрачиваемая на передачу одного отсчета**

$$EN = P_c / D$$

- *Удельная эффективность использования пропускной способности канал*

$$\eta_N = \frac{R_N}{C_N} = \eta, \quad (1.9)$$

- *Удельная энергетическая эффективность (удельная эффективность использования энергии сигнала)*

$$\beta_N = \frac{R_N}{E_N / N_0} = \beta, \quad (1.10)$$

- *Удельная скорость передачи информации*

$$\gamma_N = R_N. \quad (1.11)$$

Переход к удельным показателям эффективности означает, что процесс формирования сигналов рассматривается в ортогональном базисе. При этом удельные показатели эффективности (1.9) и (1.10) точно совпадают с каноническими (1.4), (1.5), но удается избежать неопределенности при оценке полосы частот [1, 2].

В Главе 1 учебного пособия рассматриваются цифровые виды модуляции и сигнального кодирования, их спектральная и энергетическая эффективность. Приводятся краткие теоретические сведения и практическая их реализация на базе программного обеспечения NI LabVIEW. Даны результаты исследования вероятности ошибки на бит от отношения сигнал шум в канале для различных видов модуляции FSK, MSK и GMSK, а также QAM, M-QAM, PSK, M-PSK, созвездия для многопозиционных методов модуляции, спектры, глазковые диаграммы [11] и джиттер.

В Главе 2 рассматривается пропускная способность канала связи. Кодирование источника. Теорема К.Шеннона о кодировании в дискретном канале без помех. Проводится исследование кодирования источника дискретных сообщений методами Шеннона-Фано, исследование алгоритмов Лемпеля - Зива, Фрактальные методы кодирования изображений, а также вейвлет преобразования сигналов и изображений с использованием ПО MATLAB [13].

В Главе 3 исследуется помехоустойчивое кодирование в телекоммуникационных системах. Исследуются характеристики кодов Хемминга, БЧХ (Боуза-Чоудхури-Хоквенгема) и Рида-Соломона на базе MATLAB. Проведено исследование циклического избыточного кода (Cyclic redundancy check) CRC с использованием ПО MATLAB. Рассмотрены кодеры сверточных кодов и методы их декодирования, включая метод Витерби, исследованы их характеристики с использованием ПО MATLAB. Подробно исследовано турбокодирование,

обобщенная схема турбокодера с параллельным каскадированием. Сверточные турбокоды. Декодирование турбокодов. Оценены характеристики помехоустойчивости сверточных турбокодов. Исследование турбокодов проведено с использованием ПО MATLAB. Рассмотрены низкоплотностные коды, дана классификация LDPC-кодов, показаны методы построения проверочных матриц. Реализованы алгоритмы декодирования низкоплотных кодов, проведена оценка сложности алгоритмов декодирования на базе MATLAB и NI LabVIEW, исследованы каскадные коды на базе MATLAB Simulink 2015 [1, 3, 8-11, 16].

В Главе 4 рассмотрены сигнально-кодовые конструкции в телекоммуникационных системах, критерии эффективности систем передачи информации, информационная теория сигнально-кодовых конструкций, системы многочастотной OFDM модуляции - проведено имитационное моделирование в среде LabVIEW, реализованы Треллис кодовая модуляция (TCM) и проведен ее анализ с использованием MATLAB, рассмотрены методы пространственно-временного кодирования (MIMO) и оценена их эффективность на базе NI LabVIEW [12, 16].

В Главе 5 рассмотрены модемы и кодеки современных телекоммуникационных систем, имитационное моделирование системы мобильной связи стандартов GSM на базе ПО MATLAB, имитационное моделирование системы мобильной связи стандарта CDMA на базе ПО MATLAB, имитационное моделирование системы мобильной связи стандарта IEEE 802.11 (WiFi) на базе MATLAB 2015, имитационное моделирование системы мобильной связи стандарта IEEE 802.15.4 ZigBee на базе ПО MATLAB, имитационное моделирование системы мобильной связи стандарта IEEE 802.15.1 (Bluetooth) на базе ПО MATLAB, имитационное моделирование системы мобильной связи стандарта IEEE 802.16 (WiMAX) на базе ПО MATLAB [16, 18], имитационное моделирование системы мобильной связи стандарта IEEE 802.20 LTE на базе ПО MATLAB, имитационное моделирование системы цифрового наземного телевизионного вещания DVB-T [16], имитационное моделирование системы цифрового спутникового телевизионного вещания DVB-S и системы высокоскоростного цифрового спутникового ТВ-вещания DVB-S2 [16], имитационное моделирование системы цифрового кабельного телевизионного вещания DVB-C и системы высокоскоростного цифрового кабельного ТВ-вещания DVB-C2 [16], имитационное моделирование системы цифрового мобильного телевизионного вещания DVB-H и системы высокоскоростного цифрового мобильного ТВ-вещания DVB-H2 [16, 19, 20].

ГЛАВА 1. ЦИФРОВЫЕ ВИДЫ МОДУЛЯЦИИ И СИГНАЛЬНОГО КОДИРОВАНИЯ, ИХ СПЕКТРАЛЬНАЯ И ЭНЕРГЕТИЧЕСКАЯ ЭФФЕКТИВНОСТЬ

1.1. Анализ цифровых методов модуляции

Модулятор

Известно [6, 7], что узкополосный модулированный сигнал с произвольным видом модуляции можно представить в виде:

$$s(t) = I(t) \cos(\omega t) - Q(t) \sin(\omega t), \quad (1.12)$$

где ω - несущая частота радиосигнала, $I(t)$ и $Q(t)$ называются соответственно *синфазной* и *квадратурной* составляющими модулирующего сигнала.

Таким образом, для осуществления произвольного вида модуляции сигнала необходимо

выполнить две операции: (1) сформировать синфазную и квадратурную составляющие модулирующего сигнала (вид данных составляющих будет определять вид модуляции) и (2) выполнить преобразование (1.12).

Выполнение операций (1) и (2) выполняется различными блоками передающего тракта. Операция (1) осуществляется в baseband-процессоре, а операция (2) в квадратурном (IQ) модуляторе.

Baseband-модулятор формирует низкочастотные (baseband) сигналы $I(t)$ и $Q(t)$ из закодированного информационного сигнала (последовательности нулей и единиц). Закон, по которому выполняется данное преобразование, определяет вид модуляции сигнала.

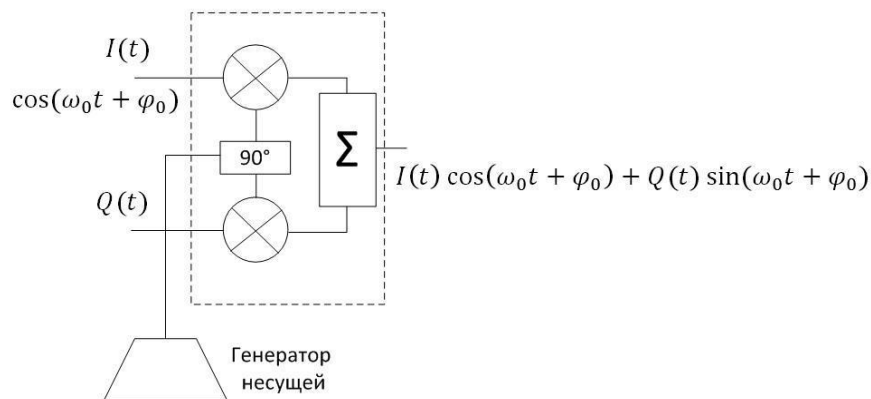


Рис. 1.3. IQ-модулятор

Для IQ-модулятора входными сигналами являются

$$I_{in}(t) = I(t); \quad Q_{in}(t) = Q(t) .$$

Сформированные baseband-процессором из исходного информационного сигнала и определяющие вид и свойства модулированного сигнала. На модулятор также поступает немодулированное несущее колебание вида $\cos(\omega t)$. На выходе модулятора образуется

сигнал (1.12) посредством умножения сигнала $I(t)$ на немодулированное несущее колебание, $Q(t)$ на немодулированное несущее колебание, сдвинутое по фазе на 90° , и последующего суммирования.

Видно, что структура IQ-модулятора является инвариантной относительно вида модуляции. Таким образом, осуществление того или иного вида модуляции определяется программой, выполняемой baseband-процессором, а именно алгоритмом формирования квадратурных составляющих из закодированного информационного сигнала. IQ-модулятор работает на высокой (несущей) частоте и, как правило, является аналоговым устройством. Таким образом, структурно передающий тракт можно разделить на цифровую и аналоговую части, разделенные цифро-аналоговым преобразователем (ЦАП), формирующим аналоговый сигнал из последовательности поступающих на него отсчетов.

Возможна также и иная реализация, когда IQ-модулятор выполнен в цифровом виде и преобразование (1.12) осуществляется baseband-процессором. В этом случае на выходе baseband-процессора формируется сигнал $s(t)$ на промежуточной частоте, существенно более низкой, чем несущая. Данный сигнал преобразуется в аналоговый с помощью ЦАП и затем его спектр переносится на несущую частоту с помощью смесителя.

Цифровые виды модуляции [7]

Цифровые виды модуляции (часто цифровая модуляция называется манипуляцией), как и аналоговые, могут быть амплитудными, фазовыми, частотными или комбинированными (например, амплитудно-фазовыми), в зависимости от того, какой из параметров немодулированного несущего колебания $s(t) = A(t)\cos(\omega t) + p(t)$ изменяется в соответствии с изменением информационного сигнала. Так как значения цифрового информационного сигнала являются дискретными (например, $\{0,1\}$), дискретным является также и возможный набор значений каждого из параметров. Однако если информационный сигнал проходит через baseband-фильтр для ограничения спектра, его значения уже не являются дискретными, поэтому реально переход от одного дискретного значения параметра колебания (например, изменение амплитуды или фазы) происходит гладко и непрерывно.

Амплитудные виды модуляции (OOK, ASK, M-ASK)

Наиболее простым видом манипуляции сигнала является амплитудная манипуляция. Модулированный сигнал имеет вид:

$$s(t) = A(c(t) + B)\cos(\omega t + \varphi_0),$$

где $c(t)$ - информационный цифровой сигнал, A , B и φ_0 - постоянные, $B \geq 0$, ω - несущая частота. Пусть множество возможных значений $c(t)$ $\{0,1\}$, $B=0$. В этом случае модулированный сигнал имеет вид $s(t) = Ac(t)\cos(\omega t + \varphi_0)$, его амплитуда принимает

значение 0 при нулевом значении информационного сигнала и A при единичном (рис.1.2). Такой тип манипуляции называется ООК (On-Off Keying, Включено-Выключено) и часто используется в системах сигнализации и охранных системах.

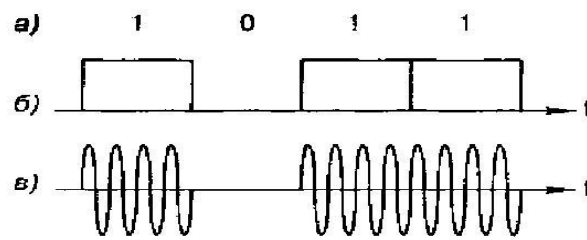


Рис. 1.4. Модуляция ООК: а - информационное сообщение; б - модулирующий цифровой сигнал; в - модулированный радиосигнал

Допустим теперь $B = 1$. В этом случае амплитуда модулированного сигнала принимает значение A при нулевом значении информационного сигнала и $2A$ при единичном. Вид модуляции, для которого $B > 0$, носит название ASK (Amplitude Shift Keying - амплитудная манипуляция). ООК является частным случаем ASK при $B=0$. Существует два основных критерия сравнения эффективности различных видов модуляции. Это *критерии спектральной и энергетической эффективности*. Спектральная эффективность характеризует полосу частот, необходимую для передачи информации с определенной скоростью. Энергетическая эффективность описывает мощность, необходимую для передачи информации с заданной достоверностью (вероятностью ошибки). Известно, что спектр модулированного сигнала на радиочастоте с точностью до постоянного множителя совпадает со спектром модулирующего (baseband) сигнала, однако, центр спектра радиосигнала размещен на несущей частоте, а не на нулевой. Поэтому, как правило, анализируются спектральные плотности модулирующих сигналов, центрированные относительно нулевой частоты. Спектральные плотности мощности ASK сигналов для различных baseband-фильтров приведены на рис.1.4. На рис.1.4 показаны соответствующие формы импульсов модулирующего сигнала после прохождения baseband-фильтра. Из сравнения рис.1.4 и рис.1.5 видно, что более гладкая форма импульса модулирующего сигнала приводит к расширению главного лепестка спектральной плотности мощности модулированного сигнала и более быстрому уменьшению амплитуды боковых лепестков.

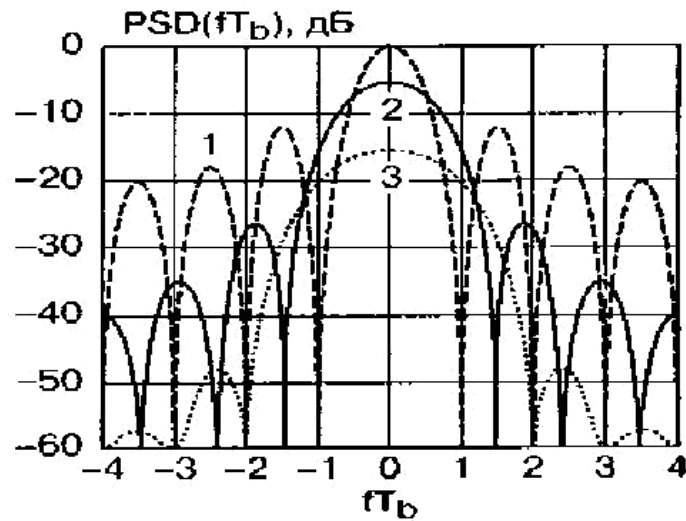


Рис. 1.5. Спектральная плотность мощности ASK-сигнала. Форма импульса модулирующего сигнала: 1 - прямоугольная, 2 - косинусоидальная, 3 - приподнятый косинус.

Множество возможных значений квадратурных компонент $I(t)$ и $Q(t)$ называется *сигнальным созвездием*. Как правило, данное множество отображают на декартовой плоскости, где по оси абсцисс отложены значения синфазной составляющей $I(t)$, а по оси ординат - квадратурной $Q(t)$. Точка на плоскости с координатами (x, y) соответствует состоянию сигнала, в котором синфазная составляющая равна x , квадратурная равна y . Таким образом, сигнальное созвездие - это диаграмма возможных состояний сигнала. Используя общий вид модулированного радиосигнала, можно показать, что амплитуда модулированного радиосигнала в текущем состоянии равна

$$A(t) = \sqrt{I^2(t) + Q^2(t)},$$

а фаза равна углу вектора, указывающего в точку (I, Q) , отсчитываемого от оси абсцисс. Для модуляций ООК и ASK сигнальное созвездие изображено на рисунке 1.6.

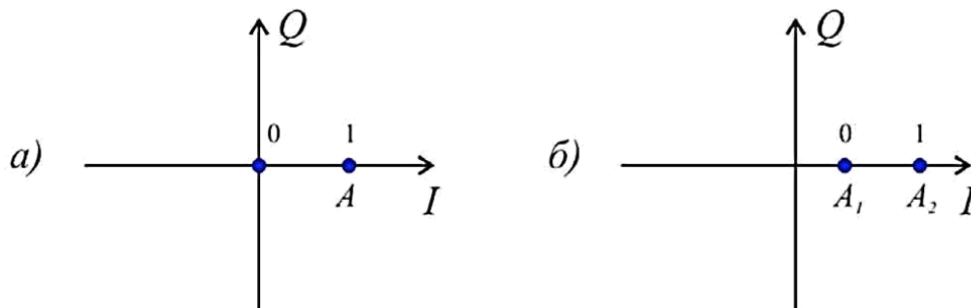


Рис. 1.6. Виды модуляции: а - сигнальное созвездие модуляции ООК, б - сигнальное созвездие модуляции ASK

Многопозиционная амплитудная модуляция (М-ASK)

При модуляции ASK множество возможных значений амплитуды радиосигнала

ограничивается двумя значениями (без учета сглаживания baseband-фильтром). Спектральная эффективность может быть существенно повышена, если использовать большее количество значений амплитуды радиосигнала.

Сгруппируем биты исходного информационного сообщения в пары. Каждая такая пара называется символом. Если каждый бит имеет множество значений $\{0,1\}$, то каждый символ имеет четыре возможных значения из множества $\{00, 01, 10, 11\}$. Сопоставим каждому из возможных значений символа значение амплитуды радиосигнала из множества $\{0, A, 2A, 3A\}$. Аналогичным образом можно группировать тройки, четверки и большее количество бит в одном символе. Получится многоуровневый (многопозиционный) сигнал M-ASK с размерностью множества возможных значений амплитуды сигнала $M = \log_2 k$, где k - число бит в одном символе.

Например, сигнал с модуляцией 256-ASK имеет 256 возможных значений амплитуды сигнала и 8 бит в одном символе.

На рисунке 1.7 изображена спектральная плотность мощности восьмиуровневого сигнала 8-ASK и спектральная плотность сигнала ASK с импульсами прямоугольной формы (без baseband-фильтрации). Многопозиционный сигнал имеет меньшую ширину главного лепестка (занимает меньшую полосу частот) и более низкий уровень боковых лепестков, т. е. имеет большую спектральную эффективность по сравнению с двухуровневым сигналом.

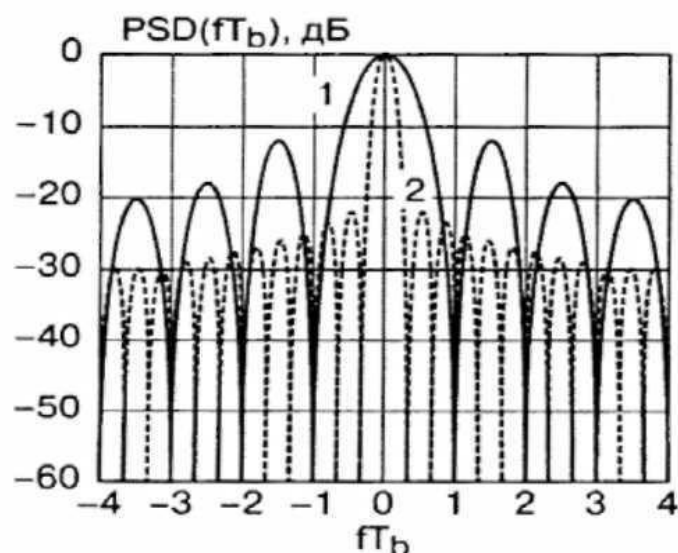


Рис. 1.7. Сравнение спектральных плотностей мощности двухуровневого и восьмиуровневого AM-сигналов: 1 - сигнал ASK, 2 – сигнал 8-ASK

Амплитудные виды модуляции имеют невысокую энергетическую эффективность (так как средний уровень мощности существенно меньше максимального), требуют высокой линейности и большого динамического диапазона усилителя мощности. Ошибка в амплитуде сигнала из-за нелинейности усилителя приведет непосредственно к символьной

ошибке, т. к. значение символа определяется амплитудой сигнала. Отношение максимальной амплитуды сигнала к минимальной достаточно высоко и требует усилителя с большим динамическим диапазоном. Влияние аддитивного шума или помехи непосредственно изменяет амплитуду сигнала, поэтому амплитудные виды модуляции не обладают высокой помехоустойчивостью. Однако они достаточно просты в реализации. Сигнальное созвездие для 8-ASK приведено на рисунок 1.8.

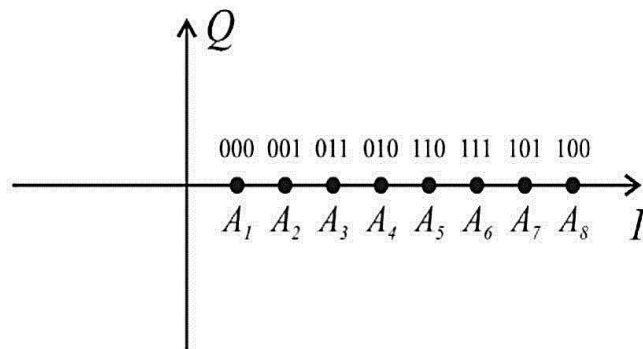


Рис. 1.8. Сигнальное созвездие модуляции 8-ASK

Фазовые виды модуляции (BPSK, QPSK, M-PSK, QAM)

Фазомодулированный сигнал имеет вид:

$$s(t) = A \cos(\omega t + \varphi(t) + \varphi_0), \quad (1.13)$$

где A и φ_0 - постоянные, ω_0 - несущая частота.

Информация кодируется фазой $\varphi(t)$. Так как при когерентной демодуляции в приемнике имеется восстановленная несущая $S_C(t) = A \cos(\omega t + \varphi_0)$, то путем сравнения сигнала с несущей вычисляется текущий сдвиг фазы $\varphi(t)$. Изменение фазы $\varphi(t)$ взаимнооднозначно связано с информационным сигналом $c(t)$.

QPSK (Quadrature Phase Shift Keying) - Квадратурная фазовая модуляция (четырёхуровневая фазовая модуляция). В данной фазовой модуляции используются четыре значения фазы несущего колебания. В этом случае фаза сигнала должна принимать четыре значения: 0° , 90° , 180° и 270° . Однако чаще используются другие значения фаз: 45° , 135° , 225° и 315° с шагом, кратным $\pi/2$.

Соотношение между сдвигом фазы модулированного колебания из множества $\{\pm\pi/4, \pm3\pi/4\}$ и множеством символов (дибитов) цифрового сообщения $\{00, 01, 10, 11\}$ устанавливается в каждом конкретном случае стандартом на радиоканал и отображается сигнальным созвездием. Сигнальное созвездие – это представление манипулированных радиосигналов на комплексной плоскости.

Модулированный сигнал можно представить как вектор на графике в полярной системе координат; длина вектора соответствует амплитуде сигнала, а его ориентация — фазе. На рисунке 1.9 представлено сигнальное созвездие для QPSK модуляции.

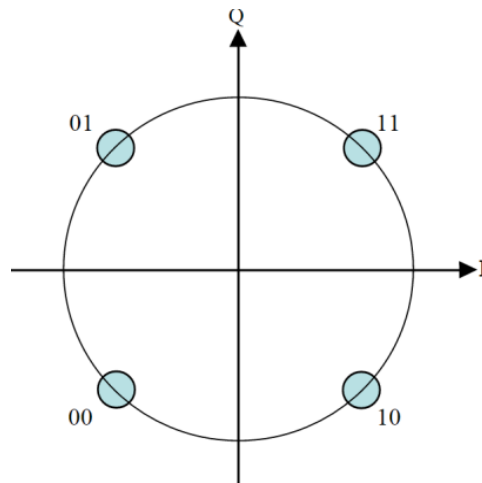


Рис. 1.9. Сигнальное созвездие для QPSK модуляции.

Фаза может переходить из одного состояния в другое. Такие переходы представлены на рисунке 1.10 чёрными стрелками.

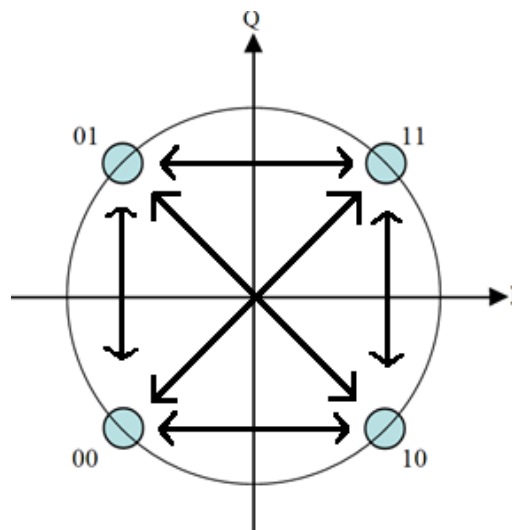


Рис. 1.10. Переходы из одного фазового состояния в другое

Из теории связи известно, что наивысшей помехоустойчивостью обладает двоичная фазовая модуляция BPSK. Однако в ряде случаев за счет уменьшения помехоустойчивости канала связи можно увеличить его пропускную способность.

Квадратурно-амплитудная модуляция - QAM (QAM – Quadrature Amplitude Modulation) может рассматриваться как расширенная многоуровневая ФМ, в которой два исходных сигнала генерируются независимо. Таким образом, здесь имеют место два полностью независимых квадратурных канала, включающие процессы кодирования и детектирования в основной полосе.

На рисунке 1.11 показано сигнально - точечное пространство для системы с 16-QAM и четырьмя уровнями в каждом квадратурном канале. Точки представляют составной сигнал, а штрихи на осях отмечают уровни амплитуды в каждом квадратурном канале.

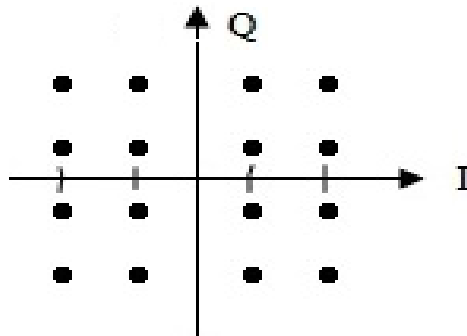


Рис. 1.11. Сигнально-точечное пространство модуляции для 16-QAM

В отличие от ФМ сигналов сигналы QAM не содержат постоянной огибающей. Наличие постоянной огибающей в ФМ объясняется поддержанием отношения уровней в квадратурных каналах. В QAM такие ограничения не вводятся ввиду того, что в каждом канале уровни независимы.

Характеристики ошибок систем QAM и ФМ модуляций сильно отличаются. При достаточно большом числе сигнальных точек системы QAM имеют, как правило, лучшие характеристики, чем системы с ФМ. Основная причина состоит в том, что расстояние между сигнальными точками на диаграмме для системы с QAM больше, чем для соответствующей системы с ФМ.

QAM имеет преимущество над системой ФМ при той же пиковой мощности.

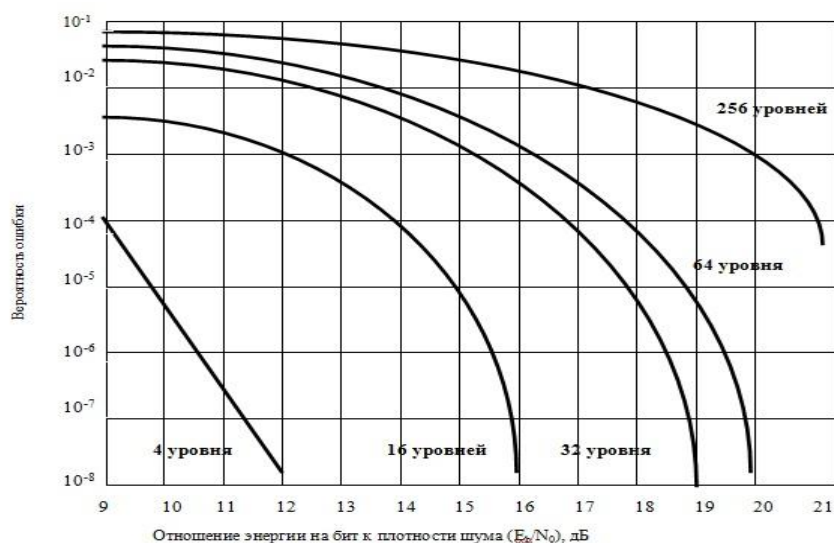


Рис. 1.12. Вероятности ошибок в системах QAM

Достоинство высоких значений номера QAM – это повышенная скорость передачи данных, поскольку таким образом большее количество битов информации может быть

передано в течении одного цикла. Однако, с другой стороны, в этом случае большее число уровней амплитуды сигнала располагаются близко друг к другу, повышая тем самым вероятность неразличимости двух уровней, и как следствие – повышая чувствительность системы к шуму.

На рисунке 1.13 представлено отношение параметра SNR к параметру BER (Bit Error Rate – Отношение Бит/Ошибка).

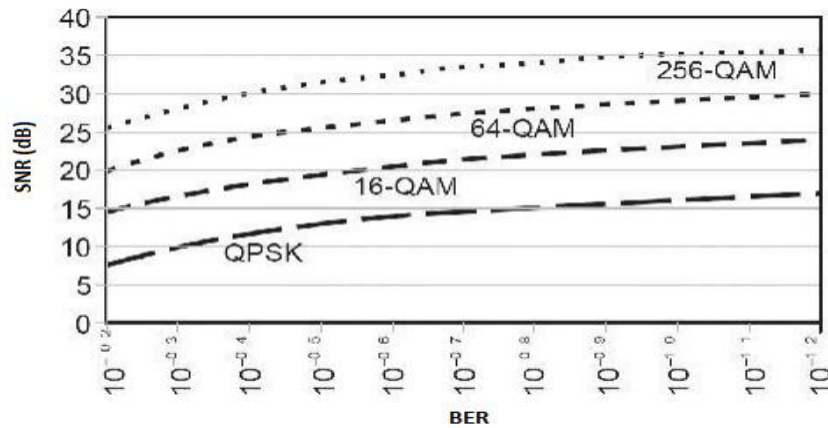


Рис. 1.13. Отношение параметра SNR к параметру BER

Двоичная фазовая модуляция (BPSK - Binary Phase Shift Keying)

Множеству значений информационного сигнала $\{0,1\}$ ставится в однозначное соответствие множество изменений фазы $\{0, \pi\}$. При изменении значения информационного сигнала фаза радиосигнала изменяется на 180° . Временная форма сигнала и его созвездие показаны на рисунке 1.14.

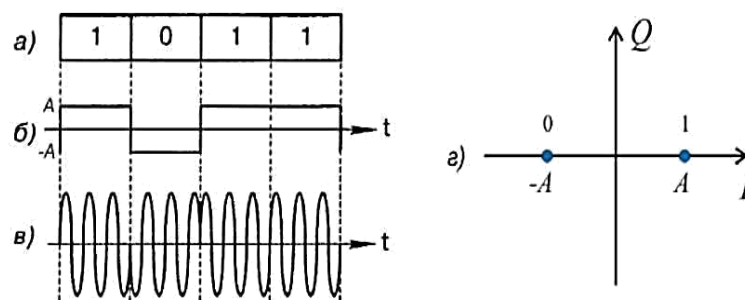


Рис. 1.14. Временная форма и сигнальное созвездие сигнала BPSK: а – цифровое сообщение; б - модулирующий сигнал; в - модулированное ВЧ-колебание; г - сигнальное созвездие

Квадратурная фазовая модуляция (QPSK - Quadrature Phase Shift Keying)

Квадратурная фазовая модуляция является четырехуровневой фазовой модуляцией ($M=4$), при которой фаза высокочастотного колебания может принимать 4 различных

значения с шагом, кратным $\pi/2$.

Соотношение между сдвигом фазы модулированного колебания из множества $\{\pm\pi/4, \pm 3\pi/4\}$ и множеством символов (дибитов) цифрового сообщения $\{00, 01, 10, 11\}$ устанавливается в каждом конкретном случае стандартом на радиоканал и отображается сигнальным созвездием, аналогичным рисунку 1.15. Стрелками показаны возможные переходы из одного фазового состояния в другое.

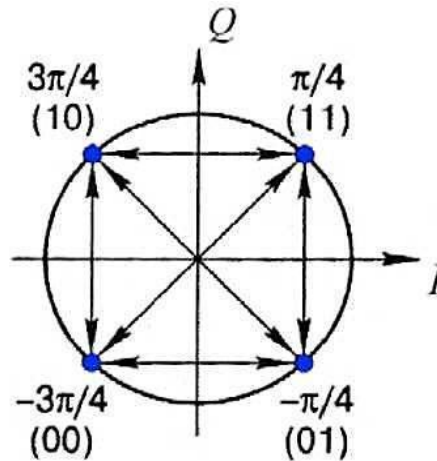


Рис. 1.15. Сигнальное созвездие модуляции QPSK

Из рисунка видно, что соответствие между значениями символов и фазой сигнала установлено таким образом, что в соседних точках сигнального созвездия значения соответствующих символов отличаются лишь в одном бите.

При передаче в условиях шума наиболее вероятной ошибкой будет определение фазы соседней точки созвездия. При указанном кодировании, несмотря на то, что произошла ошибка в определении значения символа, это будет соответствовать ошибке в одном (а не двух) бите информации. Таким образом, достигается снижение вероятности ошибки на бит. Указанный способ кодирования называется кодом Грея.

Каждому значению фазы модулированного сигнала соответствует 2 бита информации, и поэтому изменение модулирующего сигнала при QPSK - модуляции происходит в 2 раза реже, чем при BPSK-модуляции при одинаковой скорости передачи информации. Известно [5], что спектральная плотность мощности многоуровневого сигнала совпадает со спектральной плотностью мощности бинарного сигнала при замене битового интервала T_b на символьный интервал $T_s = T_b \log_2 M$. Для четырехуровневой модуляции $M=4$ и, следовательно, $T_s = 2T_b$. Расстояние между первыми нулями спектральной плотности мощности сигнала QPSK равно $Af = 1/T_b$, что в 2 раза меньше, чем для сигнала BPSK. Другими словами, спектральная эффективность квадратурной модуляции QPSK в 2 раза выше, чем бинарной модуляции BPSK.

Многопозиционная фазовая модуляция (M-PSK)

M-PSK формируется, как и другие многопозиционные виды модуляции, путем группировки $k = \log_2 M$ бит в символы и введением взаимно-однозначного соответствия между множеством значений символа и множеством значений сдвига фазы модулированного колебания. Значения сдвига фазы из множества отличаются на одинаковую величину. Для примера на рисунке 1.16 приведено сигнальное созвездие для 8-PSK с кодированием Грея.

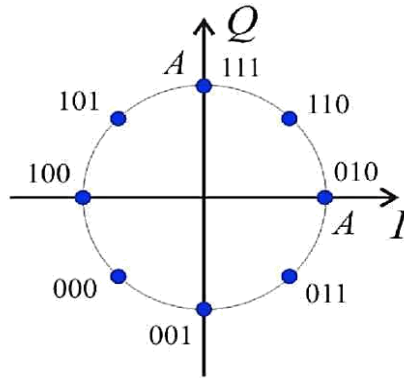


Рис. 1.16. Сигнальное созвездие модуляции 8-PSK

Амплитудно-фазовые виды модуляции (QAM)

Модуляция, при которой происходит одновременное изменение двух параметров несущего колебания - амплитуды и фазы - называется амплитудно-фазовой модуляцией.

Минимальный уровень символьных ошибок будет достигнут в случае, если расстояние между соседними точками в сигнальном созвездии будет одинаковым, т. е. распределение точек в созвездии будет равномерным на плоскости. Следовательно, сигнальное созвездие должно иметь решетчатый вид. Модуляция с подобным видом сигнального созвездия называется квадратурной амплитудной модуляцией (QAM - Quadrature Amplitude Modulation). QAM является многопозиционной модуляцией. При $M=4$ она соответствует QPSK, поэтому формально считается для QAM $M > 8$ (т.к. число бит на символ $k = \log_2 M$, $k \in \mathbb{N}$, то M может принимать только значения степеней 2: 2, 4, 8, 16 и т.д.). Для примера на рисунке 1.17 приведено сигнальное созвездие 16-QAM с кодированием Грея.

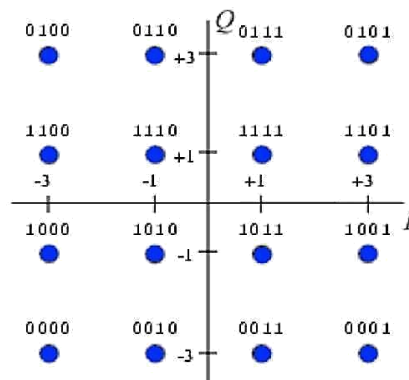


Рис. 1.17. Сигнальное созвездие модуляции 16-QAM

На практике используются большие значения M , вплоть до 1024-QAM. Такие виды модуляции позволяют достичь исключительно высокой спектральной эффективности. Однако, как видно из сигнального созвездия, так как информация кодируется в том числе амплитудой и изменения амплитуды велики, то QAM предъявляет высокие требования к линейности усилителя мощности и его динамическому диапазону, особенно для больших M . Практическое осуществление QAM-модуляции выполняется следующим образом. В памяти процессора хранится таблица значений квадратурных компонент $I(t)$ и $Q(t)$, имеющих в сигнальном созвездии и расположенных в порядке возрастания значения соответствующего символа. Процессор анализирует входную последовательность битов, разбивает ее на символы и для каждого символа выбирает соответствующие значения квадратурных компонент из таблицы. Затем выполняется baseband-фильтрация сигналов $I(t)$ и $Q(t)$.

Сравнение различных видов модуляции

Как указывалось, основными критериями эффективности различных видов модуляции являются критерии спектральной и энергетической эффективности. Энергетическая эффективность характеризует энергию, которую необходимо затратить для передачи информации с заданной достоверностью (вероятностью ошибки). Спектральная эффективность характеризует полосу частот, необходимую для того, чтобы передавать информацию с определенной скоростью. Кроме данных критериев, виды модуляции сравниваются по устойчивости к различным типам помех и искажений и сложности аппаратной реализации. Увеличение позиций (уровней) модуляции (модуляции M-ASK, M-PSK и M-QAM) увеличивает спектральную эффективность в $k = \log_2 M$ раз. Также отмечено, что наибольшей спектральной эффективностью среди частотных видов модуляции обладает модуляция MSK. MSK является спектрально в 2.6 раза менее эффективной, чем QPSK и в 1.3 раза менее эффективной, чем BPSK. Сравним виды модуляции по критерию энергетической эффективности. Для этого оценим для каждого вида модуляции требуемую энергию для передачи информации с одинаковой вероятностью ошибки на бит. В таблице приводятся зависимости вероятности ошибки на бит от отношения E_b / N_0 для различных видов модуляции.

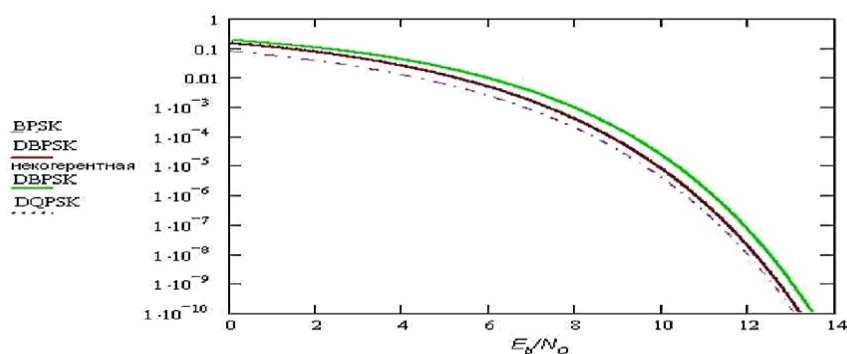


Рис. 1.18. Сравнение энергетической эффективности модуляций с относительным кодированием: DBPSK, DQPSK и некогерентной DBPSK.

Однако BER для BPSK и QPSK описываются одинаковыми формулами, при этом QPSK в 2 раза спектрально эффективнее, чем BPSK. Следовательно, QPSK всегда существенно эффективнее, чем BPSK, и, обыкновенно, имеет смысл использовать QPSK, а не BPSK.

Сравним виды модуляции по критерию энергетической эффективности. Для этого оценим для каждого вида модуляции требуемую энергию для передачи информации с одинаковой вероятностью ошибки на бит. В [2], [10] определены соотношения, связывающие вероятность битовой ошибки с величиной E_b / N_0 для различных видов модуляции:

$$BER = f\left(\frac{E_b}{N_0}\right), \quad (1.14)$$

где BER - вероятность ошибки E_b - энергия, необходимая для передачи одного бита информации, N_0 - спектральная плотность мощности белого шума в канале. Если мощность передатчика равна P , то величина энергии, приходящаяся на один бит информации, равна $E_b = PT_b$, где T_b - длительность бита. В табл. 1.1 приводятся зависимости вероятности ошибки на бит от отношения E_b / N_0 для различных видов модуляции [1].

Таблица 1.1. Вероятность ошибки для различных видов модуляции

Вид модуляции	Вероятность ошибки на бит (BER)
ОСК	$Q(\sqrt{E_b / N_0})$
M-ASK код Грея	$\frac{2(M-1)}{M \log_2 M} Q\left(\sqrt{\frac{\log_2 M}{(M-1)^2} \frac{E_b}{N_0}}\right)$
BPSK	$Q(\sqrt{2E_b / N_0})$
Некогерентная DBPSK	$\frac{1}{2} \exp(-E_b / N_0)$
Когерентная DBPSK	$2Q\left(\sqrt{\frac{2E_b}{N_0}}\right)\left(1 - Q\left(\sqrt{\frac{2E_b}{N_0}}\right)\right)$
QPSK код Грея	$Q(\sqrt{2E_b / N_0})$
Когерентная DQPSK при $E_b / N_0 \gg 1$	$2Q(\sqrt{2E_b / N_0})$

М-PSK код Грея	$\frac{2}{\log_2 M} Q\left(\sqrt{\frac{2E_b \log_2 M}{N_0} \sin^2\left(\frac{\pi}{M}\right)}\right)$
FSK	$Q\left(\sqrt{\left[1 - \frac{\sin(2\pi \cdot m)}{2\pi \cdot m}\right] \frac{E_b}{N_0}}\right)$
MSK	$Q(\sqrt{E_b / N_0})$
М-MSK	$\frac{2(M-1)}{M \log_2 M} Q\left\{\sqrt{\log_2 M \frac{E_b}{N_0}}\right\}$
QAM код Грея	<p>для $k = \log_2 M$, k – четное:</p> $BER = \frac{2P_0 - P_0^2}{\log_2 M}, \text{ где}$ $P_0 = \frac{2(\sqrt{M} - 1)}{\sqrt{M}} Q\left(\sqrt{\frac{3 \log_2 M}{M-1} \frac{E_b}{N_0}}\right)$ <p>для нечетных k:</p> $BER \leq \frac{1}{\log_2 M} \left[1 - \left(1 - 2Q\left(\sqrt{\frac{3 \log_2 M}{M-1} \frac{E_b}{N_0}}\right) \right)^2 \right]$

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-t^2/2} dt$$

В табл. - интеграл ошибок (интеграл вероятности),

M - число позиций для многопозиционных видов модуляции, m - индекс модуляции для частотной модуляции, BER - вероятность ошибки на бит.

Из табл. видно, что с увеличением позиционности модуляции, вероятность битовой ошибки увеличивается (см., например, формулы М-ASK и М-PSK, $Q(x)$ является убывающей функцией аргумента). Таким образом, как правило, при увеличении спектральной эффективности энергетическая эффективность уменьшается.

Однако BER для BPSK и QPSK описываются одинаковыми формулами, при этом QPSK в 2 раза спектрально эффективнее, чем BPSK. Следовательно, QPSK всегда существенно эффективнее, чем BPSK, и, обыкновенно, имеет смысл использовать QPSK, а не BPSK. Физически, это объясняется тем, что в случае QPSK добавляется дополнительная степень свободы: квадратурная составляющая $Q(t)$. В случае BPSK используется только синфазная составляющая $I(t)$. Квадратурная форма когерентного фазового демодулятора приводит к тому, что два канала детектора обеспечивают независимый прием двух бинарных фазомодулированных сигналов. Аналогичное явление имеет место и при сравнении модуляций DBPSK и DQPSK (с относительным кодированием). Хотя выражения для BER несколько отличаются, с высокой степенью приближения они совпадают (рис.24). Модуляции с относительным кодированием имеют небольшой энергетический проигрыш по

сравнению с обыкновенными BPSK и QPSK (0.3 - 0.9 дБ). DBPSK с некогерентным детектированием также имеет небольшой проигрыш по сравнению с DBPSK с когерентным детектированием (около 0.5 дБ), рис. 1.20. Под энергетическим выигрышем понимается разница в значениях $E_b' N_0$ при одинаковом значении вероятности ошибки на бит.

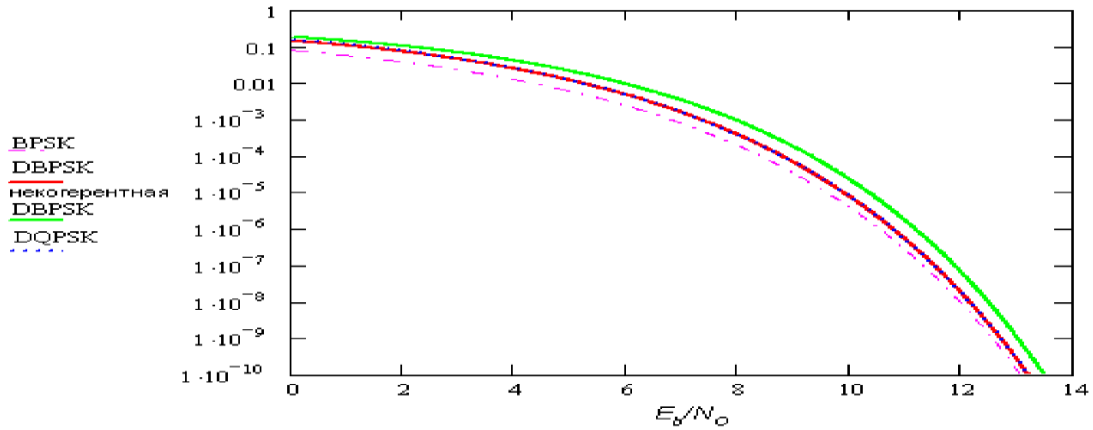


Рис. 1.20. Сравнение энергетической эффективности модуляций BPSK, DBPSK, DBPSK и DQPSK

Таким образом, имеет смысл сравнивать виды модуляции с одинаковым числом позиций. Сравним двухуровневые OOK, BPSK и MSK. Как видно из табл., OOK и MSK имеют одинаковую эффективность и уступают BPSK (и, соответственно, QPSK) по энергетической эффективности приблизительно 3 дБ (рис.1.21).

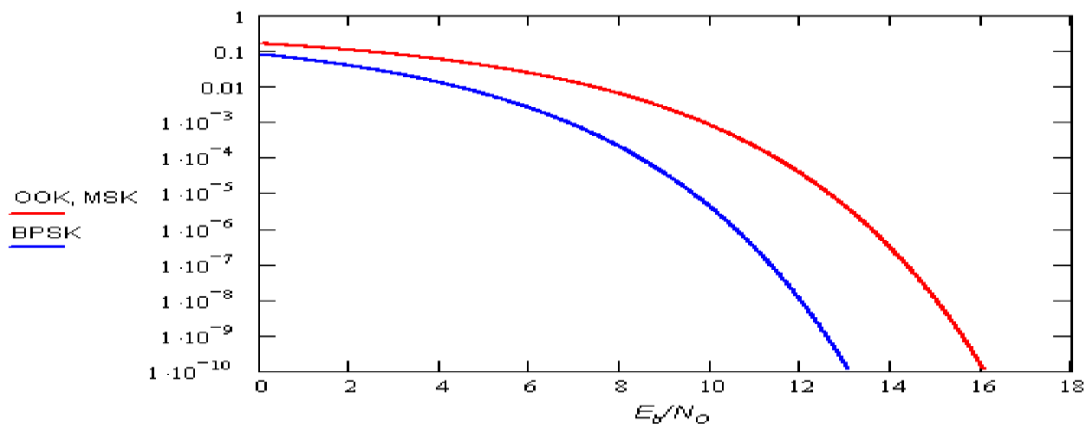


Рис. 1.21. Сравнение энергетической эффективности модуляций OOK, MSK и BPSK

По результатам данного сравнения можно сделать вывод о том, что при числе уровней до 4 включительно QPSK является спектрально и энергетически наиболее эффективным видом модуляции. Однако здесь следует сделать одно существенное замечание относительно модуляции GMSK. Ее спектральная эффективность ниже, чем QPSK, в системах с *линейным усилением*. GMSK, как частотный вид модуляции, позволяет использовать высокоэффективные нелинейные усилители и ограничители, что дает энергетический выигрыш. При прохождении QPSK через подобные устройства, ее спектр расширяется

(происходит некоторое восстановление боковых лепестков). Поэтому, в некоторых случаях, GMSK может иметь большую эффективность, чем QPSK. В частности в стандарте GSM выбор сделан в пользу GMSK, а в CDMA - OQPSK. Однако, усовершенствованные виды модуляции QPSK (например, FQPSK) в любом случае превосходят GMSK.

Сравним теперь модуляции с числом уровней $M > 4$. На рис. 1.22 изображено сравнение энергетической эффективности для амплитудной, фазовой и амплитудно-фазовой манипуляции при $M=16$ и $M=64$.

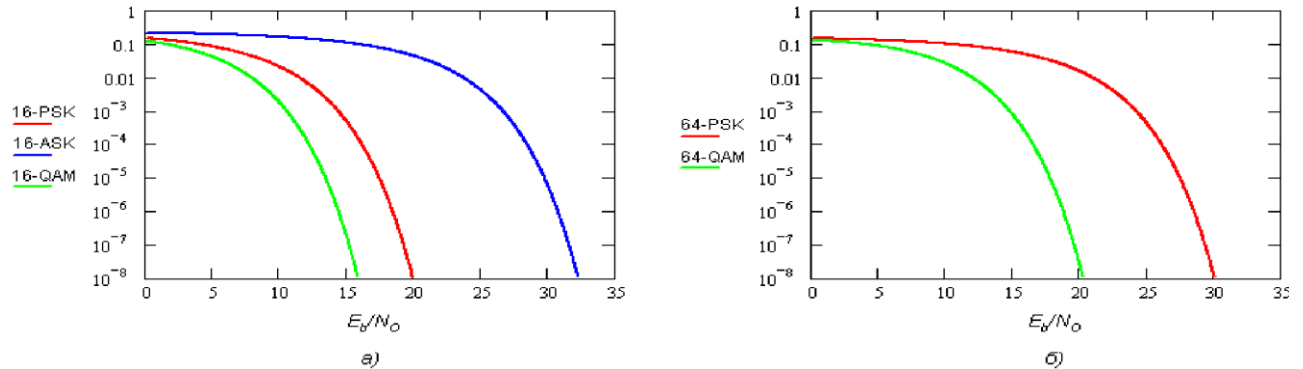


Рис. 1.22. Сравнение энергетической эффективности модуляций M-ASK, M-PSK и M-QAM: а - M=16, б - M=64

Как видно из рис. 1.22, амплитудная модуляция существенно (более 10 дБ при $M=16$) уступает фазовой и амплитудно-фазовой, поэтому при $M=64$ сравнение с ней не проводится. При сравнении M-PSK с M-QAM видно, что M-QAM превосходит по эффективности M-PSK, причем энергетический выигрыш M-QAM увеличивается с ростом M . Например, для $M=16$ выигрыш составляет около 4 дБ, а при $M=64$ около 10 дБ. Физически это объясняется тем, что расстояние между соседними точками в сигнальном созвездии M-PSK меньше, чем M-QAM. Сигнальное созвездие M-PSK представляет собой окружность с равномерно распределенными на ней точками, а созвездие M-QAM - квадрат с равномерно распределенными по его площади точками. Чем больше расстояние между точками в созвездии, тем менее вероятна ошибка в детектировании соседнего символа. Многопозиционная частотная модуляция используется гораздо реже, так как *при увеличении числа уровней и сохранении индекса модуляции ее спектр не сужается, а расширяется*, ввиду того, что вводятся новые частоты и ширина спектра растет по закону $\frac{M}{\log_2 M}$, Как видно из табл. 1.1, однако, *при увеличении числа уровней M-MSK, в отличие от всех других видов модуляции, вероятность ошибки на бит уменьшается*. Мы получаем выигрыш в энергетической эффективности за счет уменьшения спектральной эффективности.

Таким образом, при ограниченной полосе, при $M < 4$ наиболее эффективной является модуляция QPSK, а при $M > 4$ - QAM. QPSK является частным случаем QAM при $M=4$. Можно считать QAM наиболее эффективным видом модуляции при любом числе уровней. Еще больший выигрыш по сравнению с обыкновенными QPSK и QAM дают их усовершенствованные модификации, такие, как модификации Феера [11] (FQPSK, FQAM), модуляция с решетчатым кодированием (TCM), оптимизация формы сигнальных созвездий и использование многомерных сигнальных созвездий.

1.2. Модемы сотовой связи FSK, MSK GMSK и численный анализ вероятности символьной ошибки с использованием ПО LabVIEW

Частотная манипуляция (FSK)

Значениям «0» и «1» информационной последовательности соответствуют определённые частоты синусоидального сигнала при неизменной амплитуде. Частотная манипуляция весьма помехоустойчива, поскольку помехи телефонного канала искажают в основном амплитуду, а не частоту сигнала. Однако при частотной манипуляции неэкономно расходуется ресурс полосы частот телефонного канала. Поэтому этот вид модуляции применяется в низкоскоростных протоколах, позволяющих осуществлять связь по каналам с низким отношением сигнал/шум.

Существует также подвид этой модуляции GFSK. Принцип работы модулятора GFSK похож на FSK, за исключением того, что сначала полоса импульсов (-1, 1) проходит через фильтр Гаусса для сглаживания, что обеспечивает уменьшения ширины его спектра, а уже после попадает в FSK. Фильтрация Гаусса — один из самых распространенных способов уменьшения ширины спектра.

На рисунке ниже приведен график двоичной бинарной последовательности нулей и единиц и, соответствующий ему, график частотно-манипулированного сигнала. Низкому уровню бинарного двоичного сигнала соответствует частота 1 КГц, а высокому - частота 0,5 КГц несущего сигнала синусоидального типа.

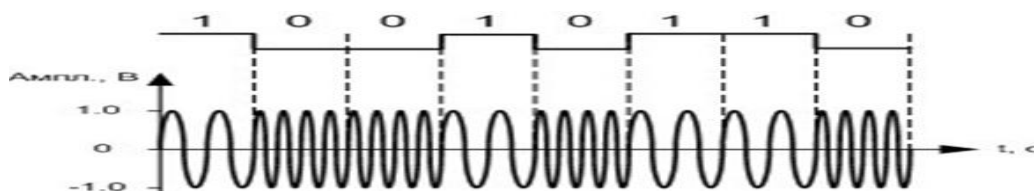


Рис. 1.23. Частотная манипуляция

В программной комплексе можно осуществить многопозиционную частотную модуляцию (MFSK) задав в поле M-FSK нужный уровень от 2 до 64. В бинарной FSK модуляции, т.е. при $M=2$, два двоичных числа представляются сигналами двух различных

частот, расположенных около несущей. Бинарная частотная модуляция менее восприимчива к ошибкам, чем амплитудная модуляция.

Более эффективной, но и более подверженной ошибкам, является схема многочастотной модуляции (Multiple FSK - MFSK), в которой используется более двух частот. В этом случае каждая сигнальная посылка представляет более одного бита. Переданный сигнал MFSK (для одного периода передачи сигнальной посылки) можно определить следующим образом:

$$s_i = A \cos(2\pi f_i t), \quad 1 \ll i \ll M$$

Здесь

$$f_i = f_c + (2i - 1 - M)f_d,$$

где f_c - несущая частота; f_d - разностная частота; M - число различных сигнальных

посылок $= 2^L$; L - количество битов на одну сигнальную посылку.

На рис. 1.24 представлен пример схемы MFSK с $M=4$. Входной поток битов кодируется по два бита, после чего передается одна из четырех возможных двухбитовых комбинаций.

Для уменьшения занимаемой полосы частот в модуляторах сигналов с фазовой модуляцией применяют сглаживающие фильтры. Применение сглаживающих фильтров приводит к увеличению эффективности использования полосы, но в то же время из-за сглаживания уменьшается расстояние между соседними сигналами, что приводит к снижению помехоустойчивости.



Рис. 1.24. Использование частоты схемой MFSK ($M = 4$)

Построение графиков глазковой диаграммы и ее вид дает много информации о джиттере сигнала, так же, как и о многих других его параметрах. Например, множество отдельных фронтов и спадов говорит о вероятном присутствии джиттера, зависящего от данных.

Глазковая диаграмма не просто предоставляет множество информации, она удобна простотой применения и тем, что может применяться для измерений в любой цепи с реальными данными.



Рис. 1.25. Параметры глазковой диаграммы

По индикаторной диаграмме можно выполнить ряд важных измерений:

- чем больше открыт глазок, тем легче различать логические 1 и 0;
- ширина открытия глазковой диаграммы (время между пересечениями линий логической 1 с логическим 0 и логического 0, с логической 1) показывает временной интервал, в течение которого сигнал может быть замерен без ошибки из-за межсимвольного влияния;
- псевдослучайной последовательности битов и отображения сигналов на запоминающем осциллографе получается структура, которая называется индикаторной (глазковой) диаграммой (eye diagram). Типичная индикаторная диаграмма приведена на рис. 3.
- высота открытия глазка измеряет запас помехоустойчивости на выходе приемника;
- ширина линий глазка к точкам пересечения в углах глазка является мерой флуктуации в системе передачи. Флуктуации вызываются разбросом времени включения и выключения лазера; искажением импульса оптическим волокном и шумом. Флуктуации выражаются в пикосекундах, градусах или в процентах интервала бита;
- толщина линий импульса наверху и внизу глазка пропорциональна шуму и искажениям в системе передачи; время перехода сигнала в схеме глазка с верхнего уровня (логического 0) в нижний (логическая 1) и наоборот указывает времена подъема и спада системы передачи. Они обычно замеряются между отметками 10 и 90%;

- времена подъема и спада важны для оценки чувствительности системы к синхронизации замеров (sample timing). Чем больше времена подъема и спада сигнала, тем более чувствительна система к ошибкам синхронизации;

- чтобы обеспечить Системе максимальную невосприимчивость к шуму, лучшим временем для замеров уровня сигнала является время, когда высота открытия индикаторной диаграммы максимальна.

Джиттер или фазовое дрожание цифрового сигнала данных — нежелательные фазовые и/или частотные случайные отклонения передаваемого сигнала. Возникают вследствие нестабильности задающего генератора, изменений параметров линии передачи во времени и различной скорости распространения частотных составляющих одного и того же сигнала.

В цифровых системах проявляется в виде случайных быстрых изменений местоположения фронтов цифрового сигнала во времени, что приводит к рассинхронизации и, как следствие, искажению передаваемой информации. Например, если фронт имеет малую крутизну или «отстал» по времени, то цифровой сигнал как бы запаздывает, сдвигается относительно значащего момента времени — момента времени, в который происходит оценка сигнала.

Джиттер является одной из основных проблем при проектировании устройств цифровой электроники, в частности, цифровых интерфейсов. Недостаточно аккуратный расчет джиттера может привести к его накоплению при прохождении цифрового сигнала по тракту и, в конечном счёте, к неработоспособности устройства.

Причины возникновения джиттера:

- Фазовые шумы петли ФАПЧ (фазовой автоподстройки частоты) устройства, синхронизируемого внешним сигналом. Джиттер, вызываемый ФАПЧ, проявляется при прослушивании материала с записывающего устройства, синхронизируемого от воспроизводящего устройства.

- АЦП. В современных цифровых системах звукозаписи и воспроизведения основным источником джиттера является АЦП. Нынешние полностью цифровые студийные синхронизаторы достаточно совершенны и часто вносят джиттер меньший, чем АЦП.

Описание прибора с FSK модуляцией (манипуляцией)

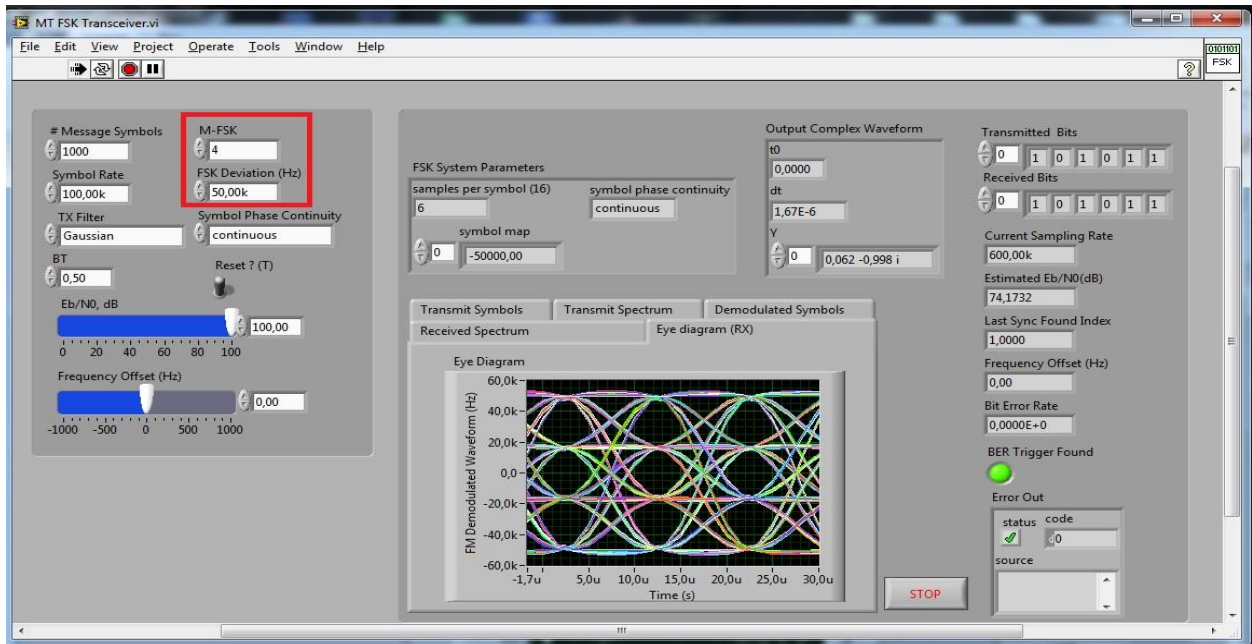


Рис. 1.26. Интерфейс прибора с FSK модуляцией

- Message Symbols – показывает количество передаваемых символов.
- M-FSK – показывает уровень модуляции. Доступны режимы: 2,4,8,16,32,64
- FSK Deviation (Гц) – отклонение по частоте частотной модуляции.
- Eb/N0- отношение сигнал/шум
- Frequency offset –задержка по частоте
- TX Filter – выбор фильтра для передатчика. Доступны следующие: Gaussian, Raised Cosine, Root Raised Cosine
- Symbol Rate – порядок символов.
- Symbol Phase Continuity –продолжительность фазы импульса. Соответственно режимы Continuous(Конечная), Discontinuous (Бесконечная).
- Transmitted Bits – переданные биты.
- Received Bits – принятые биты.
- BER Trigger Found – Вероятность шибки
- BT - безразмерная величина равная $BT = B_{-3дБ} * T$, где $B_{-3дБ}$ - полоса фильтра Гаусса по уровню -3дБ, $T = 1/B_r$ - длительность единичного импульса цифровой информации, передаваемой со скоростью B_r бит\с. Например $B_r = 20\text{кбит}/\text{с}$, тогда $T=50\text{мс}$ и при полосе фильтра Гаусса по уровню -3 дБ $B_{-3дБ} = 10\text{кГц}$ получаем $BT = 10\text{кГц} * 50\text{мкс} = 0.5$. Таким образом, параметр BT

показывает во сколько раз полоса фильтра Гаусса $B_{-3\text{дБ}}$ отличается от скорости передачи информации B_r , выраженной в единицах измерения частоты.

- Сокращения: TX – Передатчик, RX – Приемник

Исследование линии передачи с FSK модуляцией

Предварительная фильтрация

Сравним спектры сигналов и их помехоустойчивость при использовании фильтра Гаусса и без фильтра

Без фильтра при соотношении сигнал/шум = 80

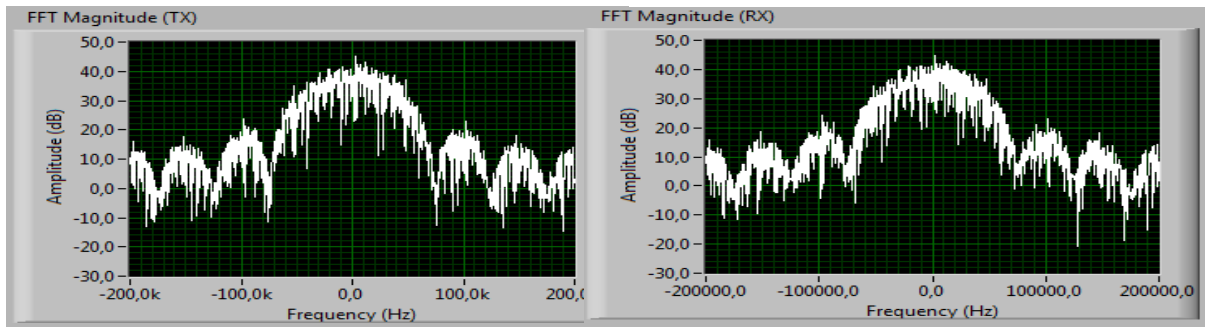


Рис. 1.27. Спектр переданного (слева) и принимаемого (справа) сигнала

Как видно из рисунка 1.27, спектр почти не изменился, но в нем присутствует высокий уровень боковых лепестков, что может повлиять на соседние каналы, также придется выделить более широкий канал в линии передачи, что не всегда является позволительным и выгодным.

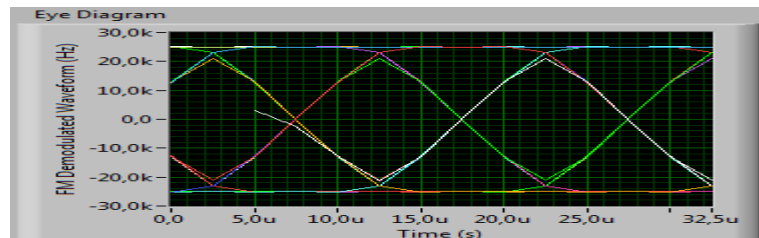


Рис. 1.28. Глазковая диаграмма

Без использования фильтра в сигнал не вносятся дополнительные искажения следовательно джиттер и помехи минимальны.

Применение фильтра Гаусса. При соотношении сигнал/шум = 80, $BT = 0,5$

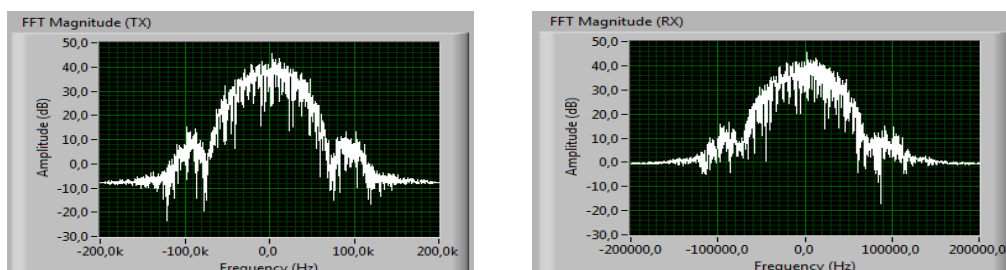


Рис. 1.29. Спектр переданного (слева) и принимаемого (справа) сигнала

Сужение полосы занимаемых частот удалось достигнуть за счет предварительной фильтрации модулирующего сигнала фильтром низкой частоты с Гауссовской импульсной характеристикой.

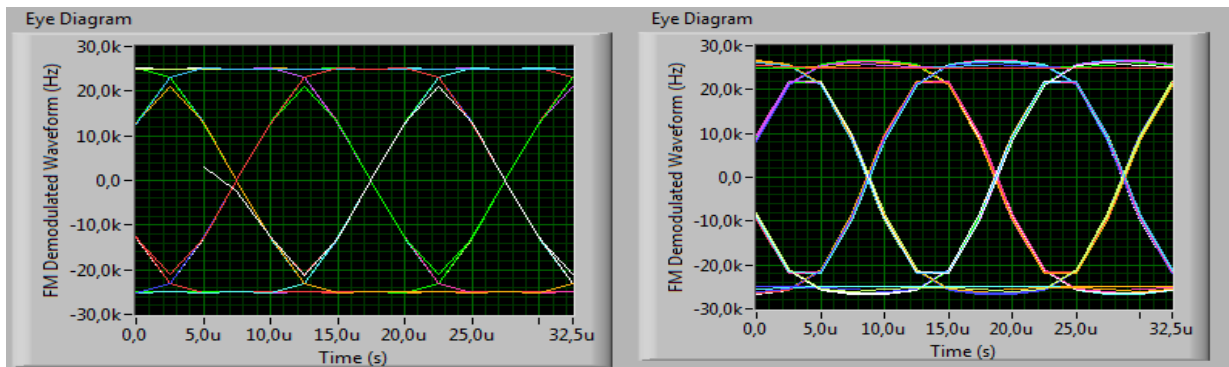


Рис. 1.30. Глазковая диаграмма без фильтра (слева) и с фильтром Гаусса (справа)

Сравнивая две глазковые диаграммы без использования фильтра и с фильтром Гаусса видно, что помехоустойчивость снизилась, появились помехи и искажения, а также увеличился джиттер.

Графики

Фильтр raisedcosine : M-fsk 4 (красный) и 2 (синий)

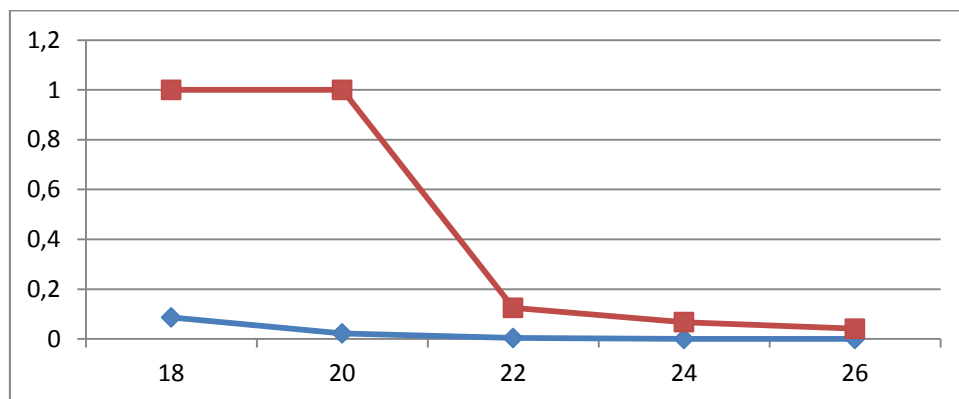


Рис. 1.31. Зависимость вероятности ошибки от отношения сигнал/шум

Фильтр rootraisedcos: M-fsk2 (красный) и 4 (синий)

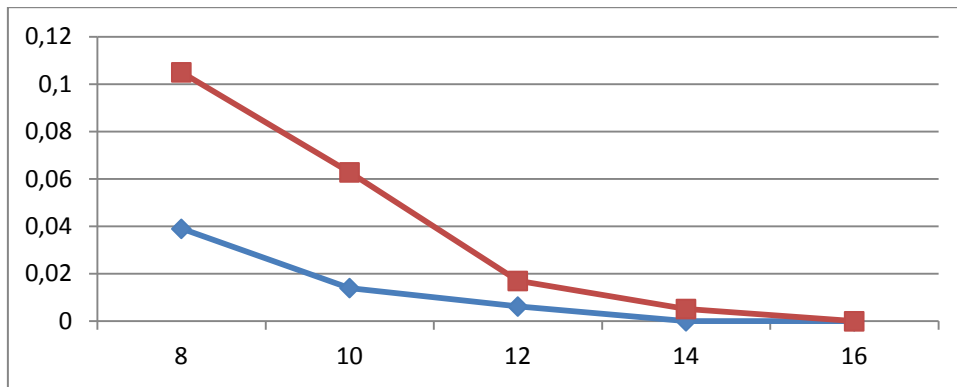


Рис. 1.32. Зависимость вероятности ошибки от отношения сигнал/шум

Фильтр Gaussian: M-fsk2 (красный) и 4 (синий)

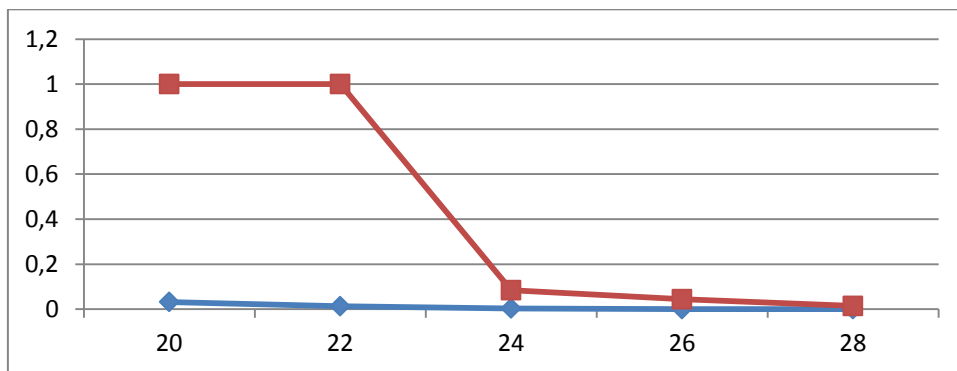


Рис. 1.33. Зависимость вероятности ошибки от отношения сигнал/шум

Частотная манипуляция с минимальным сдвигом (MSK)

MSK это частный случай сигналов с частотной манипуляцией с непрерывной фазой CPFSK с минимальным индексом частотной манипуляции $m=0.5$, поэтому ширина главного лепестка спектра MSK сигналов - минимальная из всех возможных сигналов с частотной манипуляцией.

Представляет собой способ модуляции, при котором не происходит скачков фазы и изменение частоты происходит в моменты пересечения несущей нулевого уровня. MSK уникальна потому, что значение частот соответствующих логическим «0» и «1» отличаются на величину равную половине скорости передачи данных. Другими словами, индекс модуляции равен 0,5.

Частотная модуляция с минимальным разносом частот MSK позволяет уменьшить ширину полосы частот, занимаемых цифровым радиосигналом в эфире. Однако даже этот вид модуляции не удовлетворяет всем требованиям, предъявляемым к современным радиосистемам мобильной связи. Обычно сигнал MSK в радиопередатчике дофильтровывают обычным фильтром. Именно поэтому появился еще один вид модуляции с еще более узким спектром радиочастот в эфире.

Сужение полосы занимаемых частот удалось достигнуть за счет предварительной фильтрации модулирующего сигнала фильтром низкой частоты с Гауссовской импульсной характеристикой. Ширина спектра сигнала GMSK определяется произведением длительности передаваемого символа на полосу пропускания Гауссовского фильтра ВТ. Именно полосой пропускания В и отличаются различные виды GMSK друг от друга. GMSK используется в стандарте GSM сотовой связи.

Формирование GMSK радиосигнала осуществляется таким образом, что на интервале одного информационного бита фаза несущей изменяется на 90° . Это наименьшее возможное изменение фазы, распознаваемое при данном типе модуляции. Непрерывное изменение фазы синусоидального сигнала дает в результате частотную модуляцию с дискретным изменением частоты. Применение фильтра Гаусса позволяет при дискретном изменении частоты получить "гладкие переходы". В стандарте GSM применяется GMSK-модуляция с величиной нормированной полосы ВТ - 0,3, где В - ширина полосы фильтра по уровню минус 3 дБ, Т - длительность одного бита цифрового сообщения. Основой формирователя GMSK-сигнала является квадратурный (1/Q) модулятор. Схема состоит из двух умножителей и одного сумматора. Задача этой схемы заключается в том, чтобы обеспечить непрерывную, очень точную фазовую модуляцию. Один умножитель изменяет амплитуду синусоидального, а второй косинусоидального колебания. Входной сигнал до умножителя разбивается на две квадратурные составляющие. Разложение происходит в двух обозначенных "sin" и "cos" блоках.

Модуляцию GMSK отличают следующие свойства, которые предпочтительны для подвижной связи:

- постоянная по уровню огибающая, которая позволяет использовать эффективные
- передающие устройства с усилителями мощности в режиме класса С;
- компактный спектр на выходе усилителя мощности передающего устройства,
- обеспечивающий низкий уровень внеполосного излучения;
- хорошие характеристики помехоустойчивости канала связи.

Импульсная характеристика Гауссовского фильтра описывается следующей формулой:

$$h(t) = B \sqrt{\frac{2\pi}{\ln(2)}} e^{-\frac{(B_t \pi)^2}{\ln(2)}},$$

где B — полоса пропускания фильтра по уровню 3 дБ.

GMSK с фильтром Гаусса с $BT = 1$, вырождается в классическую MSK. Для формирования сигналов GSM стандартно используется фильтр Гаусса с $BT = 0.3$.

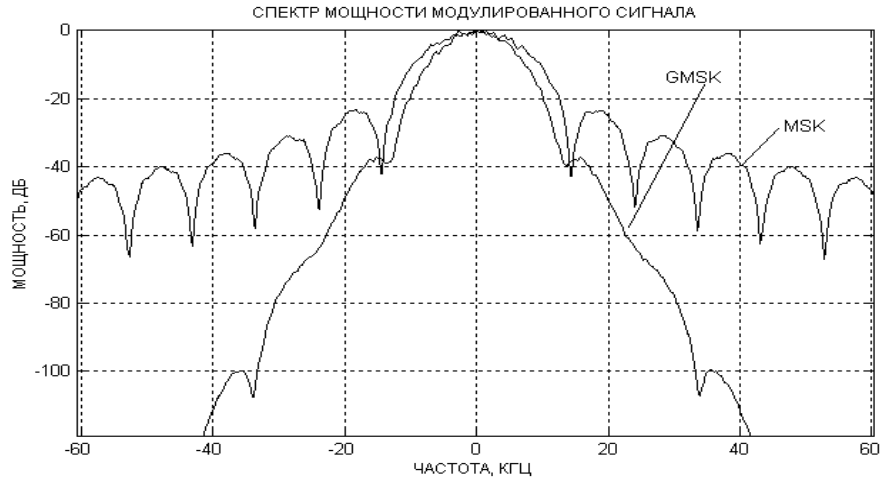


Рис. 1.34. Сравнение спектров двух видов модуляции с фильтром Гаусса (GMSK) и без фильтра (MSK)

В точках взятия отсчетов сигнал GMSK зависит от предыдущих значений передаваемого сигнала. Это вызвано действием гауссовского фильтра, формирующего спектр сигнала GMSK. В результате помехоустойчивость сигнала GMSK ниже по сравнению даже с помехоустойчивостью сигнала MSK. Конкретное значение помехоустойчивости сигналов GMSK сильно зависит от произведения BT . Пример зависимости вероятности ошибки приема сигнала GMSK в зависимости от отношения сигнал/шум на входе решающего устройства приведен на рисунке 1.35.

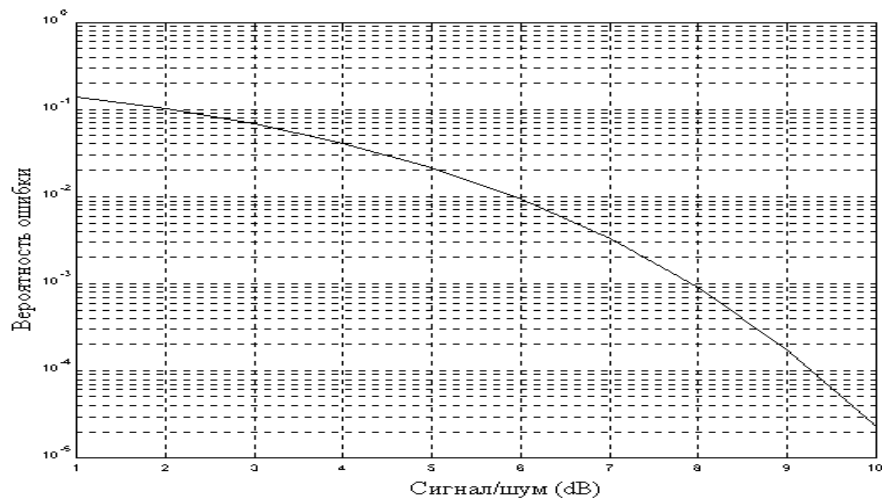


Рис. 1.35. График зависимости вероятности ошибки GMSK в зависимости от отношения сигнал/шум на входе решающего устройства

При индексе модуляции $m = 0.5$ за время передачи одного символа фаза несущего колебания успевает измениться на угол $\pm 90^\circ$. Решетка переходов фазы в MSK сигнале на протяжении двух символьных периодов приведена на рисунке 1.36.

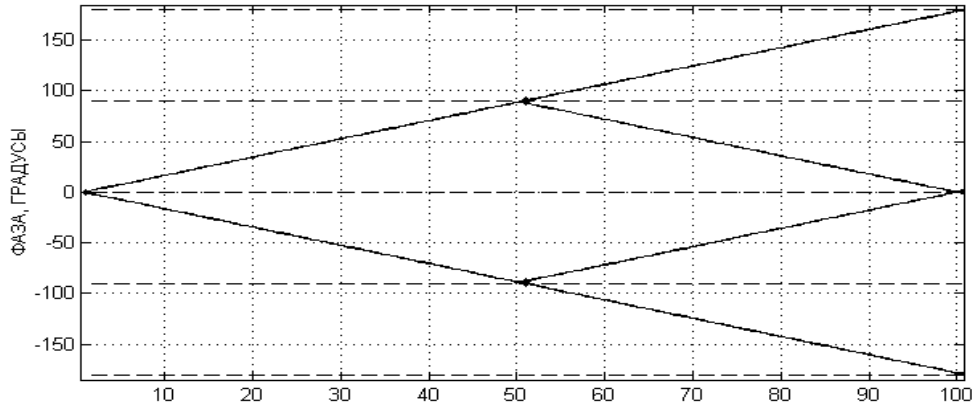


Рис. 1.36. Решетка переходов фазы в MSK сигнале

Два возможных значения фазы несущего колебания в отсчетной точке на одном временном интервале отличаются от двух возможных значений фазы несущего колебания отсчетной точке на соседнем интервале на 90° . Рассмотренная ситуация может быть проиллюстрирована на векторной диаграмме, приведенной на рисунке 1.37.

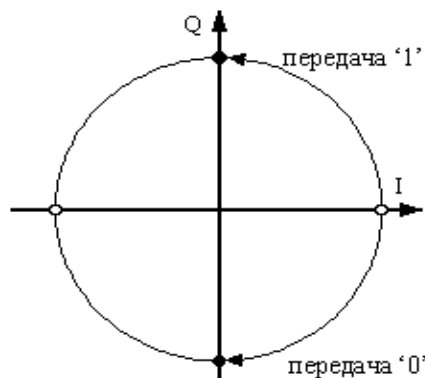


Рис. 1.37. Решетка переходов фазы в MSK сигнале.

Конец вектора, отражающий амплитуду несущего колебания на рисунке 4, в сигнале частотной модуляции может двигаться строго по окружности. На этом рисунке пара состояний сигнала показана пустыми кружочками, а пара состояний сигнала на соседнем символе — заполненными.

Рассмотренные диаграммы показывают, что при разработке радиоприемного устройства можно применить схему фазового детектора. Принимаемые двоичные символы в отсчетных точках будут отличаться друг от друга по фазе на 180° .

Описание прибора с MSK модуляцией (манипуляцией)

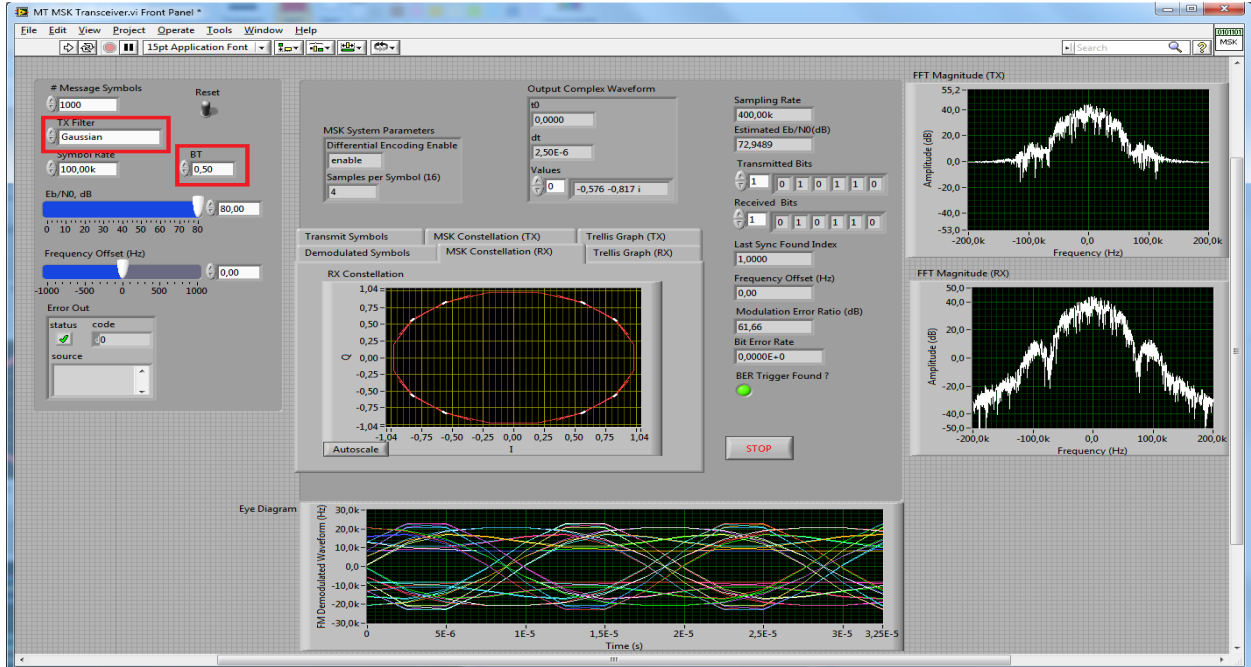


Рис. 1.38. Интерфейс программы с GMSK модуляцией

- Message Symbols – показывает количество передаваемых символов.
- Eb/N0- отношение сигнал/шум
- Frequency offset –задержка по частоте
- TX Filter – выбор фильтра для передатчика. Доступны следующие: Gaussian, Raised Cosine, Root Raised Cosine
- Symbol Rate – порядок символов.
- Transmitted Bits – переданные биты.
- Received Bits – принятые биты.
- BER Trigger Found – Вероятность ошибки
- BT - безразмерная величина равная $BT = 0,5$
- Сокращения: TX – Передатчик, RX – Приемник

Исследование линии передачи с MSK модуляцией

В качестве характеристического параметра *GMSK* используют произведение BT , где B – ширина спектра импульса $h(t)$ по уровню 3 дБ, а T - длительность одного бита. На рис. 1.39 изображена спектральная плотность мощности на выходе идеального

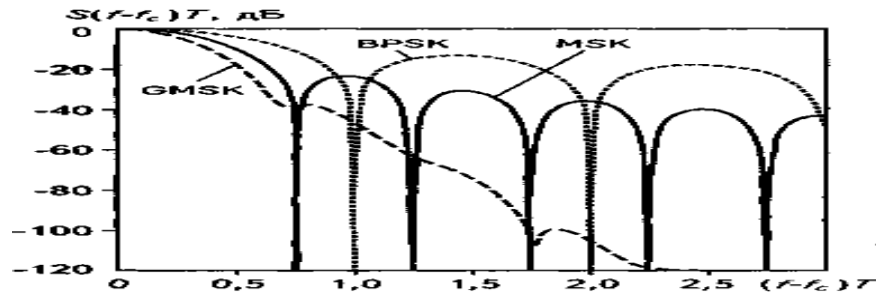


Рис. 1.39. Спектр сигнала с GMSK модуляцией

GMSK-модулятора ($BT = 0,3$), нормализованного по отношению к периоду T . Сравнение этого графика со спектральной плотностью мощности *MSK* и *BPSK*, указывает на серьезное преимущество этой модуляции, прежде всего - в части скорости спадаения внеполосного излучения, т.е. скорости снижения уровня мощности боковых спектральных лепестков.

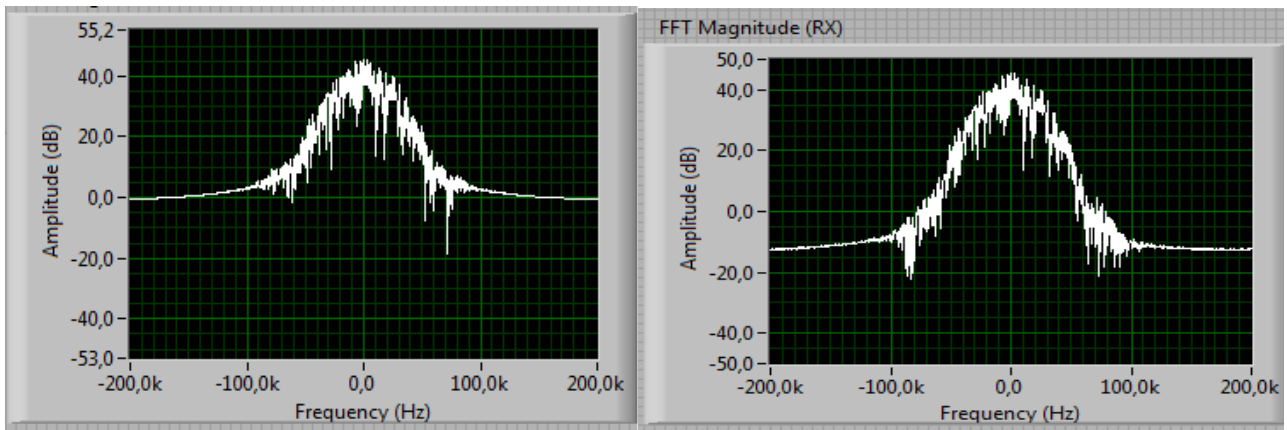


Рис. 1.40. Спектр передаваемого (слева) и принимаемого (справа) сигнала с GMSK модуляцией при $BT = 0.2$, сигнал/шум = 80

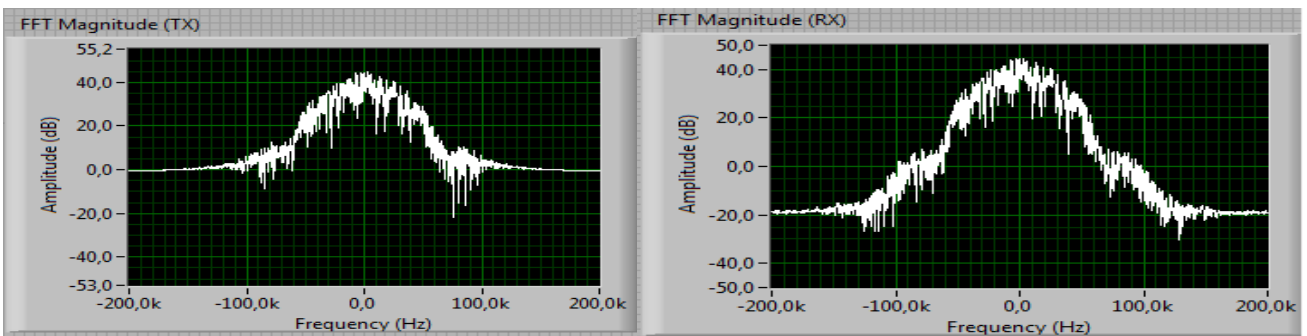


Рис. 1.41. Спектр передаваемого (слева) и принимаемого (справа) сигнала с GMSK модуляцией при $BT = 0.3$, сигнал/шум = 80

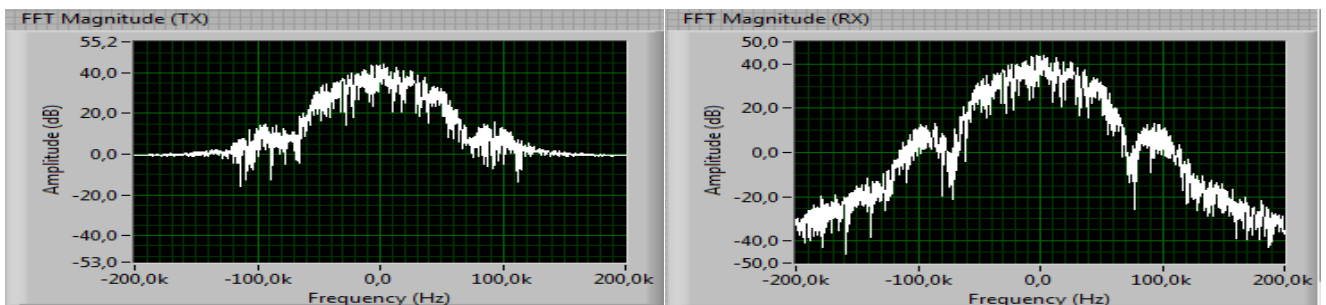


Рис. 1.42. Спектр передаваемого (слева) и принимаемого (справа) сигнала с GMSK модуляцией при $BT = 0.5$, сигнал/шум = 80

Можно видеть, что в спектре GMSK при уменьшении BT уменьшаются уровни боковых лепестков, кроме того значительно возрастает скорость убывания спектра. Так максимальный уровень бокового лепестка GMSK при $BT = 0.3$ меньше, чем при MSK модуляции, а скорость убывания линейно зависит от частоты, что обусловлено применением

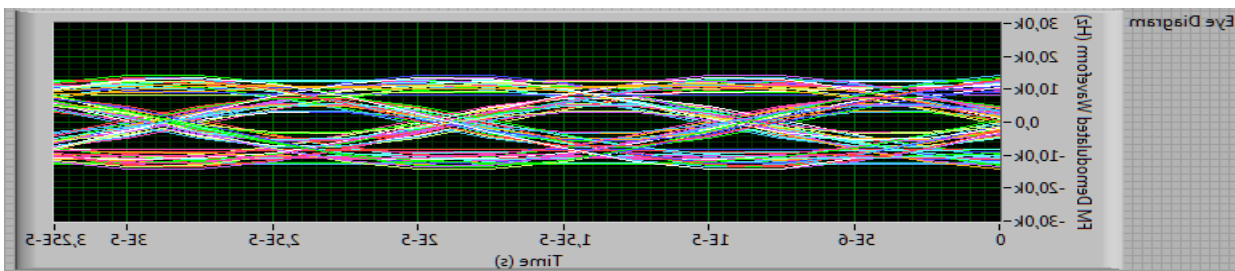


Рис. 1.43. Глазковая диаграмма сигнала с GMSK модуляцией при $BT = 0.2$, сигнал/шум = 80

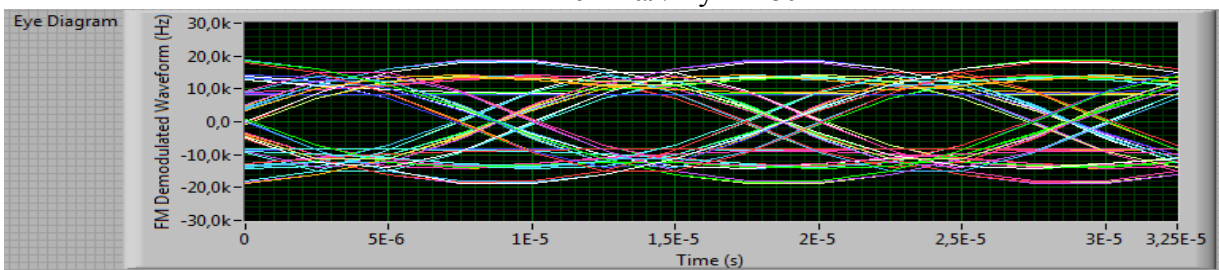


Рис. 1.44. Глазковая диаграмма сигнала с GMSK модуляцией при $BT = 0.3$, сигнал/шум = 80

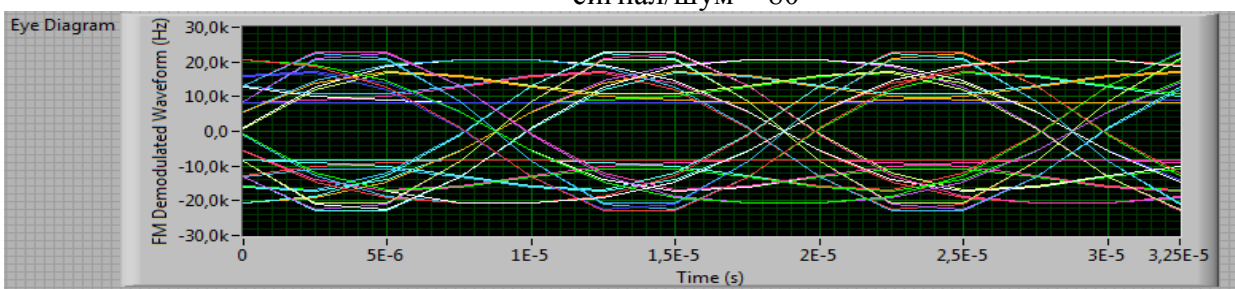


Рис. 1.45. Глазковая диаграмма сигнала с GMSK модуляцией при $BT = 0.5$, сигнал/шум = 80

Из рисунков отчетливо видно, что с уменьшением межсимвольная интерференция усиливается ввиду расширения фильтра Гаусса. Фильтр Гаусса вносит межсимвольную интерференцию и позволяет снизить уровень боковых лепестков спектра, а также

значительно увеличивает скорость убывания спектра GMSK сигнала по сравнению с MSK сигналами. На практике GMSK модуляция нашла применение в сотовой связи формата GSM при , ввиду наиболее эффективного использования радиочастотного ресурса. Дальнейшее уменьшение параметра приводит к существенному усложнению аппаратуры и сильно увеличивает вероятность ошибочного приема символа из-за недопустимо высокой межсимвольной интерференции.

Из полученных графиков можно сделать вывод что для обеспечения той же самой вероятности ошибки требуется различное отношение сигнал\шум, т.е. при использовании фильтра Гаусса требования к качеству принимаемого сигнала возрастают.

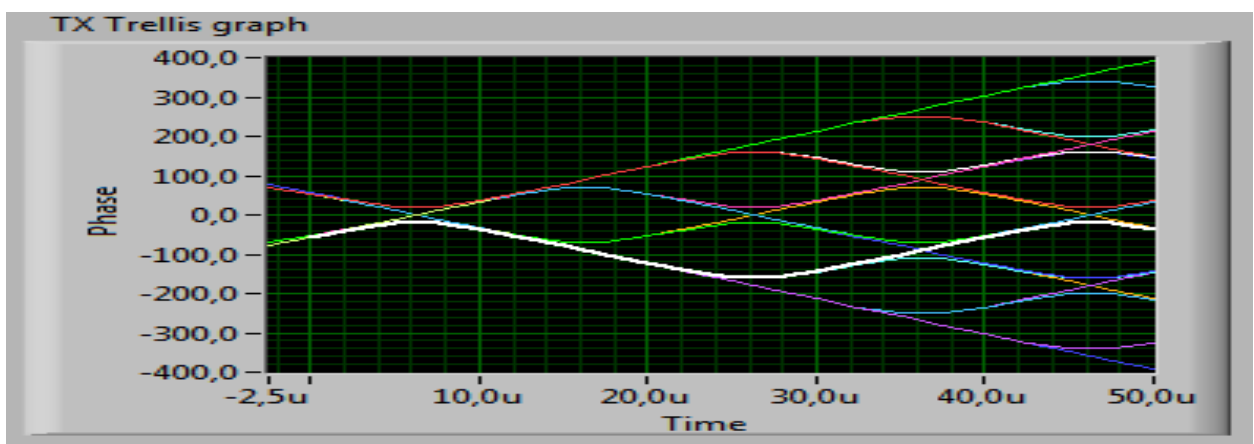


Рис. 1.46. Фазовая диаграмма передаваемого сигнала

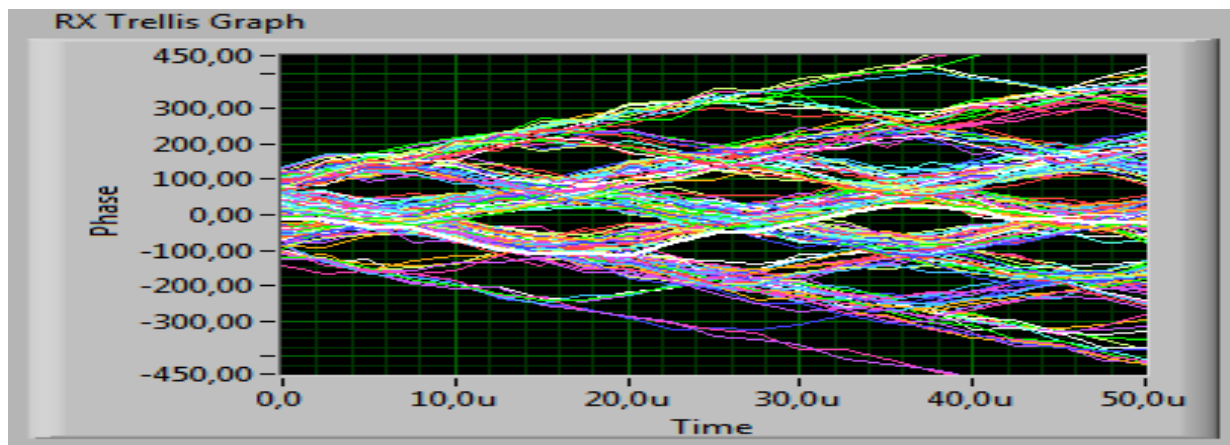


Рис. 1.47. Фазовая диаграмма полученного сигнала при $BER=2.04 \cdot 10^{-3}$,

сигнал/шум = 14

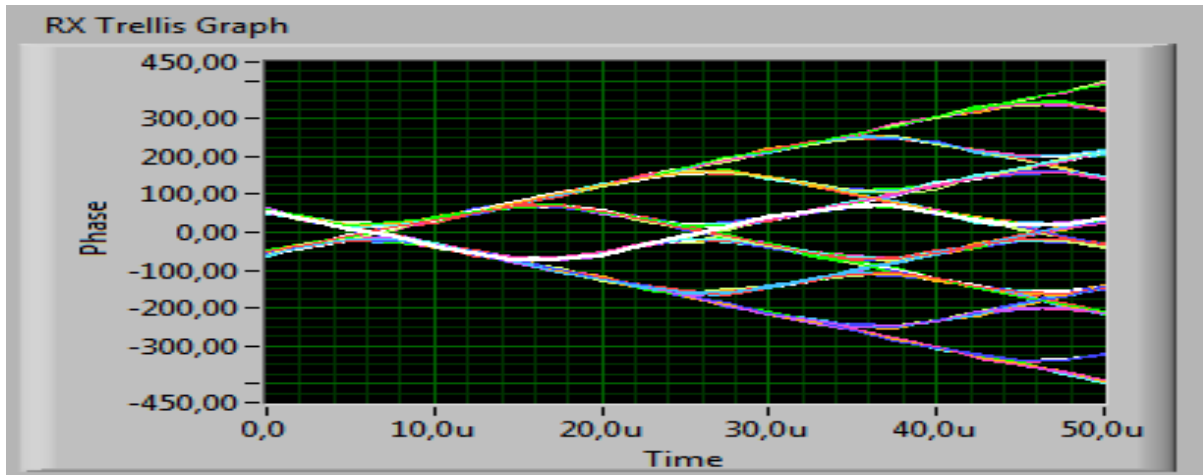


Рис. 1.48. Фазовая диаграмма полученного сигнала при $BER=10^{-4}$, сигнал/шум = 28

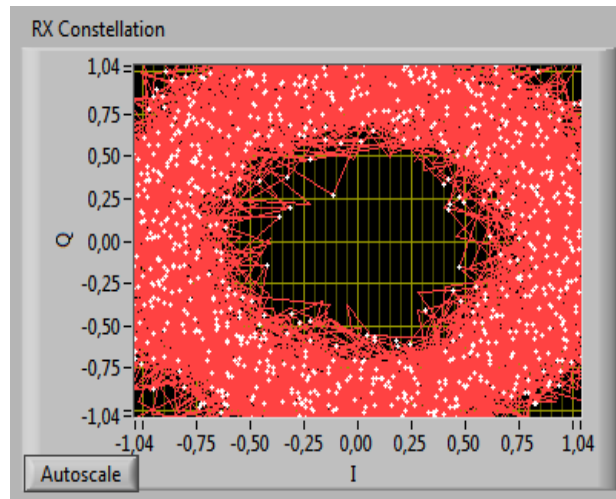


Рис. 1.49. Сигнальное созвездие сигнала при $BER=2.04 \cdot 10^{-3}$, сигнал/шум = 14

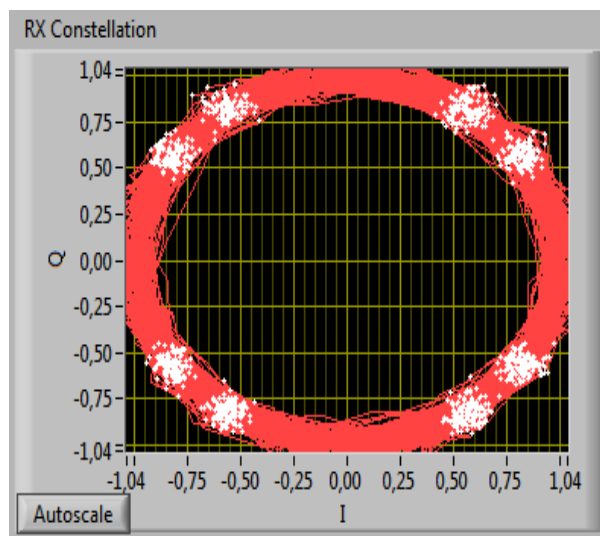


Рис. 1.50. Сигнальное созвездие сигнала при $BER=10^{-4}$, сигнал/шум = 28

TX Filter Raised Cosine (синий), Root Raised Cos (красный), Gaussian (зеленый)

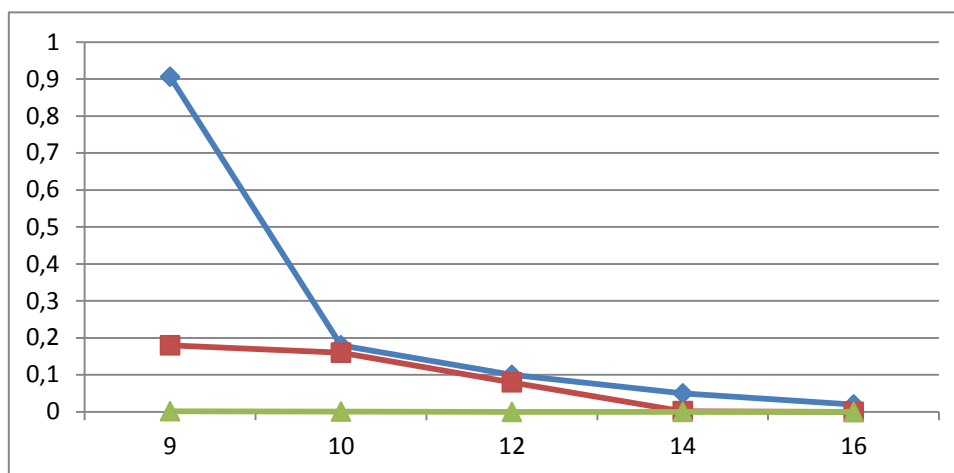


Рис. 1.51. Зависимость вероятности ошибки от отношения сигнал/шум

Таким образом, получаются следующие характеристики этих модуляций. FSK - простота реализации и относительно широкий спектр. GFSK - простота реализации меньший спектр, но и несколько худшая помехоустойчивость по отношению к FSK. MSK - очень высокая помехоустойчивость при относительно небольшом спектре, но сложная схема модулятора-демодулятора. GMSK - ширина спектра практически близка к теоретическому пределу B_T (скорость передачи цифровой информации), несколько худшая помехоустойчивость по сравнению с MSK, сложность модулятора-демодулятора такая же как у MSK.

При использовании программного обеспечения LabVIEW, были созданы две виртуальные измерительные лаборатории с различными видами модуляции сигнала FSK(Frequency Shift Keying) и MSK(Minimal Shift Keying) в каждой из которых можно применить предварительные фильтрации и получить другие подвиды GMSK или GFSK, если использовать фильтр Гаусса.

1.3. Модемы спутниковых систем связи M-QAM, M-PSK и численный анализ вероятности символьной ошибки с использованием ПО LabVIEW

В данном разделе был разработан программный комплекс для изучения многоуровневых методов модуляции применяемых в спутниковых системах разработанная в ПО LabVIEW 2011 SP 1. В разработанном ПО можно исследовать такие виды манипуляции как BPSK, QPSK, 8-PSK, 16-PSK, 32-PSK, 64-PSK и M-QAM на то, как влияет отношение сигнал/шум на вид диаграммы, посмотреть глазковые диаграммы, спектр, оценить BER, посмотреть влияние фильтров на диаграмму.

Наиболее часто используют QPSK манипуляцию. Она имеет преимущество, поскольку может работать при мощности транспондера, близкой к насыщению, то есть

эффективной отдачей энергии. В таблице 1 приведены теоретические значения отношения сигнал/шум, требуемые для достижения величины BER, равной 10^{-10} , без кодирования канала связи для различных методов манипуляции.

Таблица 1.2. Сравнение методов цифровой манипуляции для BER, равной 10^{-10}

Метод манипуляции.	Сигнал/шум без кодирования, дБ.	Эффективность использования полосы пропускания, бит/с.
BPSK	13.06	1.0
QPSK	13.06	2.0
8-PSK	16.55	3.0
16-PSK	21.09	4.0
4-QAM	13.06	2.0
16-QAM	16.98	4.0
64-QAM	21.40	6.0

Квадратурная фазовая манипуляция (QPSK)

QPSK манипуляция строится на основе кодирования двух бит передаваемой информации одним символом. При этом символьная скорость в два раза ниже скорости передачи информации. На рис. 1.43 представлена векторная диаграмма QPSK.

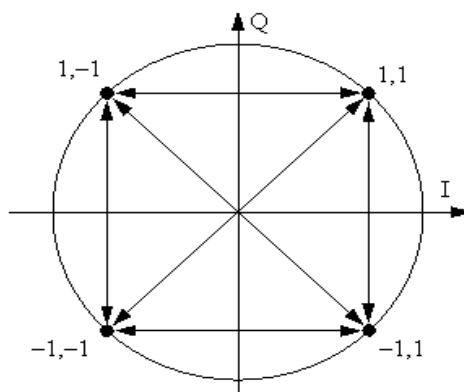


Рис. 1.52. Векторная диаграмма QPSK.

Кодирование осуществляется следующим образом: весь битовый поток разбивается на четные и нечетные биты, тогда $I(t)$ будет кодировать четные биты, а $Q(t)$ – нечетные. Два последовательно идущих друг за другом бита информации кодируются одновременно синфазным $I(t)$ и квадратурным $Q(t)$ сигналами. Это наглядно показано на осциллограммах, приведенных на рисунке 1.53 для информационного потока «1100101101100001».

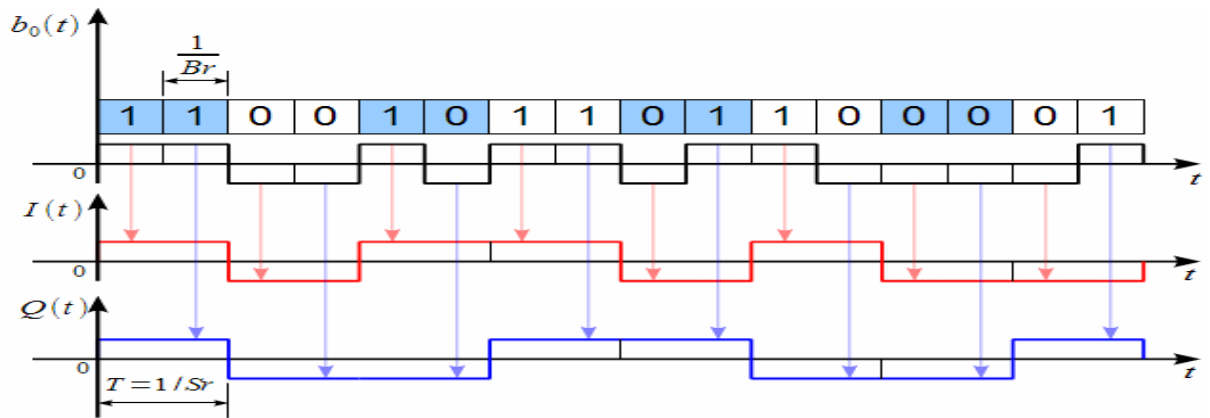


Рис. 1.53. Синфазная и квадратурная составляющие QPSK сигнала

На верхнем графике входной поток разделен на пары бит, соответствующих одной точке созвездия QPSK, показанного на рисунке 1.53. На втором графике показана осциллограмма $I(t)$, соответствующая передаваемой информации $I(t) > 0$. Если четный бит равен 1 (биты нумеруются с нуля, а не с единицы, поэтому первый в очереди бит имеет номер 0, а значит он четный по порядку), и $I(t) < 0$ если четный бит 0 (т.е. $b(t) < 0$). Аналогично строится квадратурный канал $Q(t)$, но только по нечетным битам. Длительность одного символа $T = 1/S_r$ в два раза больше длительности одного бита исходной информации. Устройство выполняющее такое кодирование $I(t)$ и $Q(t)$ согласно созвездию, QPSK условно показано на рисунке 1.54.

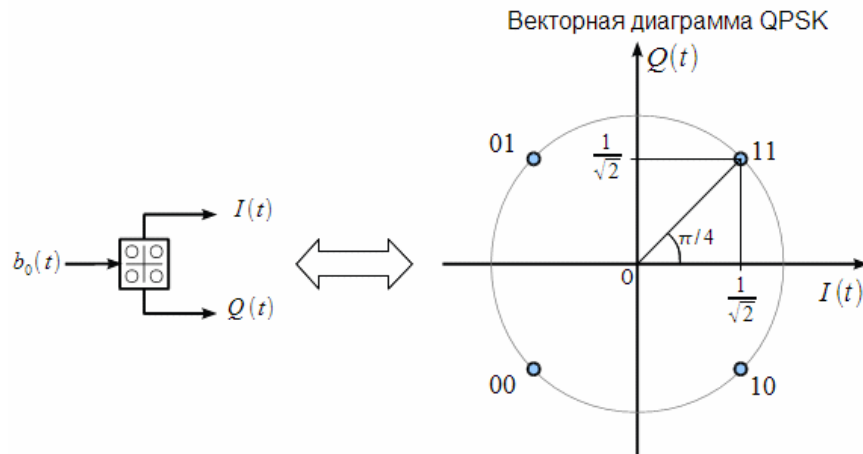


Рис. 1.54. Устройство кодирования синфазной и квадратурной составляющих на основе созвездия QPSK.

В зависимость от пары бит $b(t)$ на входе на выходе получаем постоянные в пределах длительности этой пары бит сигналы $I(t)$ и $Q(t)$, значение которых зависит от передаваемой информации.

Структурная схема QPSK модулятора.

Структурная схема QPSK модулятора на основе универсального квадратурного модулятора показана на рисунке 1.55.

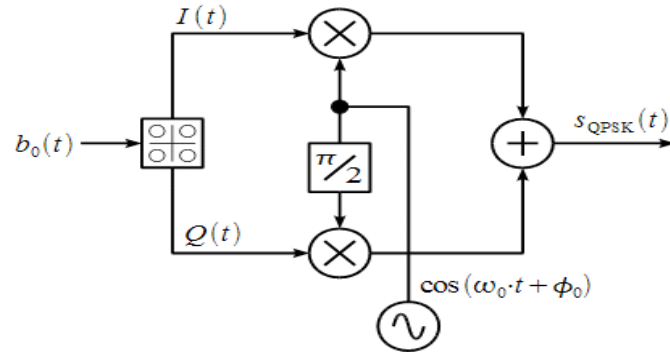


Рис. 1.55. Структурная схема QPSK модулятора.

Сигнал $S_{QPSK}(t)$ имеет вид:

$$S_{QPSK}(t) = I(t) \cdot \cos(\omega_0 t + \varphi_0) - Q(t) \cdot \sin(\omega_0 t + \varphi_0)$$

Синфазная $I(t)$ и квадратурная $Q(t)$ составляющие это ничто иное, как реальная и мнимая части комплексной огибающей QPSK сигнала $Z(t) = I(t) + j \cdot Q(t)$, которые являются входными сигналами квадратурного модулятора. Тогда можно представить $S_{QPSK}(t)$ через его комплексную огибающую $Z(t)$: $S_{QPSK}(t) = R[Z(t) \cdot \exp(j\omega_0 t)]$. Из комплексной огибающей можно выделить фазовую огибающую как:

$$\varphi(t) = \arctan\left(\frac{j[Z(t)]}{R[Z(t)]}\right) = \arctan\left(\frac{Q(t)}{I(t)}\right)$$

Важно отметить, что арктангенс должен вычисляться с учетом четверти комплексной плоскости (функции арктангенс 2). Вид фазовой огибающей $\varphi(t)$ для информационного потока «1100101101100001» показан на рисунке 1.56.

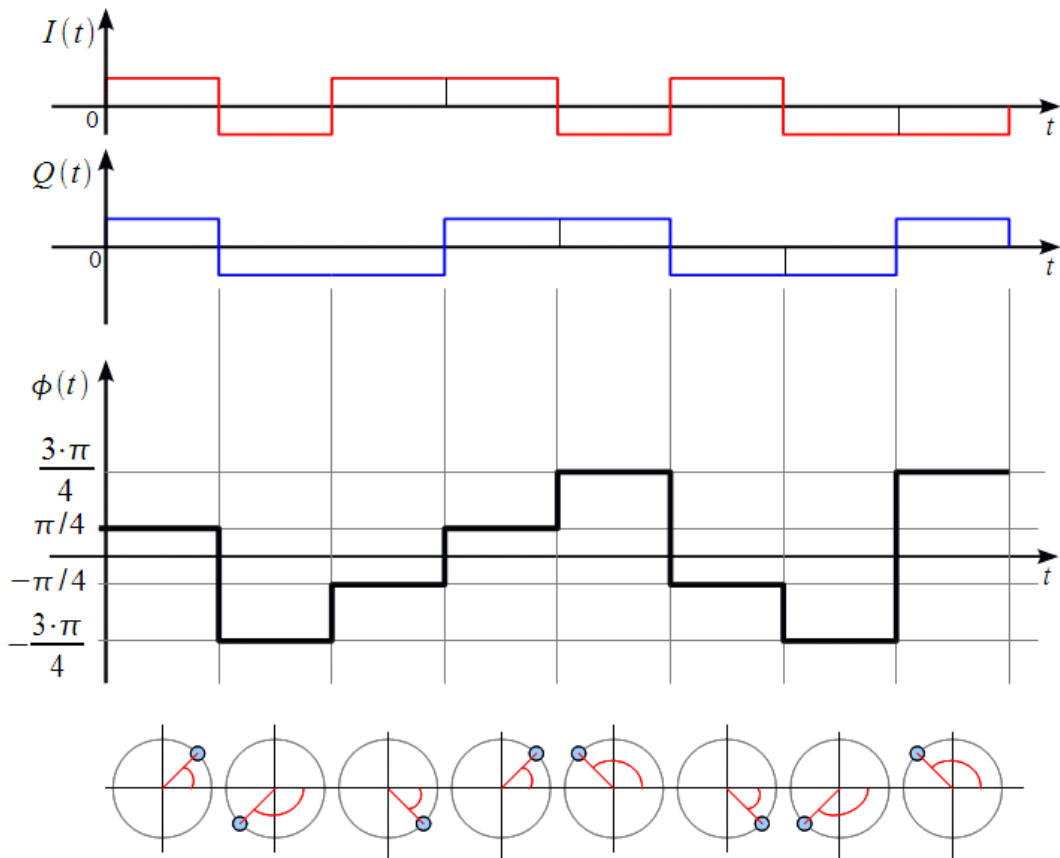


Рис. 1.56. Фазовая огибающая QPSK сигнала.

Фазовая огибающая представляет собой ступенчатую функцию времени, претерпевающую разрывы в моменты смены символа QPSK (один символ QPSK несет два бита информации). При этом в пределах одного символа векторная диаграмма QPSK находится всегда в одной точке созвездия, как это показано внизу, а при смене символа – скачкообразно переходит в точку соответствующую следующему символу. Поскольку у QPSK всего четыре точки в созвездии, то фазовая огибающая может принимать всего четыре значения: $\mp \frac{\pi}{4}$ и $\mp \frac{3\pi}{4}$.

Амплитудная огибающая QPSK сигнала $a(t)$ также может быть получена из комплексной огибающей $Z(t)$:

$$a(t) = \sqrt{I^2(t) + Q^2(t)}$$

Амплитудная огибающая QPSK сигнала равна единице всюду, за исключением моментов смены передаваемых символов, т. е. в моменты перескока фазы и перехода очередной точке созвездия.

Пример осциллограммы QPSK сигнала при входном битовом потоке «1100101101100001» при скорости передачи информации $V_{tr}=10$ кбит/с и несущей частоте 20 кГц показан на рисунке 1.57.

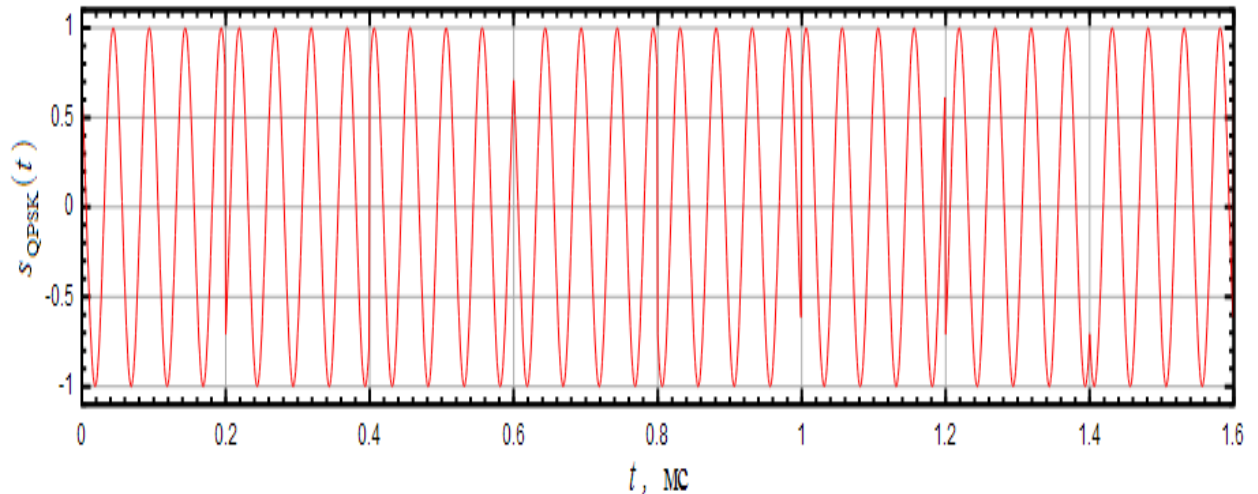


Рис. 1.57. Осциллограмма QPSK сигнала.

Фаза несущего колебания может принимать четыре значения: $\mp \frac{\pi}{4}$ и $\mp \frac{3\pi}{4}$ радиан. При этом фаза следующего символа относительно предыдущего может не измениться, или измениться на $\mp \frac{\pi}{2}$ или на $\mp \pi$ радиан. Также отметим, что при скорости передачи информации $B_r=10$ кбит/с мы имеем символьную скорость $S_r=B_r/2=5$ кбит/с, и длительность одного символа $T=1/S_r=0.2$ мс, что отчетливо видно на осциллограмме (скачок фазы происходит через 0.2 мс).

Формирование спектра QPSK сигнала с помощью фильтров Найквиста.

На рисунке 1.58 показан спектр QPSK $|S_{QPSK}(f)|^2$ сигналов при $B_r=10$ кбит/с и несущей частоте 100 кГц. Можно заметить, что ширина главного лепестка, а также боковых лепестков QPSK сигнала вдвое меньше чем у BPSK сигнала при одной скорости передачи информации. Это обусловлено тем, что символьная скорость S_r QPSK сигнала вдвое меньше скорости передачи информации B_r .

Использование формирующих фильтров дает возможность передавать 0.5 символа в секунду на 1 Гц полосы, или 1 бит/с цифровой информации на 1 Гц полосы при использовании фильтра с АЧХ вида приподнятого косинуса.

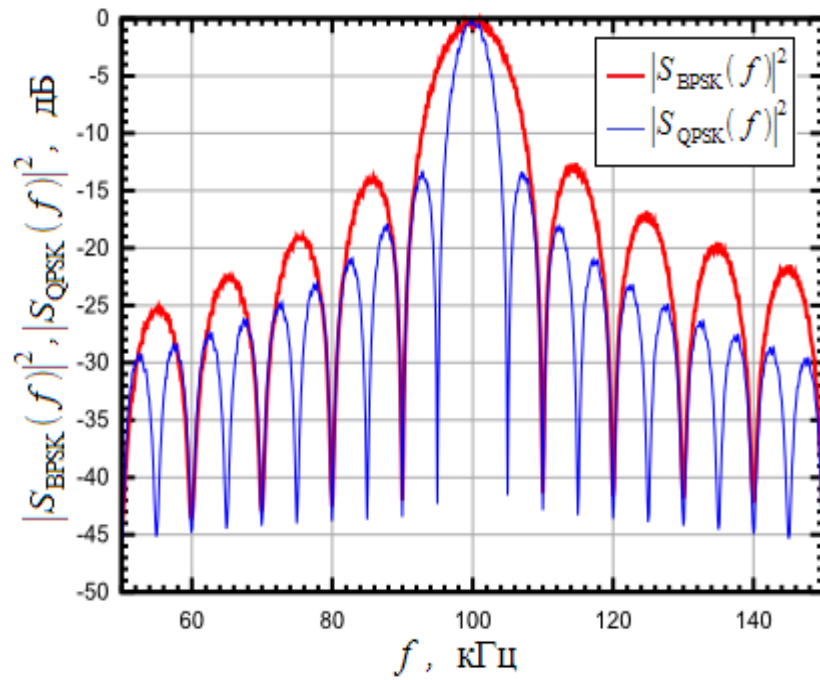


Рис. 1.58. Спектр QPSK сигнала.

На рисунке 1.58 черным показан спектр QPSK сигнала без использования формирующего фильтра. Видно, что применение фильтра Найквиста позволяет полностью подавить боковые лепестки в спектре QPSK сигналов. Структурная схема QPSK модулятора при использовании формирующего фильтра показана на рисунке 1.59.

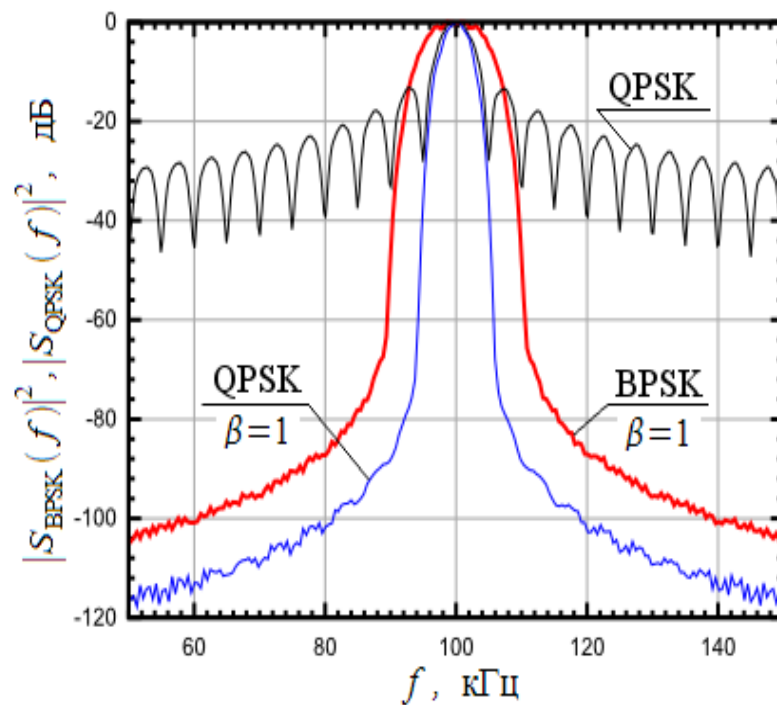


Рис. 1.59. Спектр QPSK сигнала с формирующим фильтром Найквиста

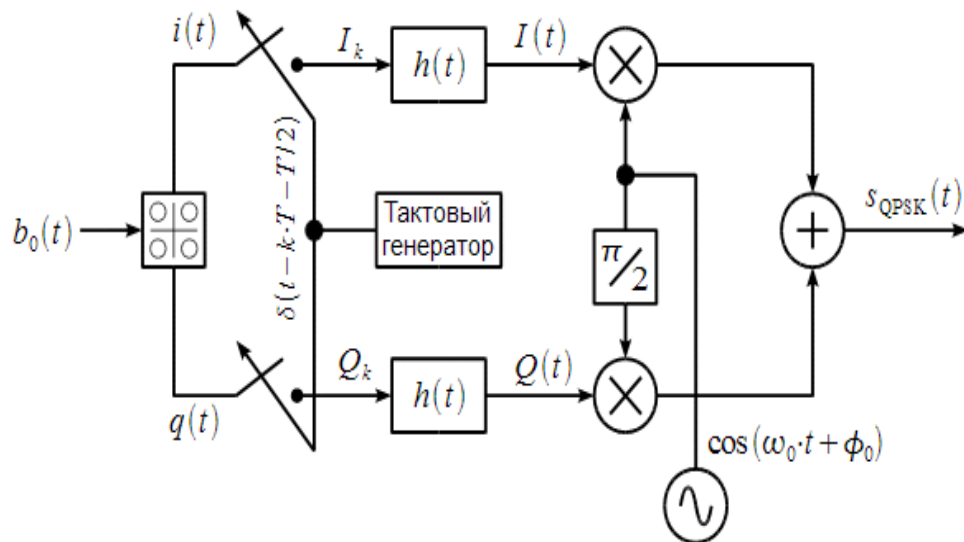


Рис. 1. 60. Структурная схема QPSK модулятора с использованием формирующего фильтра

Графики, поясняющие работу QPSK модулятора показаны на рисунке 1.61.

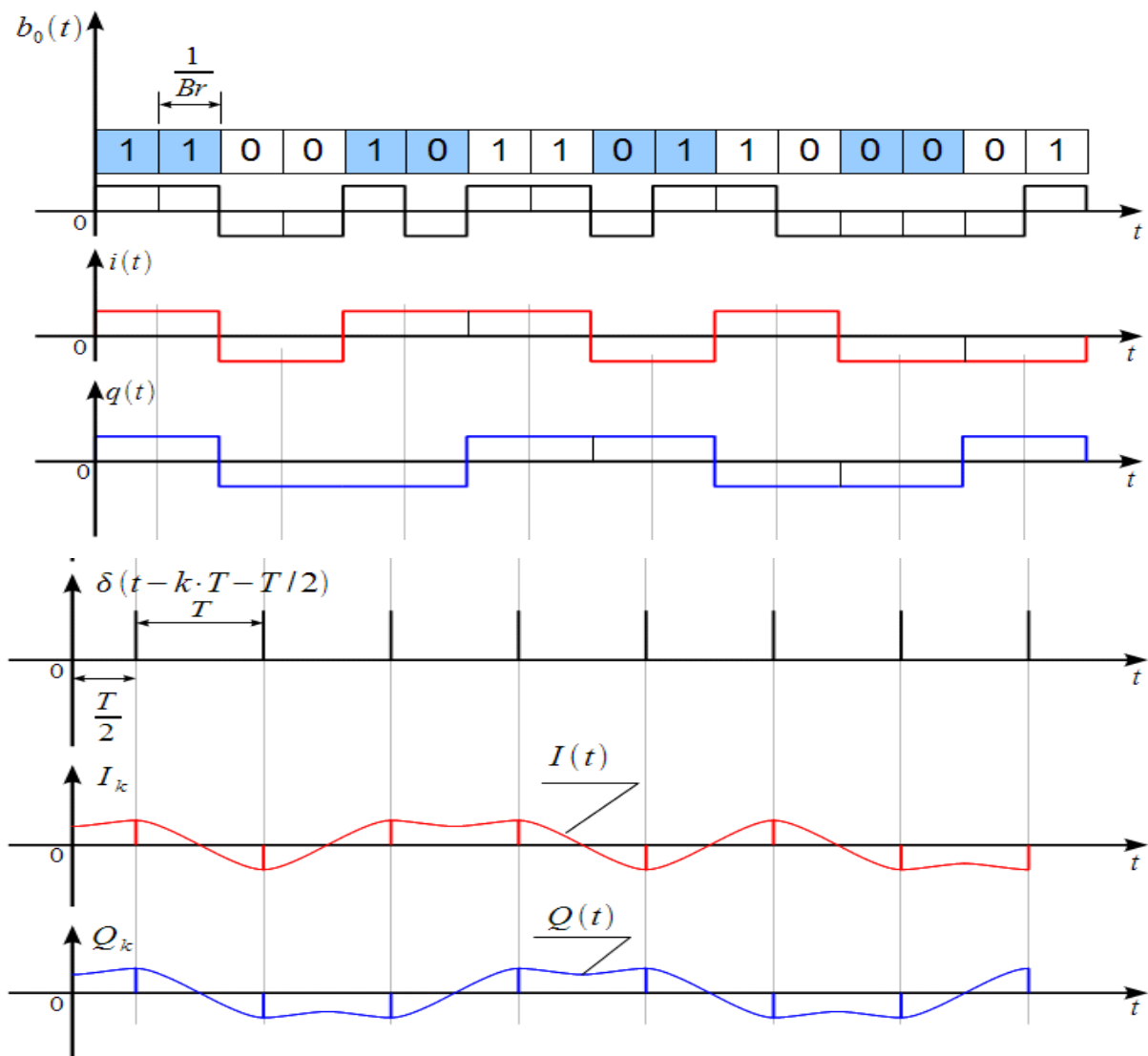


Рис. 1.61. Поясняющие графики.

Цифровая информация поступает со скоростью V_r и преобразуется в символы $i(t)$ и $q(t)$ в соответствии с созвездием QPSK, длительность одного передаваемого символа равна $T=1/S_r=2/V_r$. Тактовый генератор выдает последовательность дельта-импульсов с периодом T , но отнесенных к центру импульса $i(t)$ и $q(t)$, как это показано на четвертом графике. Импульсы тактового генератора стробируют $i(t)$ и $q(t)$ при помощи ключей и получаем отсчеты I_k и Q_k , показанные на двух нижних графиках, которые возбуждают формирующий фильтр интерполятор с импульсной характеристикой $h(t)$ и на выходе имеем синфазную $I(t)$ и квадратурную $Q(t)$ составляющие комплексной огибающей, которые подаются на универсальный квадратурный модулятор. На выходе модулятора получаем QPSK сигнал с подавлением боковых лепестков спектра.

Обратим внимание, что синфазная $I(t)$ и квадратурная $Q(t)$ составляющие становятся непрерывными функциями времени, в результате вектор комплексной огибающей QPSK уже не находится в точках созвездия, перескакивая во время смены символа, а непрерывно движется комплексной плоскости как это показано на рисунке 1.62 при использовании фильтра приподнятого косинуса с различными параметрами β .

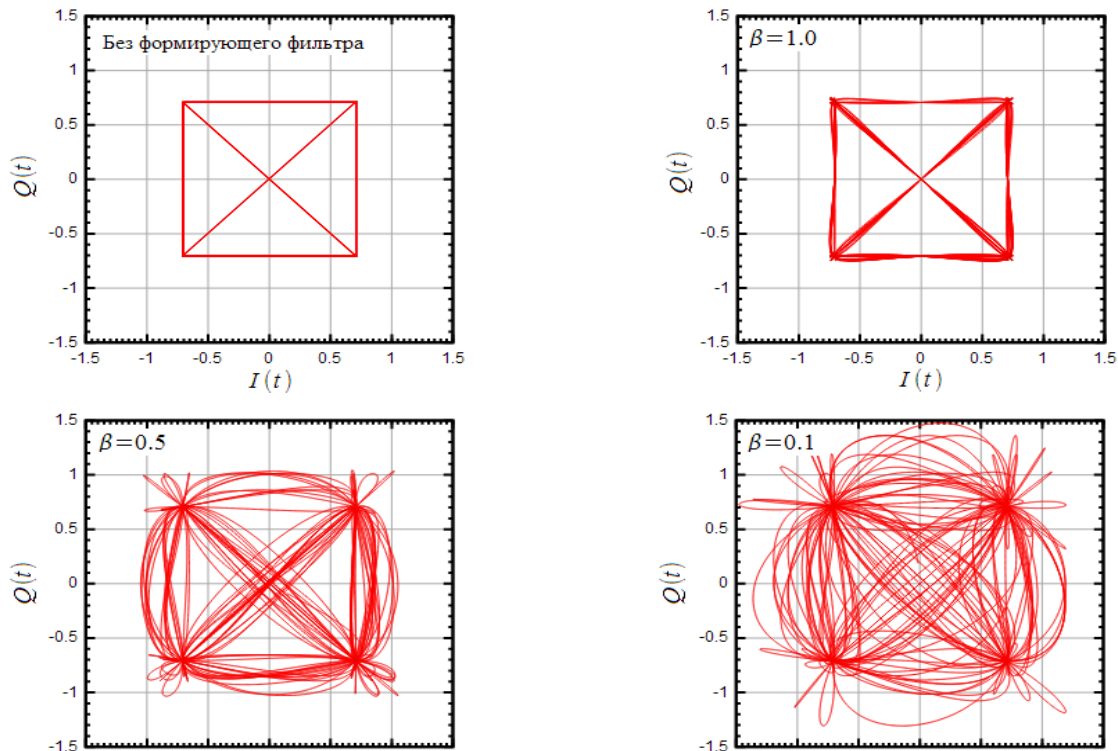


Рис. 1.62. Траектории движения вектора комплексной огибающей QPSK сигнала при различных параметрах формирующего фильтра Найквиста.

На верхнем левом графике показана векторная диаграмма при отсутствии формирующего фильтра при скачкообразном перемещении вектора комплексной огибающей. Если включить фильтр Найквиста, то при уменьшении β векторная диаграмма

превращается в «клубок». При $\beta=1$ получаем наилучшее приближение к идеальной диаграмме.

При непрерывном движении вектора комплексной огибающей его амплитуда начинает меняться во времени, значит у QPSK сигнала при использовании формирующего фильтра начинает меняться амплитудная огибающая

$$a(t) = \sqrt{I^2(t) + Q^2(t)},$$

что наглядно демонстрируется осциллограммой QPSK сигнала, показанной на рисунке 1.63.

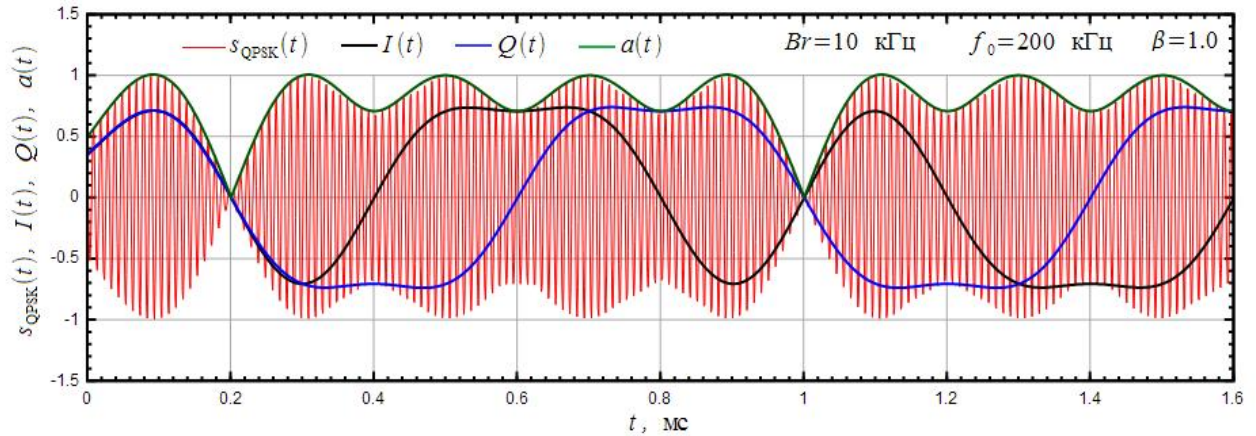


Рис. 1.63. Осциллограмма QPSK сигнала при использовании формирующего фильтра Найквиста

Видно, что фильтр Найквиста приводит к появлению паразитной амплитудной модуляции. При этом в точках когда и синфазная $I(t)$ и квадратурная $Q(t)$ составляющие комплексной огибающей равны нулю, амплитуда $a(t)$ QPSK сигнала также падает до нуля, и фаза поворачивается на π радиан. Глубокая амплитудная модуляция — это негативный эффект, который устраняется офсетной QPSK (OQPSK) модуляции. Важно отметить, что при непрерывных $I(t)$ и $Q(t)$ фазовая огибающая также становится непрерывной функцией времени и перестает меняться скачкообразно, а плавно перетекает от символа к символу, что и приводит к сужению спектра QPSK сигнала при использовании формирующего фильтра.

Практическая часть.

Внешний вид разработанного программного обеспечения представлен на рисунке 1.64.

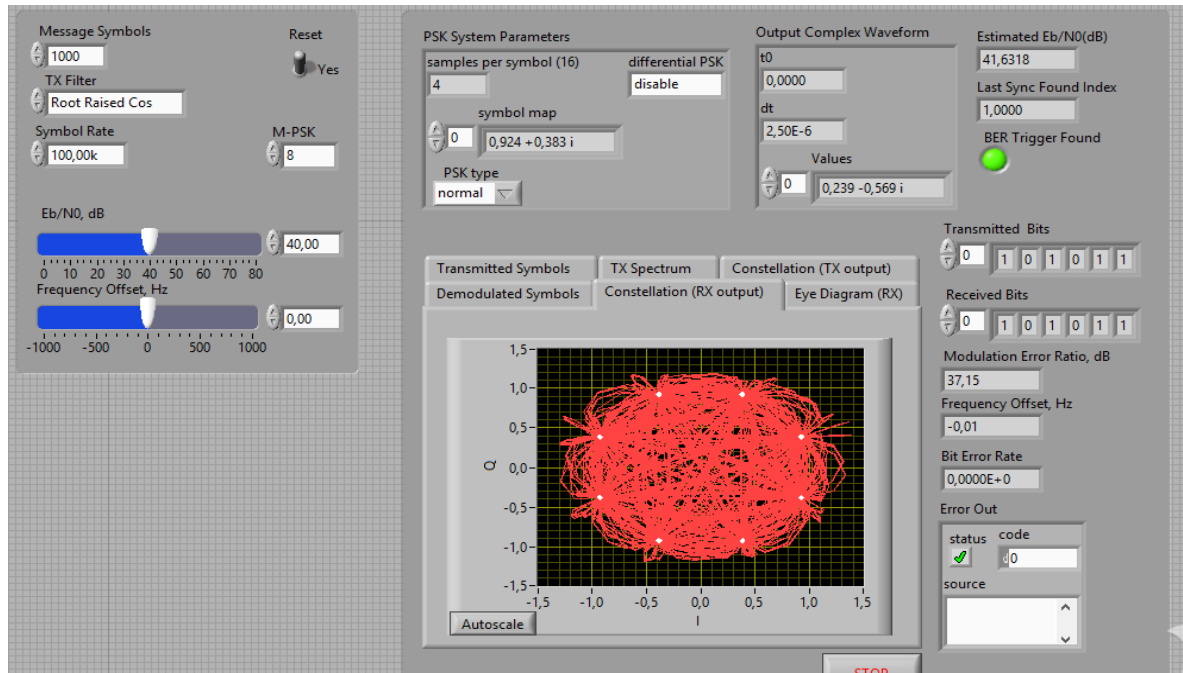


Рис. 1.64. Внешний вид разработанного ПО для исследования QPSK.

Программа имеет два блока:

1. Блок настройки (слева);
2. Блок отображения результата (справа).

В блоке настройки можно задавать количество символов в передаваемом сообщении (поле Message Symbols), чем больше символов в сообщении, тем дольше оно будет передаваться. На рисунке 1.65 показано как меняется диаграмма QPSK снятая с приемника в зависимости от количества символов сообщения.

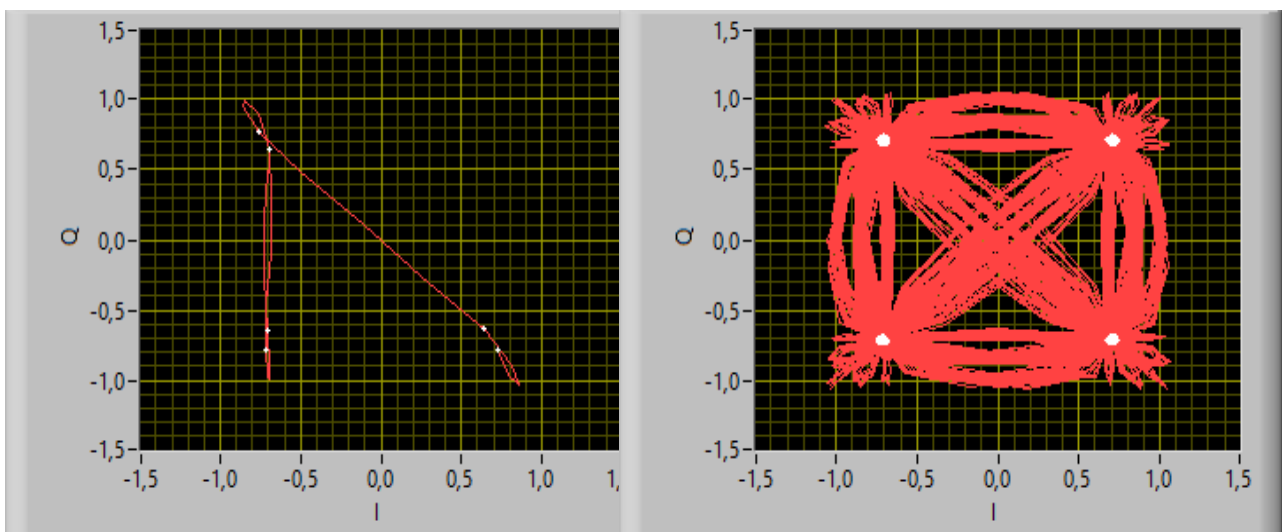


Рис. 1.65. Диаграмма QPSK. 10 символов в сообщении (слева), 1000 символов в сообщении (справа).

Можно задать значение отношения сигнал/шум (поле E_b/N_0 , dB). На рисунке 16 представлена диаграмма QPSK для разных значений сигнал/шум.

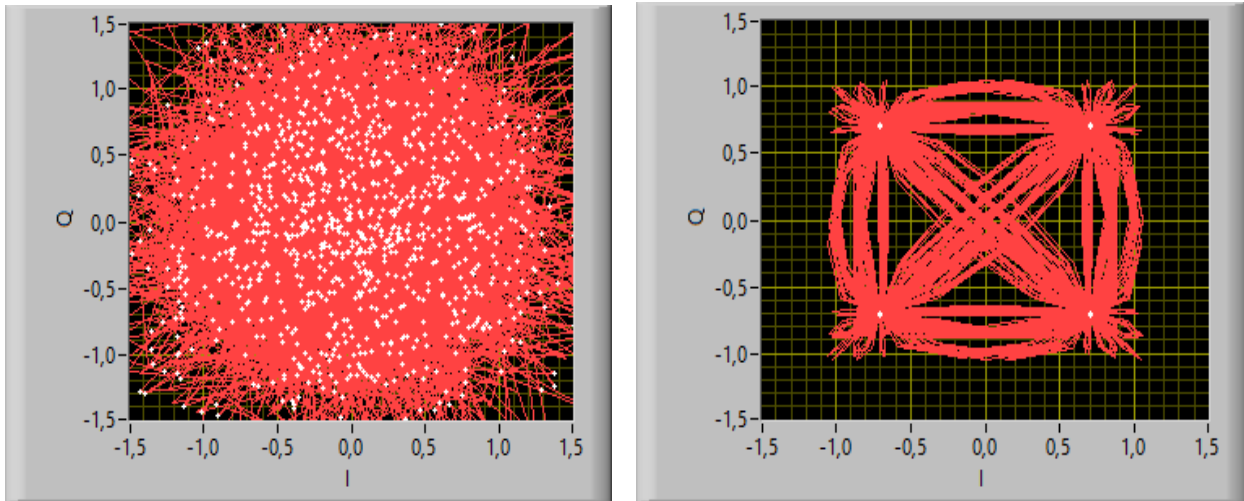


Рис. 1.66. Диаграмма QPSK. Отношение сигнал/шум 0 дБ (слева), Отношение сигнал/шум 60 дБ (справа).

Из рисунка 1.66 видно, что чем больше отношение сигнал/шум, тем четче выглядит диаграмма QPSK.

Так же можно установить:

1. Сдвиг по частоте (поле Frequency Offset, Hz) в диапазоне [-1000,1000] Гц с минимальным шагом 0.01 Гц;
2. Вид фильтра (поле TX Filter).

Можно задать следующие фильтры:

- None (без фильтра);
- Raised Cosine (Приподнятый косинус);
- Root Raised Cos();
- Gaussian (Гаусса).

3. Число точек M в созвездии (поле M-PSK). Можно задать следующие значения:

- 2 (BPSK);
- 4 (QPSK);
- 8 (8-BSK);
- 16 (16-BSK);
- 32 (32-BSK);
- 64 (64-BSK).

4. передаваемых бит (поле Transmitted Bits).

Последовательность

5. бит (поле Received Bits).

Последовательность принимаемых

6. Rate).

Скорость передачи (поле Symbol

На рисунке 1.60 показано как изменяется диаграмма QPSK в зависимости от выбранного значения M .

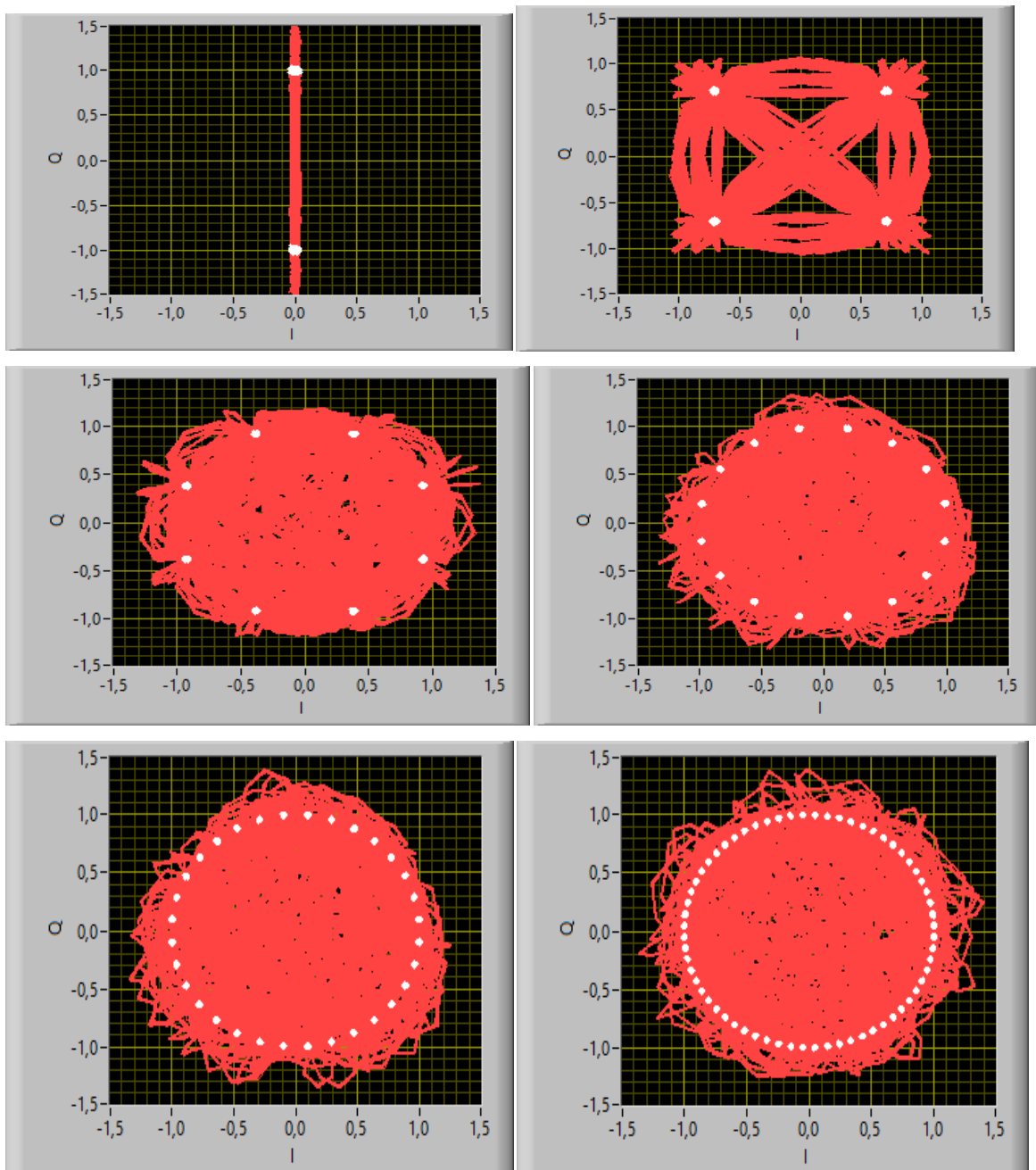


Рис. 1.67. Диаграмма QPSK. Слева на право сверху вниз - 2,4,8,16,32,64.

В разработанном программном обеспечении можно снимать спектр передатчика (вкладка TX Spectrum). На рисунке 1.67 представлены спектрограммы QPSK для разных фильтров.

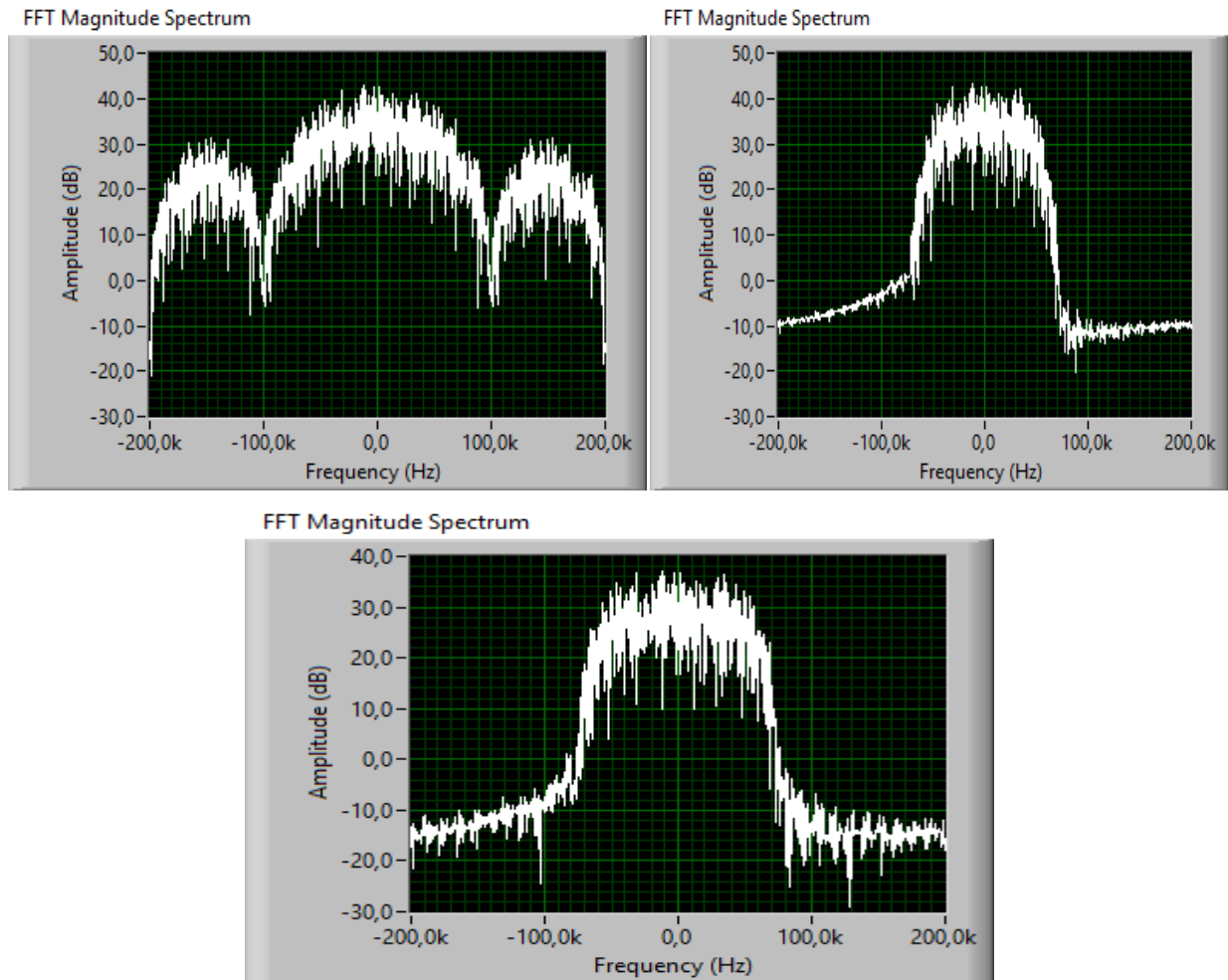


Рис. 1.68. Спектрограммы QPSK с выставленными значениями фильтра (слева на право - сверху вниз): без фильтра, приподнятый косинус, Root Raised Cos

В разработанном программном обеспечении можно снимать глазковую диаграмму с приемника (вкладка Eye Diagramm(RX)). На рисунке 1.68 представлены глазковые диаграммы при разных значения М для оси Q, а на рисунке 1.69 представлены глазковые диаграммы при разных значения М для оси I

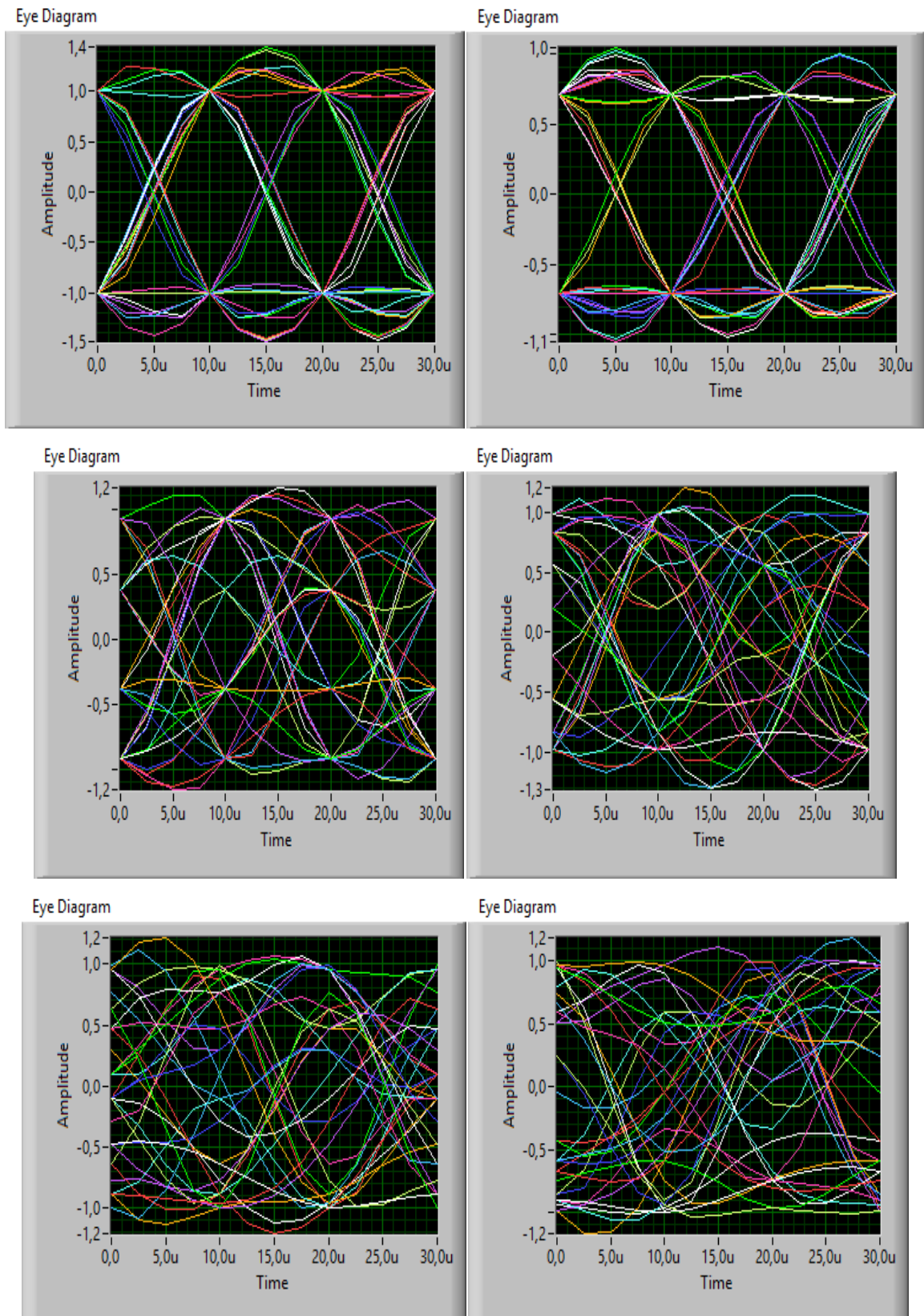


Рис. 1.69. Глазковые диаграммы для разных значений M на оси Q (слева на право – сверху вниз): 2,4,8,16,32,64

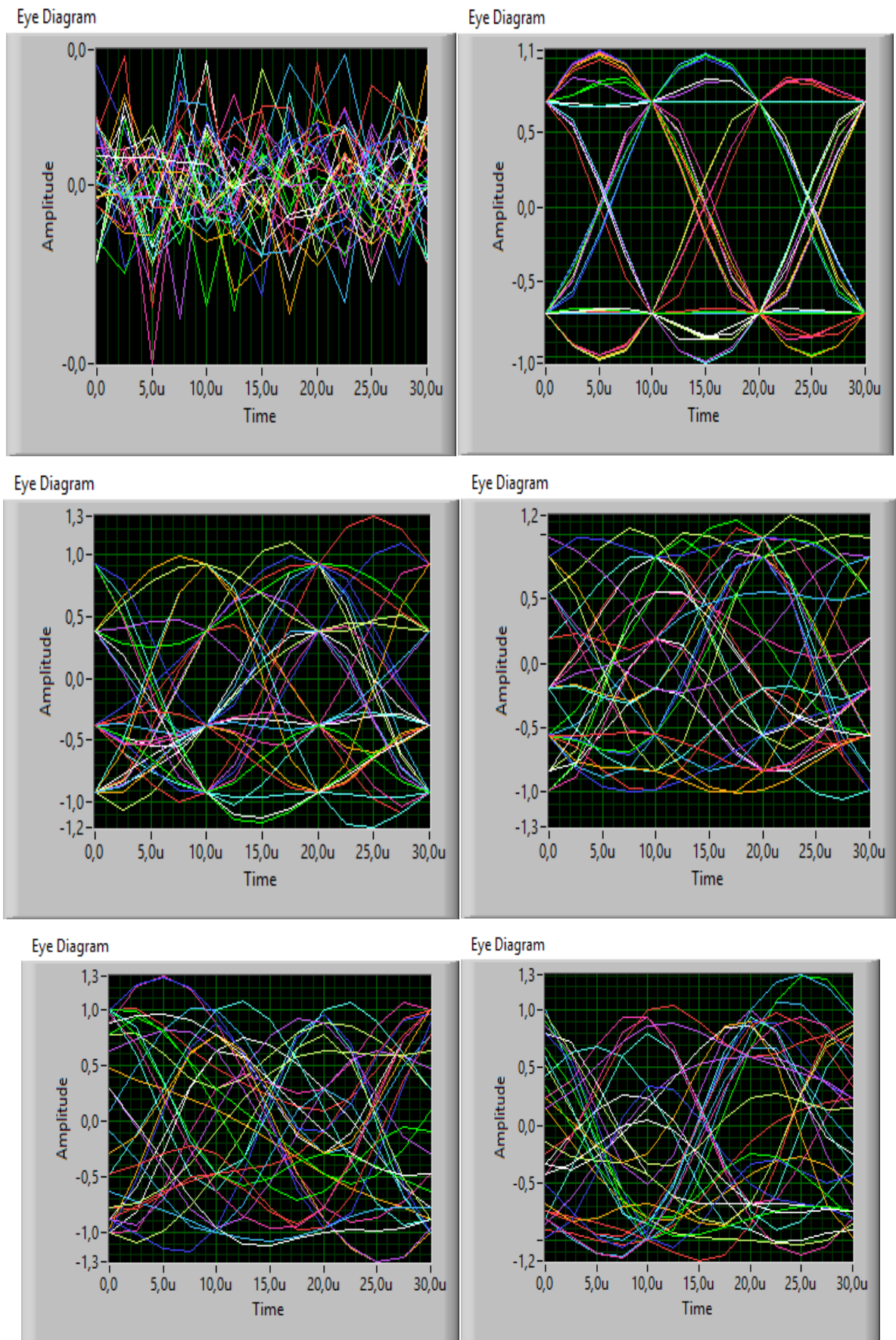


Рис. 1.70. Глазковые диаграммы для разных значений M на оси I (слева на право – сверху вниз): 2,4,8,16,32,64

Ось выбирается на вкладке Eye Diagramm (RX) в поле Display I/Q.

Так же программное обеспечение рассчитывает BER (Bit Error Rate – битовая вероятность ошибки). Это значение можно увидеть в поле Bit Error Rate.

Быстро определить есть ошибки или нет можно если посмотреть на зеленую кнопку с подписью BER Trigger Found, если она будет гореть значит битовая вероятность ошибки присутствует.

В правом нижнем углу есть информационное окно Error Out, в которое будет выводиться код и сообщение ошибки в случае не корректной работы программы.

На вкладках Transmitted Symbols и Demodulated Symbols можно посмотреть переданные и демодулированные символы соответственно.

Квадратурно-амплитудная модуляция - QAM (QAM – Quadrature Amplitude Modulation) может рассматриваться как расширенная многоуровневая ФМ, в которой два исходных сигнала генерируются независимо. Таким образом, здесь имеют место два полностью независимых квадратурных канала, включающие процессы кодирования и детектирования в основной полосе.

На рисунке показано сигнально - точечное пространство для системы с 16-QAM и четырьмя уровнями в каждом квадратурном канале. Точки представляют составной сигнал, а штрихи на осях отмечают уровни амплитуды в каждом квадратурном канале. Основная схема модулятор – демодулятора 16-QAM представлена на рисунке .

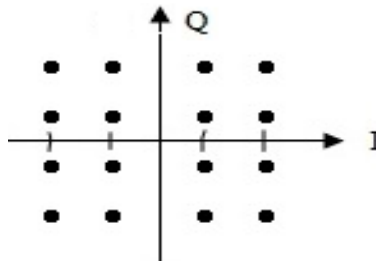


Рис.1.71. Сигнально-точечное пространство модуляции для 16-QAM

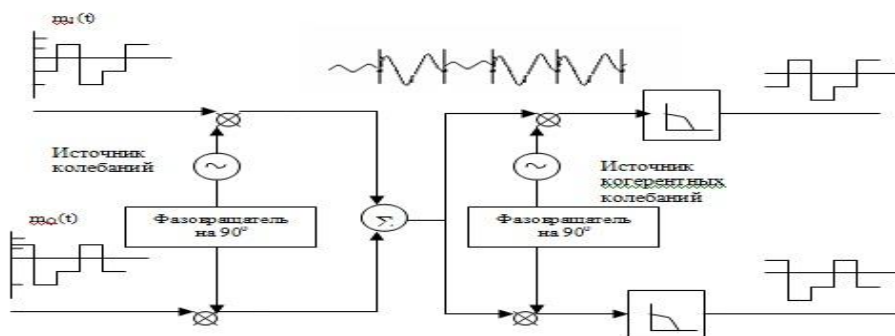


Рис. 1.72. Схема модулятора-демодулятора QAM

В отличие от ФМ сигналов сигналы QAM, показанные на рисунке не содержат постоянной огибающей. Наличие постоянной огибающей в ФМ объясняется поддержанием отношения уровней в квадратурных каналах. В QAM такие ограничения не вводятся ввиду того, что в каждом канале уровни независимы.

Характеристики ошибок систем QAM и ФМ модуляций сильно отличаются. При достаточно большом числе сигнальных точек системы QAM имеют, как правило, лучшие характеристики, чем системы с ФМ. Основная причина состоит в том, что расстояние между сигнальными точками на диаграмме для системы с QAM больше, чем для соответствующей системы с ФМ.

Расстояние d между соседними точками в системе QAM с нормированной к единице пиковой амплитудой и числом уровней L может быть представлено в виде:

$$d = \frac{\sqrt{2}}{L-1} \quad (1.14)$$

На рисунке представлено сравнение систем QAM-16 и ФМ-16 работающих на одинаковой пиковой мощности, по расстоянию между точками.

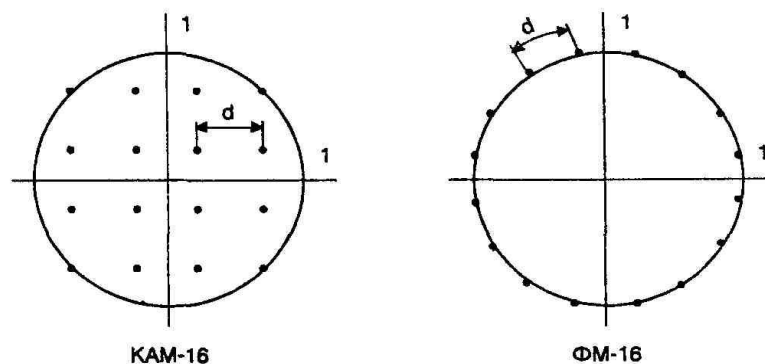


Рис. 1.73. Сравнение систем QAM-16 и ФМ-16 работающих на одинаковой пиковой мощности, по расстоянию между точками

QAM имеет преимущество над системой ФМ при той же пиковой мощности.

В настоящее время для передачи пользуются системами 256-QAM. Надо отметить, что надежное функционирование высокоплотных форматов модуляции, таких как 256-QAM требует строгой линейности усилителей, для возможности обработки широкого диапазона амплитуд сигналов. Соотношения для характеристик ошибок методов 4-, 16-, 64- и 256-QAM

в зависимости от отношения функции $\frac{E_b}{N_0}$ приведены на рисунке.

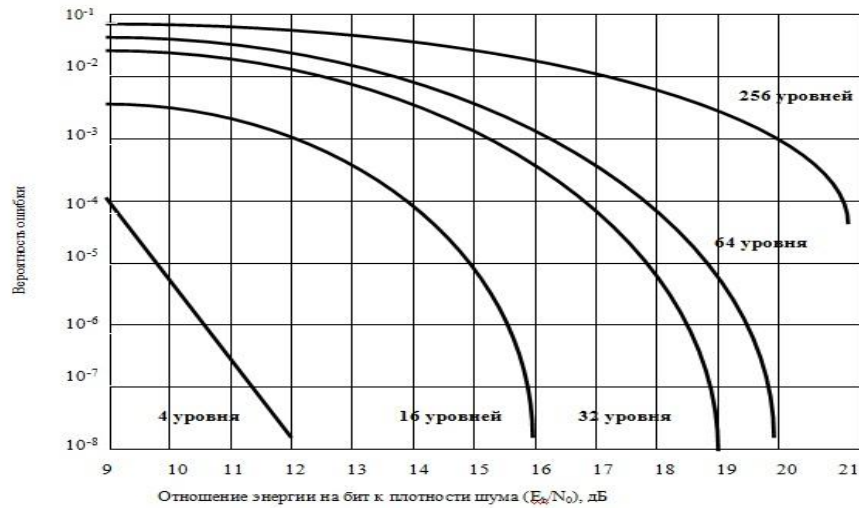


Рис. 1.74. Вероятности ошибок в системах QAM.

Достоинство высоких значений номера QAM – это повышенная скорость передачи данных, поскольку таким образом большее количество битов информации может быть передано в течении одного цикла. Однако, с другой стороны, в этом случае большее число уровней амплитуды сигнала располагаются близко друг к другу, повышая тем самым вероятность неразличимости двух уровней, и как следствие – повышая чувствительность системы к шуму. Таким образом, высокие значения номера QAM более требовательны к параметру SNR (Signal Noise Ratio – Отношение Сигнал/Шум).

Практическая часть QAM

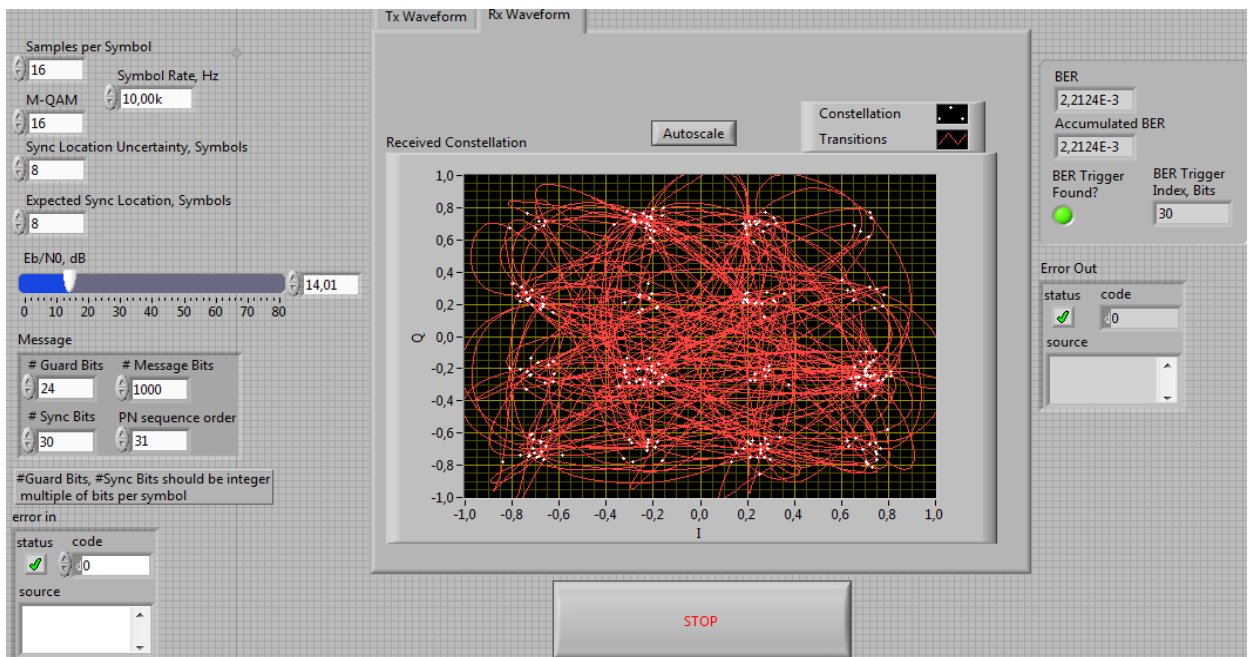


Рис. 1.75. Внешний вид разработанного ПО для исследования QAM

Число посылок 2000 и различные отношения сигнал/шум

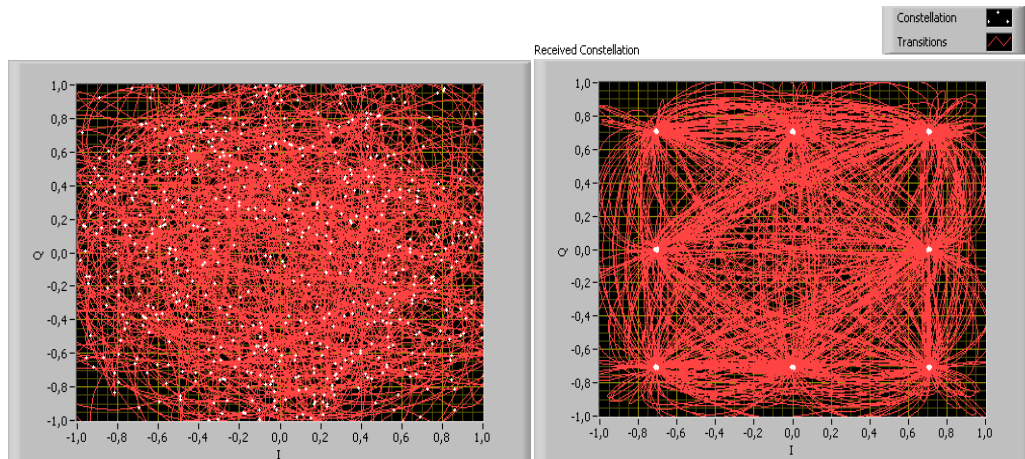


Рис. 1.76. Созвездия для QAM-8 передаваемого сигнала.

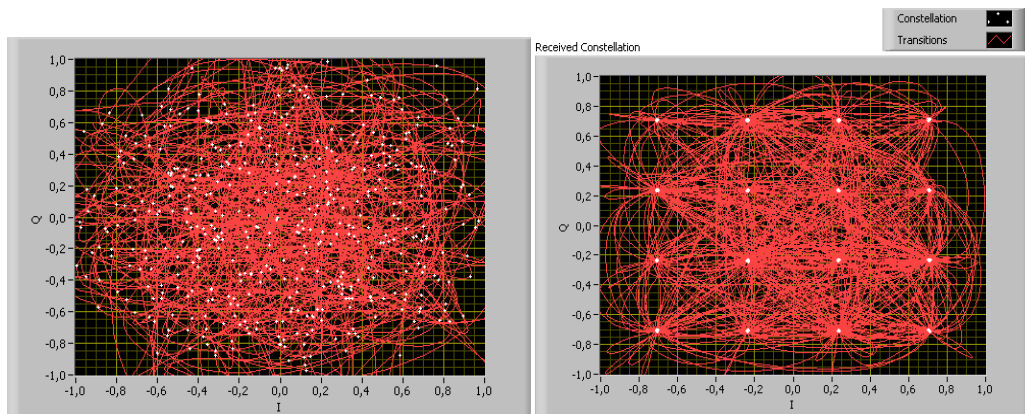


Рис. 1.77. Созвездия для QAM-16 передаваемого сигнала.

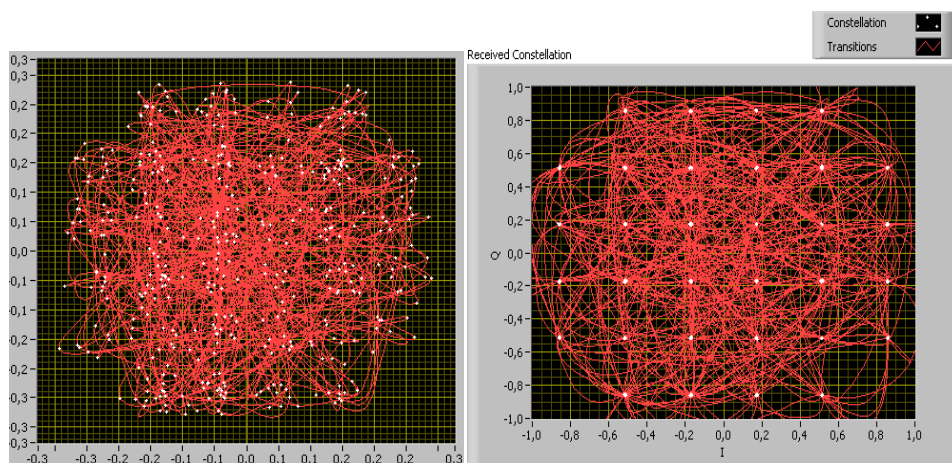


Рис. 1.78 Созвездия для QAM-32 передаваемого сигнала.

На рисунке 1.79. приведены глазковые диаграммы для QAM-8 сигнала.

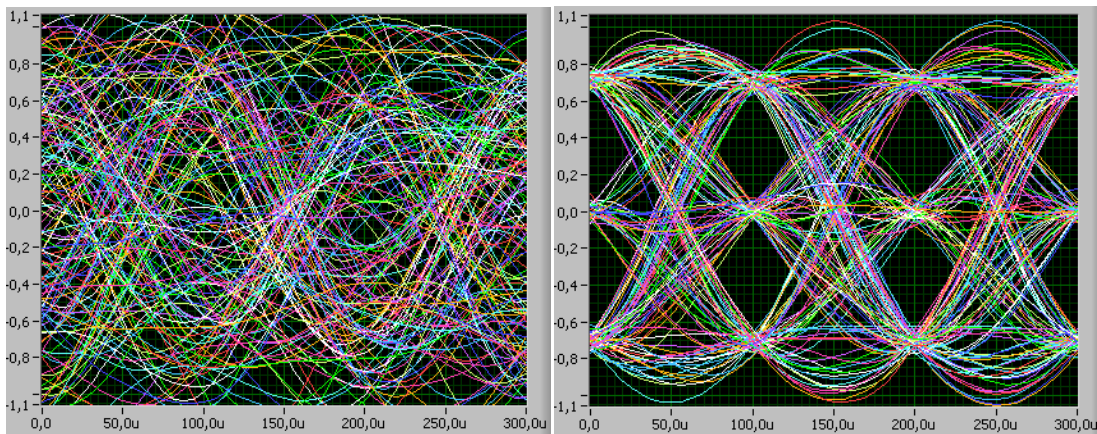


Рис. 1.79. Глазковые диаграммы при малом отношении сигнал/шум и при наилучшем отношении сигнал/шум

На рисунке 1.80. приведены глазковые диаграммы для QAM-16 сигнала.

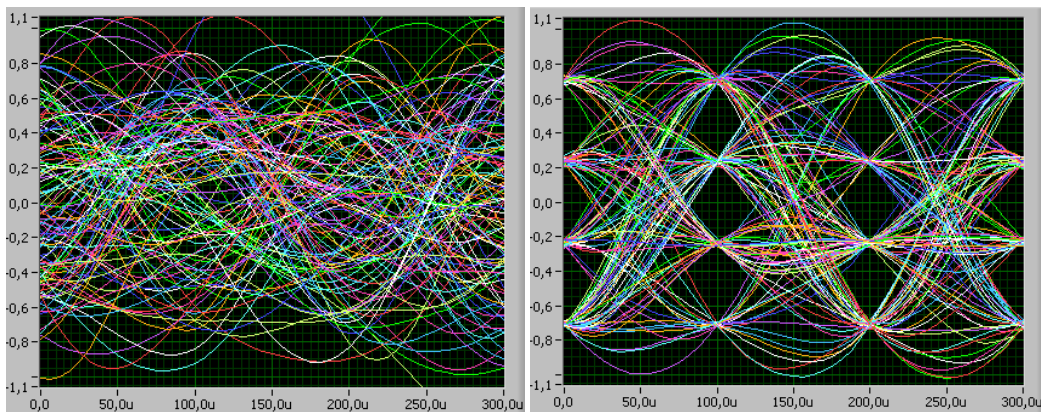


Рис. 1.81. Глазковые диаграммы при малом отношении сигнал/шум и при наилучшем отношении сигнал/шум

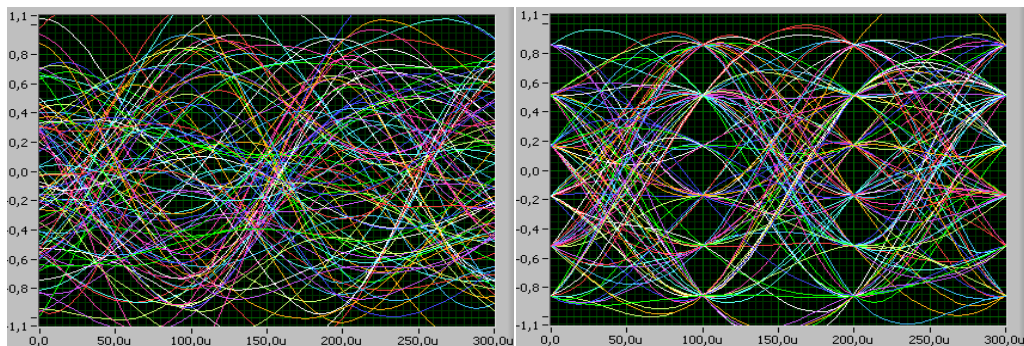


Рис. 1.82. Глазковые диаграммы при малом отношении сигнал/шум и при наилучшем отношении сигнал/шум

На рисунке 1.83 приведены спектрограммы для QAM-8 сигнала.

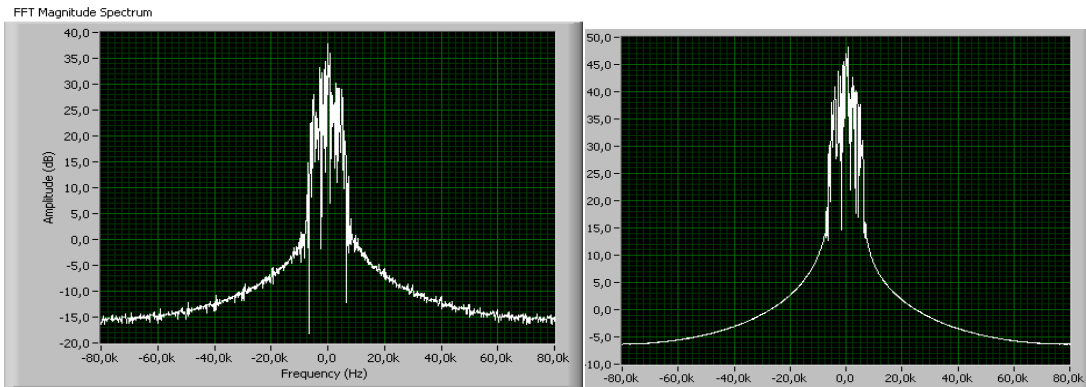


Рис. 1.83. Спектрограммы на входе и на выходе канала

На рисунке 1.84. приведены спектрограммы для QAM-16 сигнала.

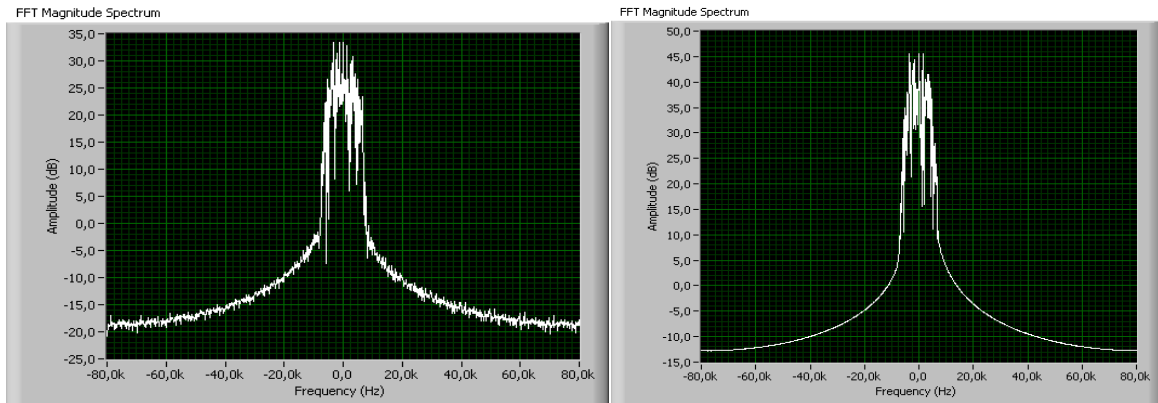


Рис. 1.85. Спектрограммы на входе и на выходе канала

На рисунке 1.86 приведены спектрограммы для QAM-32 сигнала.

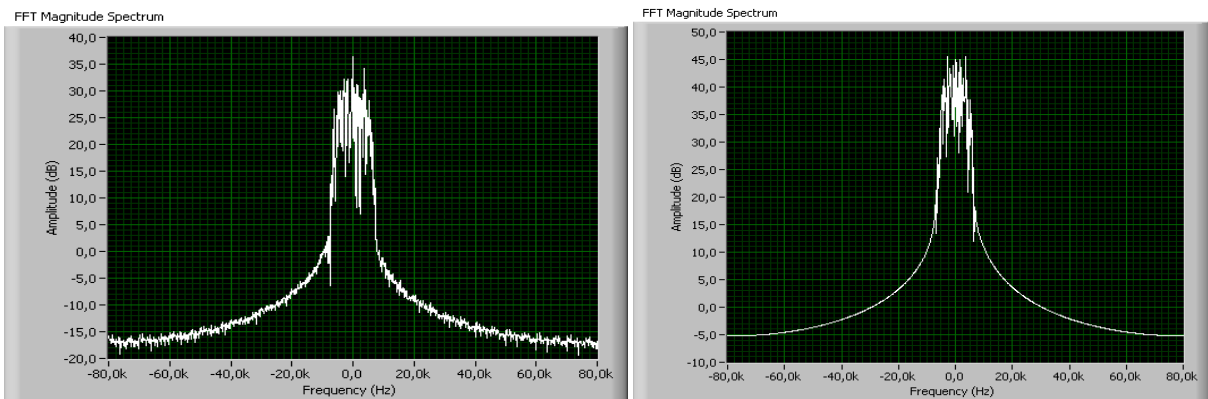


Рис. 1.86. Спектрограммы на входе и на выходе канала

Ползунком E_b/N_0 устанавливается уровень отношения сигнал/шума, в поле BER отображается количество обнаруженных ошибок при передаче. Из полученных данных

можно построить график зависимости, сравнить показатели у различных видов модуляции и подтвердить/опровергнуть теорию, описанную выше.

В таблице ниже представлены данные для рассмотренных видов модуляций.

BER	1	1	1	1	1	1,00E+00	1,19E-02	3,97E-03	7,94E-03	1,98E-03
BER	1	1	1	1	4,51E-01	1,00E+00	1,72E-02	7,81E-03	2,08E-03	2,08E-03
BER	1	1	1	1	1,00E+00	2,27E-01	2,42E-02	2,10E-03	5,25E-04	5,25E-04
Eb/NO, dB	0	2	4	5	6	8	10	12	14	16

На рисунке представлен график зависимости параметра BER от Eb/NO для 8-QAM, 16-QAM, 32-QAM

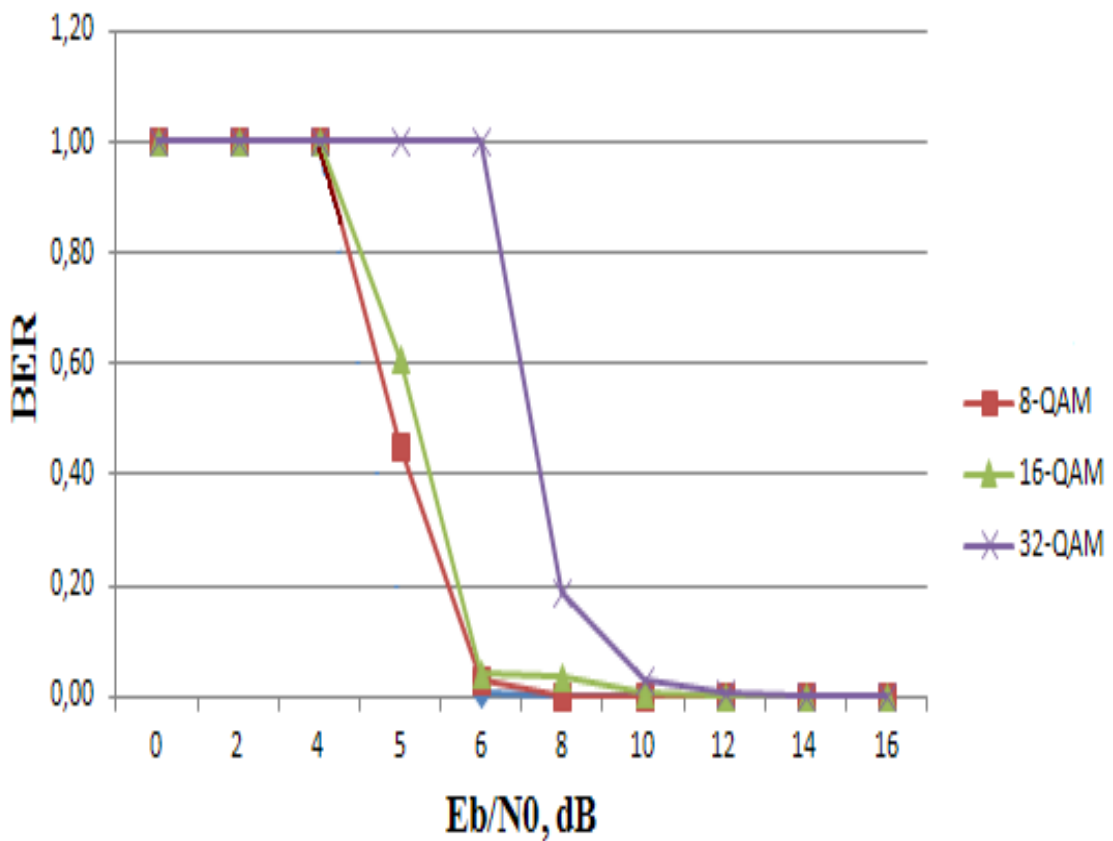


Рис. 1.87. График зависимости параметра BER от Eb/No для 8-QAM, 16-QAM, 32-QAM

ГЛАВА 2. ПРОПУСКНАЯ СПОСОБНОСТЬ КАНАЛА СВЯЗИ. КОДИРОВАНИЕ ИСТОЧНИКА

2.1. Пропускная способность канала связи. Объем сигнала и емкость канала связи, условия их согласования

Рассматривается вопрос согласования дифференциальных характеристик источника дискретной информации (ИДИ) и предоставленного дискретного канала связи (КС) в терминах потока информации, выводимой из ИДИ, и пропускной способности КС. В данном разделе эта задача решается на уровне интегральных характеристик сигнала и канала в виде объема сигнала и емкости канала связи.

Сигнал как модель сообщения (информации) имеет «габаритные размеры», характеризующие объем сигнала, аналитически определяемый выражением

$$V_c = T_c F_c W_c = T_c F_c \log_2 \left(1 + \frac{P_c}{P_n}\right) \quad (2.1)$$

где T_c - временная длительность сигнала, F_c - эффективный спектр сигнала, определяемый эффективным спектра элементарного сигнала кода и типом модуляции. В (2.1) компонент $W_c = \log_2 \left(1 + \frac{P_c}{P_n}\right)$ именуется логарифмическим превышением сигнала над помехой, в котором P_c - мощность сигнала, P_n - мощность помехи, сопровождающей процесс формирования сигнала.

Аналогичным образом канал связи как транспортная среда характеризуется емкостью канала связи, аналитически задаваемую выражением

$$V_k = T_k F_k W_k = T_k F_k \log_2 \left(1 + \frac{P_c}{P_n}\right), \quad (2.2)$$

где T_k - длительность интервала времени, на который предоставлен канал связи, F_k - эффективная полоса пропускания канала связи, которая может быть определена аналитически или экспериментально по амплитудной частотной характеристике четырехполосника, который представляет собой канал связи. В (2.2) компонент $W_k = \log_2 \left(1 + \frac{P_c}{P_n}\right)$ именуется логарифмическим превышением сигнала помехой, в котором P_c - мощность сигнала, фиксируемая в канальной среде, P_n - мощность помехи в канальной среде.

Нетрудно понять, что передача сигнала по предоставленному каналу связи возможна только тогда, когда размеры транспортируемого средства (сигнала) не превышают размеров транспортной среды (канала связи).

Таким образом необходимым условием согласования сигнала с предоставленным каналом связи является выполнение неравенства

$$V_c \leq V_k. \quad (2.3)$$

Достаточными условиями согласования сигнала с предоставленным каналом связи является выполнение неравенств

$$T_c \leq T_k, \quad F_c \leq F_k, \quad W_c \leq W_k, \quad (2.4)$$

И наконец, условием эффективного использования предоставленного канала связи является выполнение равенства

$$V_c = V_k. \quad (2.5)$$

Конструктивным инструментом согласования сигнала с предоставленным каналом связи путем уменьшения объема сигнала V_c за счет уменьшения компонента T_c является использование возможностей эффективного кодирования.

Теорема К.Шеннона об эффективном кодировании символов ИДИ (основная теорема К.Шеннона)

Пусть источник дискретной информации (ИДИ) генерирует алфавит

$X = \{x_i : p(x_i); i = \overline{1, n}\}$ составленный из n дискретных статистически независимые

$$H(X) = - \sum_{i=1}^n p(x_i) \log_2 p(x_i)$$

символов, характеризующийся энтропией , тогда:

1. существует такой способ кодирования символов, при котором среднее на символ

$$l_{cp}[K(X)] = \sum_{i=1}^n p(x_i) l[K(x_i)]$$

число двоичных разрядов кода будет сколь угодно близким к энтропии $H(X)$ ИДИ;

2. не существует такого способа кодирования, при котором среднее на символ число двоичных разрядов кода $l_{cp}[K(X)] = \sum_{i=1}^n p(x_i) l[K(x_i)]$ будет меньше энтропии $H(X)$

ИДИ.

Наиболее распространенными алгоритмами эффективного кодирования, основанного на использовании основной теоремы К.Шеннона, являются:

- Алгоритм К.Шеннона - Р.Фано,
- Алгоритм Д. Хаффмэна.

Следует сказать, что содержательный момент этих алгоритмов, несмотря на процедурные различия, позволяющий построить эффективный двоичный код, средняя на

символ длина которого максимально приближена к энтропии источника, состоит в том, что символам с большей вероятностью появления на выходе ИДИ ставятся в соответствие коды меньшей длины, а символам с меньшей вероятностью - коды большей длины. Таким образом, эффективное кодирование реализуется в классе неравномерных кодов.

Эффективные коды принято характеризовать двумя показателями кода: избыточность и эффективность

Определение 2.1. Избыточностью эффективного кода называется показатель D_k , определяемый выражением

$$D_k = \frac{l_{cp}[K(X)] - l_{\min}[K(X)]}{l_{cp}[K(X)]} = \frac{\sum_{i=1}^n p(x_i) l[K(x_i)] - H(X)}{\sum_{i=1}^n p(x_i) l[K(x_i)]}. \quad (2.6)$$

Определение 2.2. Эффективностью эффективного кода называется показатель η , определяемый выражением

$$\eta = \frac{l_{\min}[K(X)]}{l_{cp}[K(X)]} = \frac{H(X)}{\sum_{i=1}^n p(x_i) l[K(x_i)]}. \quad (2.7)$$

Нетрудно видеть, что избыточность и эффективность эффективного кода связаны соотношением

$$D_k = 1 - \eta. \quad (2.8)$$

2.2. Исследование кодирования источника дискретных сообщений методом

Шеннона Фано

Для удобства расположим все имеющиеся n букв в один столбик в порядке убывания вероятностей. Затем все эти буквы следует разбить на две группы – верхнюю и нижнюю – так, чтобы суммарная вероятность первой группы была наиболее близка к суммарной вероятности второй группы. Для букв первой группы в качестве первой цифры кодового обозначения используется цифра 1, а для букв второй группы – цифра 0. Далее, каждую из двух групп подобным образом снова надо разделить на две части и в качестве второй цифры кодового обозначения мы будем использовать цифру 1 или 0 в зависимости от того, принадлежит ли наша группа к первой или ко второй из этих подгрупп. Затем, каждая из содержащих более одной буквы групп снова делится на две части возможно более близкой суммарной вероятности и т.д.; процесс повторяется до тех пор, пока мы не придем к группам, каждая из которых содержит по одной единственной букве.

Например, если наш алфавит содержит всего шесть букв, вероятность которых (в порядке убывания) равны 0,4, 0,2, 0,2, 0,1, 0,05 и 0,05, то на первом этапе деления букв на группы мы отщепим лишь одну первую букву (1-я группа), оставив во второй группе все остальные. Далее, вторая буква составит 1-ю подгруппу 2-й группы; 2-я же подгруппа той же группы, состоящая из оставшихся четырех букв, будет и далее последовательно делиться на части так, что каждый раз 1-я часть будет состоять лишь из одной буквы.

Таблица 2.1.

№ буквы	вероятность	разбиение на подгруппы (римские цифры обозначают номера групп и подгрупп)					кодировочное обозначение	
1	0,4	} 1					1	
2	0,2	} 0	} 1				01	
3	0,2		} 1				001	
4	0,1		} 0	} 0	} 1			0001
5	0,05				} 0	} 0	} 1	
6	0,05						} 0	

Основной принцип, положенный в основу кодирования по методу Шеннона – Фано, заключается в том, что при выборе каждой цифры кодировочного обозначения мы стараемся, чтобы содержащееся в ней количество информации было наибольшим, т. е. чтобы независимо от значений всех предыдущих цифр, эта цифра принимала оба возможных для нее значения 0 и 1 по возможности с одинаковой вероятностью.

Процесс декодирования

Теперь рассмотрим алгоритм декодирования кодов Шеннона-Фано. Процесс усложняется тем, что невозможно, как в случае кодирования, заменять каждые 8 бит входного потока, кодом переменной длины. При восстановлении исходной последовательности необходимо провести обратные операции - заменить код переменной длины символом длиной 8 бит. В данном случае, лучше всего будет использовать бинарное дерево, ячейками которого будут являться символы.

Пример кодирования сообщений

Для наглядной иллюстрации алгоритма, сделаем сжатие предложения "Карлсон, который живет на крыше.". Вычислим, сколько раз встречается каждый символ, и занесем данные в таблицу 2.2.

Таблица 2.2.

Символ	Количество появлений
Пр	4
обел	3
р	3
о	2
к	2
а	2
е	2
т	2
ы	2
н	1
К	1
й	1
л	1
в	1
и	1
ш	1
ж	1
с	1
.	1
,	

Далее, разделив таблицу на две части, таким образом, чтобы число появлений символов в верхней половине таблицы примерно равнялось общему числу в нижней половине. В предложении 32 символа, а значит, верхняя часть таблицы должна содержать около 16 появлений. Мы можем получить этот результат, поместив разделительную линию (далее р.л.) между строками "е" и "т". В итоге верхняя часть будет содержать появление 16 символов, а нижняя, соответственно, оставшиеся 16.

Далее проделаем то же с каждой из частей таблицы: размесим линии так, чтобы разделить 2 полученные части на приблизительно равные по количеству появлений. Продолжаем процесс деления до того момента, пока каждый символ не будет отделен от другого.

Результирующее дерево.

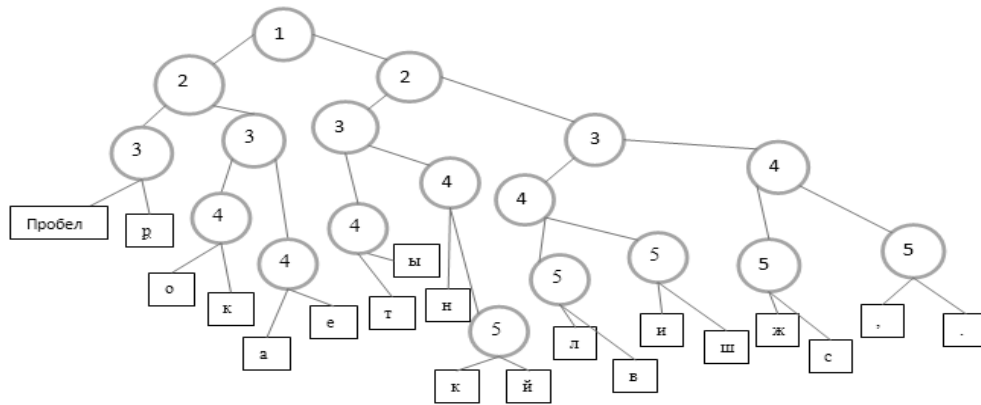


Рис.2.1. Результирующее дерево

Если принять, что перемещение влево эквивалентно нулевому биту, а вправо - единичному, можно создать таблицу кодирования.

Таблица 2.3

Символ	Кодировка
Пробел	000
р	001
о	0100
к	0101
а	0110
е	0111
т	1000
ы	1001
н	1010
К	1011
й	11000
л	11001
в	11010
и	11011
ш	11100
ж	11101
с	11110
.	11111
,	

Таблица содержит 137 бит, а оригинальная фраза "Карлсон, который живет на крыше." занимает 256 бит, следовательно, у нас получается коэффициент сжатия 53%.

Практическая часть

1. Запустить программный комплекс [6].

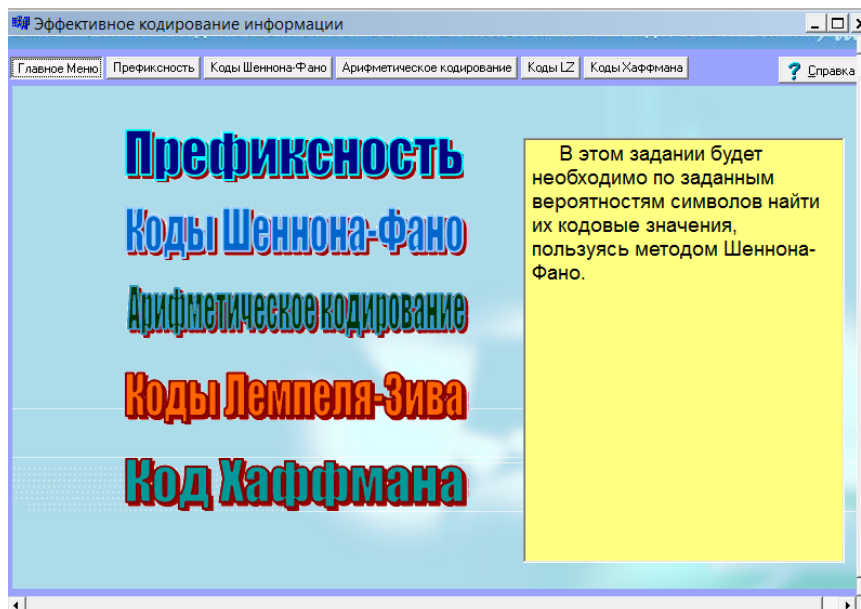


Рис. 2.2. Панель программного комплекса

2. Выбрали в меню пункт задания «Коды Шеннона-Фано». Изучили предложенную таблицу 16-символьного алфавита. Символы отсортированы в порядке убывания вероятностей их появления.

3. Пользуясь теоретическим материалом, найти двоичные коды для каждого из предложенных символов, записали их в графу Код кодовой таблицы. Следуя правилу приписывания очередного символа, указанному в окне программы над кодовой таблицей.

Закодировать приведенные символы алфавита кодом Шеннона-Фано. При разбиении суммы вероятностей в каждой из групп должны быть максимально одинаковыми.
При разбиении на группы всем символам верхней половины в качестве первого бита приписывают 0, а всем нижним - 1

Символ	Вероятность	Код
А	0,11	000
Б	0,09	0010
В	0,08	0011
Г	0,08	010
Д	0,07	0110
Е	0,07	0111
Ж	0,07	1000
З	0,07	1001
И	0,06	1010
Й	0,06	1011
К	0,06	1100
Л	0,05	1101
М	0,05	1110
Н	0,04	11110
О	0,03	111110
П	0,01	111111

Проверить таблицу

Рис. 2.3. Кодовая таблица для данного сообщения

Закодируйте приведенные символы алфавита кодом Шеннона-Фано. При разбиении суммы вероятностей в каждой из групп должны быть максимально одинаковыми.

При разбиении на группы всем символам верхней половины в качестве первого бита приписывайте 0, а всем нижним - 1

Символ	Вероятность	Код
А	0,11	000
Б	0,09	0010
В	0,08	0011
Г	0,08	010
Д	0,07	0110
Е	0,07	0111
Ж	0,07	1000
З	0,07	1001
И	0,06	1010
Й	0,06	1011
К	0,06	1100
Л	0,05	1101
М	0,05	1110
Н	0,04	11110
О	0,03	111110
П	0,01	111111

Подсказка: Вы можете проверять результат построчно(посимвольно), для этого после каждой введенной вами строки (кода символа) нажимайте "Проверить таблицу". Программа сотрет строки результат которых введен не правильно.

Лабораторная работа

Верно! Вы отлично справились с заданием, продолжайте в том же духе.

OK

Проверить таблицу

Рис. 2.4. Процесс проверки кодовой таблицы

Закодируйте приведенные символы алфавита кодом Шеннона-Фано. При разбиении суммы вероятностей в каждой из групп должны быть максимально одинаковыми.

При разбиении на группы всем символам верхней половины в качестве первого бита приписывайте 0, а всем нижним - 1

Символ	Вероятность	Код
А	0,11	000
Б	0,09	0010
В	0,08	0011
Г	0,08	010
Д	0,07	0110
Е	0,07	0111
Ж	0,07	1000
З	0,07	1001
И	0,06	1010
Й	0,06	1011
К	0,06	1100
Л	0,05	1101
М	0,05	1110
Н	0,04	11110
О	0,03	111110
П	0,01	111111

Подсказка: Вы можете проверять результат построчно(посимвольно), для этого после каждой введенной вами строки (кода символа) нажимайте "Проверить таблицу". Программа сотрет строки результат которых введен не правильно.

Декодируйте приведенную последовательность при помощи таблицы кода Шеннона-Фано.

001110100110011111110

Результат:

Проверить

Рис. 2.5. Закодированная последовательность

4. Декодировали при помощи составленной таблицы предложенной последовательности символов. Ответ записали в поле результат. По завершении проверили.

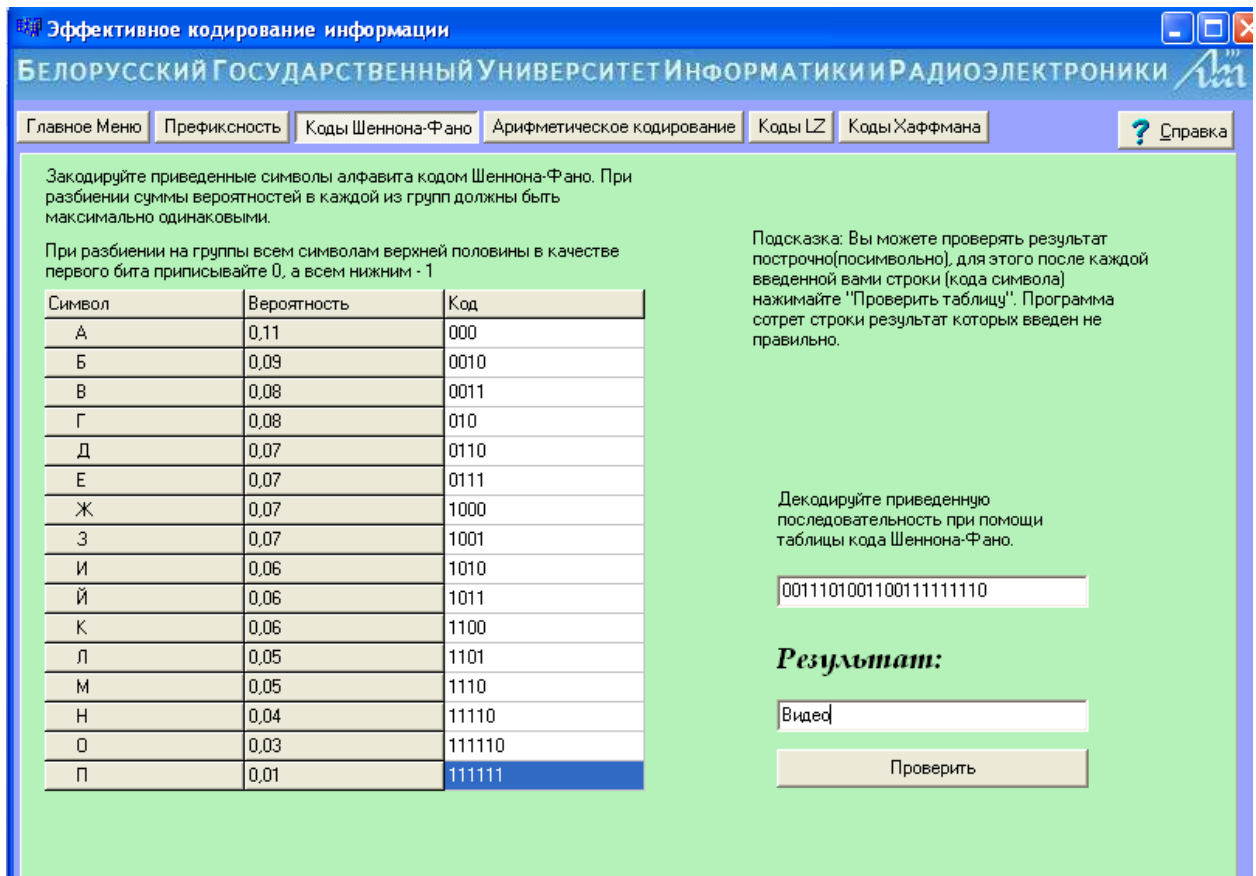


Рис. 2.6. Декодирование закодированной последовательности

Полученную последовательность в соответствии с таблицей символов декодировали самостоятельно и в результате получили слово: «Видео».

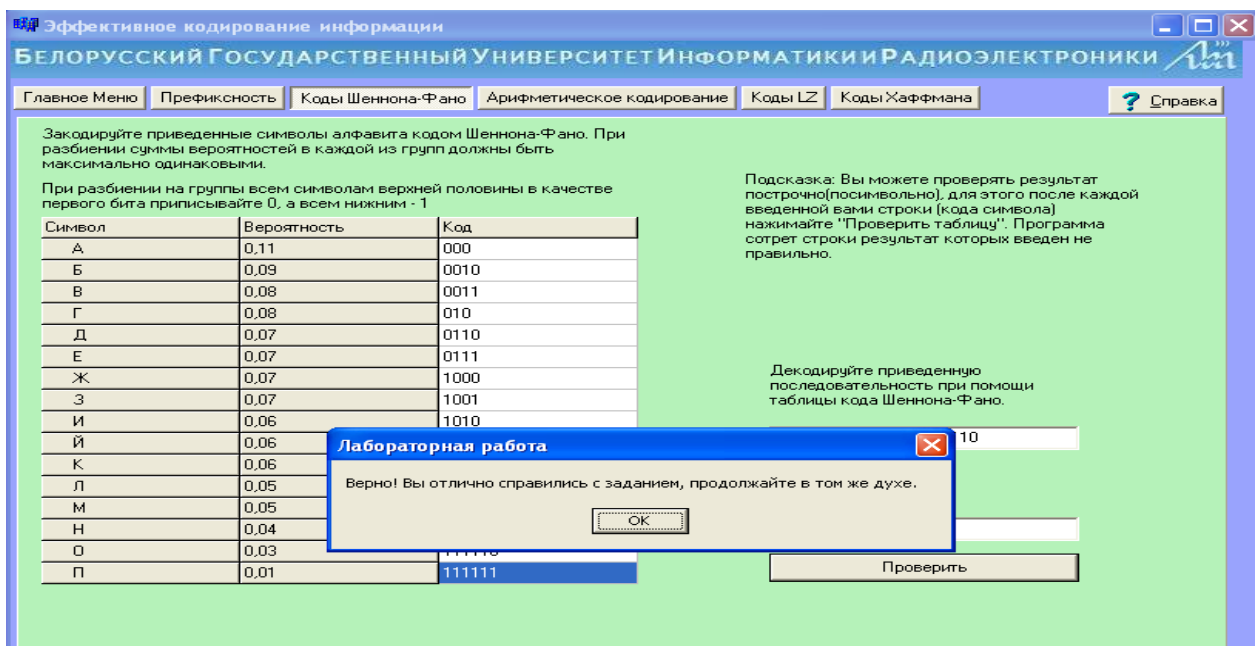


Рис. 2.7. Результат проверки

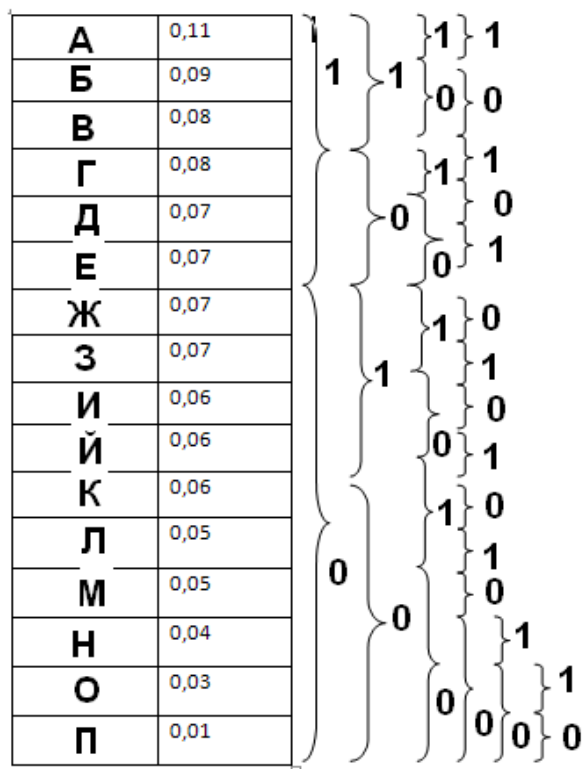


Рис. 2.8. Кодовое дерево

Преимуществом данного метода является его простота и, как следствие этого, высокая скорость кодирования и декодирования. Данный метод является простым в понимании, в реализации, а как следствие довольно эффективным. В настоящее время не используется, так как был усовершенствован Хаффманом, алгоритм которого в данный момент широко используется в связках с другими алгоритмами при кодировании/декодировании информации.

2.3. Исследование алгоритма Лемпеля - Зива

Основная идея алгоритма Лемпеля-Зива состоит в замене появления фрагмента в данных (группы байт) ссылкой на предыдущее появление этого фрагмента. Существуют два основных класса методов, названных в честь Якоба Зива (Jakob Ziv) и Абрахама Лемпеля (Abraham Lempel), впервые предложивших их в 1977 (алгоритм LZ77) и 1978 годах (алгоритм LZ78).

Алгоритм LZ77

LZ77 использует уже просмотренную часть сообщения как словарь. Чтобы добиться сжатия, он пытается заменить очередной фрагмент сообщения на указатель в содержимое словаря.

В качестве модели данных LZ77 использует “скользящее” по сообщению окно, разделенное на две неравные части. Первая, большая по размеру, включает уже

просмотренную часть сообщения. Вторая, намного меньшая, является буфером, содержащим еще не закодированные символы входного потока. Обычно размер окна составляет несколько килобайтов. Буфер намного меньше, обычно не более ста байтов. Алгоритм пытается найти в словаре фрагмент, совпадающий с содержимым буфера.

Алгоритм LZ77 выдает коды, состоящие из трех элементов:

- смещение в словаре относительно его начала подстроки, совпадающей с содержимым буфера;
- длина подстроки;
- первый символ в буфере, следующий за подстрокой.

Пример. Размер окна – 20 символов, словаря – 12 символов, а буфера -8. Кодировается сообщение «ПРОГРАММНЫЕ ПРОДУКТЫ ФИРМЫ MICROSOFT». Пусть словарь уже заполнен. Тогда он содержит строку «ПРОГРАММНЫЕ_», а буфер строку «ПРОДУКТЫ». Просматривая словарь, алгоритм обнаружит, что совпадающей подстрокой будет «ПРО», в словаре она расположена со смещением 0 и имеет длину 3 символа, следующим символом в буфере является «Д». Таким образом, выходным кодом будет тройка (0, 3, «Д»). После этого алгоритм сдвигает влево все содержимое окна на длину совпадающей подстроки +1 и одновременно считывает столько же символов из входного потока в буфер. Получаем в словаре строку «РАММНЫЕ ПРОД», в буфере – «УКТЫ ФИР». В данной ситуации совпадающей подстроки обнаружить не удастся и алгоритм выдаст код (0, 0, «У»), после чего сдвинет окно на один символ. Затем словарь будет содержать «АММНЫЕ ПРОДУ», а буфер – «КТЫ ФИРМ». и т.д.

Проблемы LZ77

Очевидно, что быстродействие кодера LZ77 сильно зависит от того, каким образом будет осуществляться поиск совпадающей подстроки в словаре. Если искать совпадение полным перебором всех возможных вариантов, то очевидно, что сжатие будет очень медленным. Причем при увеличении размеров окна для повышения степени сжатия скорость работы будет пропорционально уменьшаться. Для декодера это неважно, так как при декодировании не осуществляется поиск совпадения.

Быстродействие и кодера, и декодера зависит от того, как реализовано “скольжение” окна по содержимому сообщения. Рационально было бы для работы с окном пользоваться кольцевым буфером, организованным на физически сплошном участке памяти, а не реальным сдвигом содержимого окна. Хотя для поддержания кольцевого буфера необходимы дополнительные затраты на сохранение целостности индексов в нем, в целом это дает очень существенный выигрыш потому, что отсутствует постоянное сдвигание большого блока памяти. Если размер кольцевого буфера равен степени двойки, то

стандартная для кольцевого буфера операция “смещение по модулю размер”, может быть заменена побитовой логической операцией, что еще больше повышает эффективность.

Помимо проблем с быстродействием, у алгоритма LZ77 возникают проблемы с самим сжатием. Они появляются, когда кодер не может найти совпадающую подстроку в словаре и выдает стандартный 3-компонентный код, пытаясь закодировать один символ. Если словарь имеет длину $4K$, а буфер — 16 байтов, то код $\langle 0, 0, \text{символ} \rangle$ будет занимать 3 байта. Кодирование одного байта в три имеет мало общего со сжатием и существенно понижает производительность алгоритма.

Упрощенный алгоритм LZ77

пока (буфер_предпросмотра не пуст) найти наиболее длинное соответствие (позиция_начала, длина) в буфере_предыстории и в буфере_предпросмотра; если (длина > минимальной_длины), то вывести в кодированные данные пару (позиция_начала, длина); иначе вывести в кодированные данные первый символ буфера_предпросмотра; изменить указатель на начало буфера_предпросмотра и продолжить.

Алгоритм LZ78

Этот алгоритм генерирует на основе входных данных словарь фрагментов, внося туда фрагменты данных (последовательности байт) по определенным правилам и присваивает фрагментам коды (номера). При сжатии данных (поступлении на вход программы очередной порции) программа на основе LZ78 пытается найти в словаре фрагмент максимальной длины, совпадающий с данными, заменяет найденную в словаре порцию данных кодом

фрагмента и дополняет словарь новым фрагментом. При заполнении всего словаря (размер словаря ограничен по определению) программа очищает словарь и начинает процесс заполнения словаря снова. Реализации этого метода различаются конструкцией словаря, алгоритмами его заполнения и очистки при переполнении.

Обычно, при инициализации словарь заполняется исходными (элементарными) фрагментами всеми возможными значениями байта от 0 до 255.

Это гарантирует, что при поступлении на вход очередной порции данных будет найден в словаре хотя бы однобайтовый фрагмент.

Алгоритм LZ78 резервирует специальный код, назовем его «Reset», который вставляется в упакованные данные, отмечая момент сброса словаря. Значение кода Reset обычно принимают равным 256.

Таким образом при начале кодирования минимальная длина кода составляет 9 бит. Максимальная длина кода зависит от объема словаря – количества различных фрагментов, которые туда можно поместить. Если объем словаря измерять в байтах (символах), то очевидно, что максимальное количество фрагментов равно числу символов, а,

следовательно, максимальная длина кода равна \log_2 (объем словаря в байтах).

Проведение эксперимента и обработка результатов

Кодирование фразы алгоритмом LZ77 со следующими размерами СЛОВАРЬ/БУФЕР

Таблица 2.3. Варианты расчетных заданий

ФРАЗА	СЛОВАРЬ/БУФЕР			
«IDI TYDA, NE ZNAU KUDA»	8/5	12/5	12/8	16/8

Используем ввод фразы с клавиатуры в программе LZ77.EXE для разных размеров словаря/буфера:

```

C:\DOCUME~1\Major\0016-1\1A1\11__3_-1\Z\Z77.EXE
Алгоритм Лемпела - Зива (LZ77)
Введите входной поток :IDI TYDA, NE ZNAU KUDA
Введите размер словаря:8
Введите размер буфера :5

!Словарь !Буфер!Код
!-----!IDI T!<0,0,"I">!
!-----!IDI TY!<0,0,"D">!
!-----!ID! I TYD!<6,1," " >!
!-----!IDI !TYDA,!<0,0,"T">!
!----!IDI T!YDA, !<0,0,"Y">!
!--!IDI TY!DA, N!<3,1,"A">!
!IDI TYDA!, NE !<0,0," ">!
!DI TYDA, ! NE Z!<2,1,"N">!
! TYDA, N!E ZNA!<0,0,"E">!
!TYDA, NE! ZNAU!<5,1,"Z">!
!DA, NE Z!NAU K!<4,1,"A">!
!, NE ZNA!U KUD!<0,0,"U">!
! NE ZNAU! KUDA!<0,1,"K">!
!E ZNAU K!UDA--!<5,1,"D">!
!ZNAU KUD!A----!<0,0,"A">!
Длина кода исходного сообщения:176
Длина кода полученного сообщения:210
  
```

Рис. 2.9. Ввод фразы с клавиатуры для размеров 8/5

```

C:\DOCUME~1\Major\0016-1\1A1\11__3_-1\Z\Z77.EXE
Алгоритм Лемпела - Зива (LZ77)
Введите входной поток :IDI TYDA, NE ZNAU KUDA
Введите размер словаря:12
Введите размер буфера :5

!Словарь !Буфер!Код
!-----!IDI T!<0,0,"I">!
!-----!I !DI TY!<0,0,"D">!
!-----!ID! I TYD!<10,1," " >!
!-----!IDI !TYDA,!<0,0,"T">!
!-----!IDI T!YDA, !<0,0,"Y">!
!----!IDI TY!DA, N!<7,1,"A">!
!--!IDI TYDA!, NE !<0,0," ">!
!--!IDI TYDA, ! NE Z!<6,1,"N">!
!IDI TYDA, N!E ZNA!<0,0,"E">!
!IDI TYDA, NE! ZNAU!<3,1,"Z">!
!I TYDA, NE Z!NAU K!<8,1,"A">!
!TYDA, NE ZNA!U KUD!<0,0,"U">!
!YDA, NE ZNAU! KUDA!<4,1,"K">!
!A, NE ZNAU K!UDA--!<9,1,"D">!
! NE ZNAU KUD!A----!<0,0,"A">!
Длина кода исходного сообщения:176
Длина кода полученного сообщения:225
  
```

Рис. 2.10. Ввод фразы с клавиатуры для размеров 12/5

```

C:\DOCUME~1\Major\0016-1\1A1\11__3_-1\Z\Z77.EXE
Алгоритм Лемпела - Зива <LZ??>
Введите входной поток :IDI TYDA, NE ZNAU KUDA
Введите размер словаря:12
Введите размер буфера :8

!Словарь      !Буфер      !Код      !
!-----!
!IDI TYDA, NE ZNAU KUDA!<0,0,"I">!
!IDI TYDA, NE ZNAU KUDA!<0,0,"D">!
!IDI TYDA, NE ZNAU KUDA!<10,1," ">!
!IDI TYDA, NE ZNAU KUDA!<0,0,"T">!
!IDI TYDA, NE ZNAU KUDA!<0,0,"Y">!
!IDI TYDA, NE ZNAU KUDA!<7,1,"A">!
!IDI TYDA, NE ZNAU KUDA!<0,0," ">!
!IDI TYDA, NE ZNAU KUDA!<6,1,"N">!
!IDI TYDA, NE ZNAU KUDA!<0,0,"E">!
!IDI TYDA, NE ZNAU KUDA!<3,1,"Z">!
!IDI TYDA, NE ZNAU KUDA!<8,1,"A">!
!IDI TYDA, NE ZNAU KUDA!<0,0,"U">!
!IDI TYDA, NE ZNAU KUDA!<4,1,"K">!
!IDI TYDA, NE ZNAU KUDA!<9,1,"D">!
!IDI TYDA, NE ZNAU KUDA!<0,0,"A">!
Длина кода исходного сообщения:176
Длина кода полученного сообщения:240

```

Рис. 2.11. Ввод фразы с клавиатуры для размеров 12/8

```

C:\DOCUME~1\Major\0016-1\1A1\11__3_-1\Z\Z77.EXE
Алгоритм Лемпела - Зива <LZ??>
Введите входной поток :IDI TUDA, NE ZNAU KUDA
Введите размер словаря:16
Введите размер буфера :8

!Словарь      !Буфер      !Код      !
!-----!
!IDI TUDA, NE ZNAU KUDA!<0,0,"I">!
!IDI TUDA, NE ZNAU KUDA!<0,0,"D">!
!IDI TUDA, NE ZNAU KUDA!<14,1," ">!
!IDI TUDA, NE ZNAU KUDA!<0,0,"T">!
!IDI TUDA, NE ZNAU KUDA!<0,0,"U">!
!IDI TUDA, NE ZNAU KUDA!<11,1,"A">!
!IDI TUDA, NE ZNAU KUDA!<0,0," ">!
!IDI TUDA, NE ZNAU KUDA!<10,1,"N">!
!IDI TUDA, NE ZNAU KUDA!<0,0,"E">!
!IDI TUDA, NE ZNAU KUDA!<7,1,"Z">!
!IDI TUDA, NE ZNAU KUDA!<12,1,"A">!
!IDI TUDA, NE ZNAU KUDA!<5,1," ">!
!IDI TUDA, NE ZNAU KUDA!<0,0,"K">!
!IDI TUDA, NE ZNAU KUDA!<0,0,"U">!
Длина кода исходного сообщения:176
Длина кода полученного сообщения:224

```

Рис. 2.12. Ввод фразы с клавиатуры для размеров 16/8

Используем ввод текста из файла для разных размеров словаря/буфера:

Текст: «Полуадаптированное моделирование решает эту проблему, используя для каждого текста свою модель, которая строится еще до самого сжатия на основании результатов предварительного просмотра текста (или его образца). Перед тем, как окончено формирование сжатого текста, модель должна быть передана декодировщику. Несмотря на дополнительные затраты по передаче модели, эта стратегия в общем случае окупается благодаря лучшему соответствию модели тексту.

Адаптированное (или динамическое) моделирование уходит от связанных с этой передачей расходов. Первоначально и кодировщик, и декодировщик присваивают себе некоторую пустую модель, как если бы символы все были равновероятными. Кодировщик

использует эту модель для сжатия очередного символа, а декодировщик - для его разворачивания. Затем они оба изменяют свои модели одинаковым образом (например, наращивая вероятность рассматриваемого символа). Следующий символ кодируется и достается на основании новой модели, а затем снова изменяет модель. Кодирование продолжается аналогичным декодированию образом, которое поддерживает идентичную модель за счет применения такого же алгоритма ее изменения, обеспеченным отсутствием ошибок во время кодирования. Используемая модель, которую к тому же не нужно передавать явно, будет хорошо соответствовать сжатому тексту.

Адаптированные модели представляют собой элегантную и эффективную технику, и обеспечивают сжатие по крайней мере не худшее производимого неадаптированными схемами. Оно может быть значительно лучше, чем у плохо соответствующих текстам статичных моделей. Адаптированные модели, в отличие от полуадаптированных, не производят их предварительного просмотра, являясь поэтому более привлекательными и лучшесжимаемыми. Т.о. алгоритмы моделей, описываемые в подразделах, при кодировании и декодировании будут выполняться одинаково. Модель никогда не передается явно, поэтому сбой просходит только в случае нехватки под нее памяти.»

```

C:\DOCUME~1\Major\0016-1\1A1\11__3_-1\ZLZ77_F.EXE
тэю, яю"Сюье ёсюц яЕюёЮюфшС Сююфью т ё <0,0,"ь">
эю, яю"Сюье ёсюц яЕюёЮюфшС Сююфью т ё <1,1,"ь">
яю"Сюье ёсюц яЕюёЮюфшС Сююфью т ёме <0,0,"т">
яю"Сюье ёсюц яЕюёЮюфшС Сююфью т ёмеу <8,2,"ь">
"Сюье ёсюц яЕюёЮюфшС Сююфью т ёмеу <4,1,"з">
юье ёсюц яЕюёЮюфшС Сююфью т ёмеу <0,0,"р">
ье ёсюц яЕюёЮюфшС Сююфью т ёмеу <0,0,"х">
е ёсюц яЕюёЮюфшС Сююфью т ёмеу <1,1,"э">
ёсюц яЕюёЮюфшС Сююфью т ёмеу <29,1,"т">
сюц яЕюёЮюфшС Сююфью т ёмеу <20,1,"р">
юц яЕюёЮюфшС Сююфью т ёмеу <9,1,"ь">
юц яЕюёЮюфшС Сююфью т ёмеу <6,1,"ь">
ёЮюфшС Сююфью т ёмеу <0,0,"я">
ёЮюфшС Сююфью т ёмеу <1,2,"ь">
юфшС Сююфью т ёмеу <24,2,"р">
шС Сююфью т ёмеу <0,0,"ь">
юфшС Сююфью т ёмеу <0,0,"ь">
юфшС Сююфью т ёмеу <16,1,"ш">
шС Сююфью т ёмеу <0,0,"ь">
СююфшС Сююфью т ёмеу <0,1,"ш">
юфшС Сююфью т ёмеу <0,0,"ш">
юфшС Сююфью т ёмеу <0,0,"ш">
Длина кода исходного сообщения :15544 буТ <1943 бауТ>
Длина кода полученного сообщения:15984 буТ <1998 бауТ>

```

Рис. 2.13. Ввод текста из файла для размеров 32/5

```

C:\DOCUME~1\Major\0016-1\1A1\11__3_-1\ZLZ77_F.EXE
р эх яхЕхфрхСё тэю, яю"Сюье ёсюц яЕюёЮюфшС Сююфью т ёмеу <17,1,"р">
эх яхЕхфрхСё тэю, яю"Сюье ёсюц яЕюёЮюфшС Сююфью т ёмеу <10,1,"ь">
яхЕхфрхСё тэю, яю"Сюье ёсюц яЕюёЮюфшС Сююфью т ёмеу <38,1,"ь">
хЕхфрхСё тэю, яю"Сюье ёсюц яЕюёЮюфшС Сююфью т ёмеу <16,2,"ф">
фрхСё тэю, яю"Сюье ёсюц яЕюёЮюфшС Сююфью т ёмеу <51,3,"х">
ё тэю, яю"Сюье ёсюц яЕюёЮюфшС Сююфью т ёмеу <8,2,"р">
тэю, яю"Сюье ёсюц яЕюёЮюфшС Сююфью т ёмеу <11,1,"ь">
эю, яю"Сюье ёсюц яЕюёЮюфшС Сююфью т ёмеу <7,1,"ш">
яю"Сюье ёсюц яЕюёЮюфшС Сююфью т ёмеу <0,0,"ь">
яю"Сюье ёсюц яЕюёЮюфшС Сююфью т ёмеу <0,1,"ш">
ю"Сюье ёсюц яЕюёЮюфшС Сююфью т ёмеу <0,0,"ш">
Длина кода исходного сообщения :15544 буТ <1943 бауТ>
Длина кода полученного сообщения:15011 буТ <1876 бауТ>

```

Рис. 2.14. Ввод текста из файла для размеров 64/5

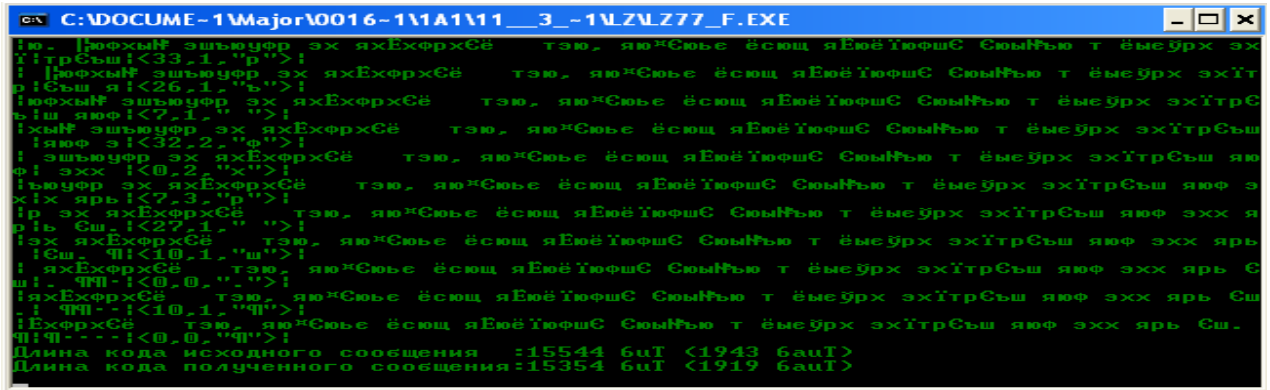


Рис. 2.15. Ввод текста из файла для размеров 80/5



Рис. 2.16. Ввод текста из файла для размеров 128/5

Таблица 2.4. Результаты измерений

D_словарь	32	64	80	128
L _{вх} /L _{вых}	0,972	1,036	1,012	1,141

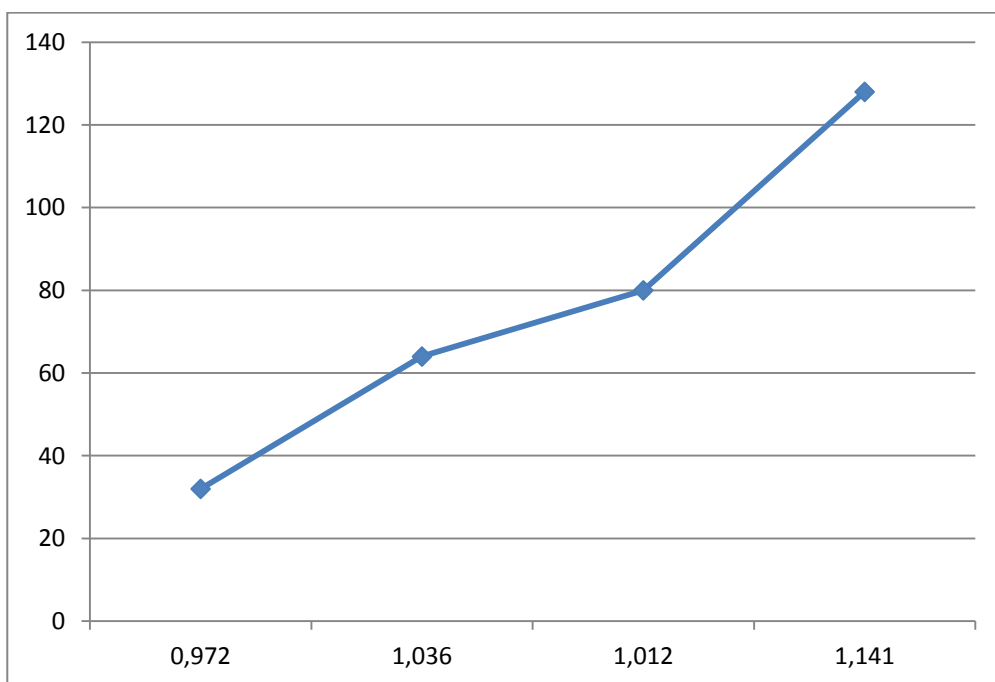


Рис. 2.17. График зависимости коэффициента сжатия от размера словаря
 Кодирование для разной длины текста и одинакового размера словаря 64/5:

```

C:\DOCUMENT-1\Major\0016-1\1A1\11__3_-1\Z\Z77_F.EXE
>| .....Т хёе Еюшрёё! зюю!<50,1,"з"
>| .....Т хёе Еюшрёё! зюю!<49,1,"ю"
>| .....Т хёе Еюшрёё! зюю!<0,0,"з"
>| .....Т хёе Еюшрёё! зюю!<0,0,"з"
>| .....Т хёе Еюшрёё! зюю!<55,1,""
>| .....Т хёе Еюшрёё! зюю!<0,0,"т"
>| .....Т хёе Еюшрёё! зюю!<41,5,""
>| .....Т хёе Еюшрёё! зюю!<42,1,"э"
>| .....Т хёе Еюшрёё! зюю!<53,2,"Е"
>| .....Т хёе Еюшрёё! зюю!<37,1,"ё"
>| .....Т хёе Еюшрёё! зюю!<0,0,"ы"
>|
Длина кода исходного сообщения :312 байт <39 байт>
Длина кода полученного сообщения:391 байт <49 байт>
  
```

Рис. 2.18. Ввод текста длиной 39 байт

```

C:\DOCUMENT-1\Major\0016-1\1A1\11__3_-1\Z\Z77_F.EXE
>| "Ср ёррехуш т юс*хь ёмеурх юеярхСё сыр-уюфЕ меу°хье ёюС!тхСёС!<16,1,"х"
>| "Ср ёррехуш т юс*хь ёмеурх юеярхСё сыр-уюфЕ меу°хье ёюС!тхСёС!<35,2,"С"
>| "ёррехуш т юс*хь ёмеурх юеярхСё сыр-уюфЕ меу°хье ёюС!тхСёС!<11,1,"ш"
>| "ёррехуш т юс*хь ёмеурх юеярхСё сыр-уюфЕ меу°хье ёюС!тхСёС!<0,0,"!"
>| "ёррехуш т юс*хь ёмеурх юеярхСё сыр-уюфЕ меу°хье ёюС!тхСёС!<7,1,"ь"
>| "ехуш т юс*хь ёмеурх юеярхСё сыр-уюфЕ меу°хье ёюС!тхСёС!<36,2,"х"
>| "ш т юс*хь ёмеурх юеярхСё сыр-уюфЕ меу°хье ёюС!тхСёС!<12,1,"ш"
>| "т юс*хь ёмеурх юеярхСё сыр-уюфЕ меу°хье ёюС!тхСёС!<0,1,"С"
>| "юс*хь ёмеурх юеярхСё сыр-уюфЕ меу°хье ёюС!тхСёС!<4,1,"ь"
>| "с*хь ёмеурх юеярхСё сыр-уюфЕ меу°хье ёюС!тхСёС!<48,2,"е"
>| "ь ёмеурх юеярхСё сыр-уюфЕ меу°хье ёюС!тхСёС!<0,0,""
>|
Длина кода исходного сообщения :3616 байт <452 байт>
Длина кода полученного сообщения:3638 байт <455 байт>
  
```

Рис. 2.19. Ввод текста длиной 452 байта

```

C:\DOCUMENT-1\Major\0016-1\1A1\11__3_-1\Z\Z77_F.EXE
>| "р эх яхЕхрхСё тэю, яю"Сюье ёсюц яЕюёЮфшС СюмНью т ёмеурх эх!трСьш!<17,1,"р"
>| "эх яхЕхрхСё тэю, яю"Сюье ёсюц яЕюёЮфшС СюмНью т ёмеурх эх!трСьш я!<10,1,"ь"
>| "яхЕхрхСё тэю, яю"Сюье ёсюц яЕюёЮфшС СюмНью т ёмеурх эх!трСьш яюф!<38,1,""
>| "хЕхрхСё тэю, яю"Сюье ёсюц яЕюёЮфшС СюмНью т ёмеурх эх!трСьш яюф э!<16,2,"ф"
>| "рхСё тэю, яю"Сюье ёсюц яЕюёЮфшС СюмНью т ёмеурх эх!трСьш яюф! эхх!<51,3,"х"
>| "ё тэю, яю"Сюье ёсюц яЕюёЮфшС СюмНью т ёмеурх эх!трСьш яюф эхх! ярь!<8,2,"р"
>| "тэю, яю"Сюье ёсюц яЕюёЮфшС СюмНью т ёмеурх эх!трСьш яюф эхх! ярь!<11,1,""
>| "эю, яю"Сюье ёсюц яЕюёЮфшС СюмНью т ёмеурх эх!трСьш яюф эхх! ярь!<7,1,"ш"
>| "яю"Сюье ёсюц яЕюёЮфшС СюмНью т ёмеурх эх!трСьш яюф эхх! ярь!<0,0,""
>| "яю"Сюье ёсюц яЕюёЮфшС СюмНью т ёмеурх эх!трСьш яюф эхх! ярь!<0,1,""
>| "ю"Сюье ёсюц яЕюёЮфшС СюмНью т ёмеурх эх!трСьш яюф эхх! ярь!<0,0,""
>|
Длина кода исходного сообщения :11904 байт <1488 байт>
Длина кода полученного сообщения:11475 байт <1434 байт>
  
```

Рис. 2.20. Ввод текста длиной 1488 байт

```

C:\> C:\DOCUMENTS\Major\0016-111A1111__3_-1\Z\Z77_F.EXE
{>| яояем ЕэЧьш т эрёсю ^хх тЕхь ьхСюфрьш, Срьшьш, эр-яЕшьхЕ, ьрь ёцрСшх !<17,1,"ц
{>| яояем ЕэЧьш т эрёсю ^хх тЕхь ьхСюфрьш, Срьшьш, эр-яЕшьхЕ, ьрь ёцрСшх !<14,1,"С
{>| ем ЕэЧьш т эрёсю ^хх тЕхь ьхСюфрьш, Срьшьш, эр-яЕшьхЕ, ьрь ёцрСшх ||штр!<7,1,"х"
{>| ЕэЧьш т эрёсю ^хх тЕхь ьхСюфрьш, Срьшьш, эр-яЕшьхЕ, ьрь ёцрСшх! ||штр!<6,1,"||"
{>| эЧьш т эрёсю ^хх тЕхь ьхСюфрьш, Срьшьш, эр-яЕшьхЕ, ьрь ёцрСшх ||штр-т!<3,1,"т"
{>| ьш т эрёсю ^хх тЕхь ьхСюфрьш, Срьшьш, эр-яЕшьхЕ, ьрь ёцрСшх ||штр-трь!<40,2,"т"
{>| т эрёсю ^хх тЕхь ьхСюфрьш, Срьшьш, эр-яЕшьхЕ, ьрь ёцрСшх ||штр-т!хьяхм!<14,2,"я
{>| рёсю ^хх тЕхь ьхСюфрьш, Срьшьш, эр-яЕшьхЕ, ьрь ёцрСшх ||штр-трь!хьяхмр. !<6,1,"м"
{>| сю ^хх тЕхь ьхСюфрьш, Срьшьш, эр-яЕшьхЕ, ьрь ёцрСшх ||штр-трь!хьяхмр. !<18,1,"."
{>| ^хх тЕхь ьхСюфрьш, Срьшьш, эр-яЕшьхЕ, ьрь ёцрСшх ||штр-трь!хьяхмр. !<4,1,"!"
{>| хх тЕхь ьхСюфрьш, Срьшьш, эр-яЕшьхЕ, ьрь ёцрСшх ||штр-трь!хьяхмр. !<0,0,"!"
{>|
Длина кода исходного сообщения :34336 байт (4292 байт)
Длина кода полученного сообщения:33388 байт (4174 байт)

```

Рис. 2.21. Ввод текста длиной 4292 байта

Таблица 2.5. Результаты измерений

$L_{ВХ}$	39	452	1488	4292
$L_{ВХ}/L_{ВЫХ}$	0,795	0,993	1,037	1,028

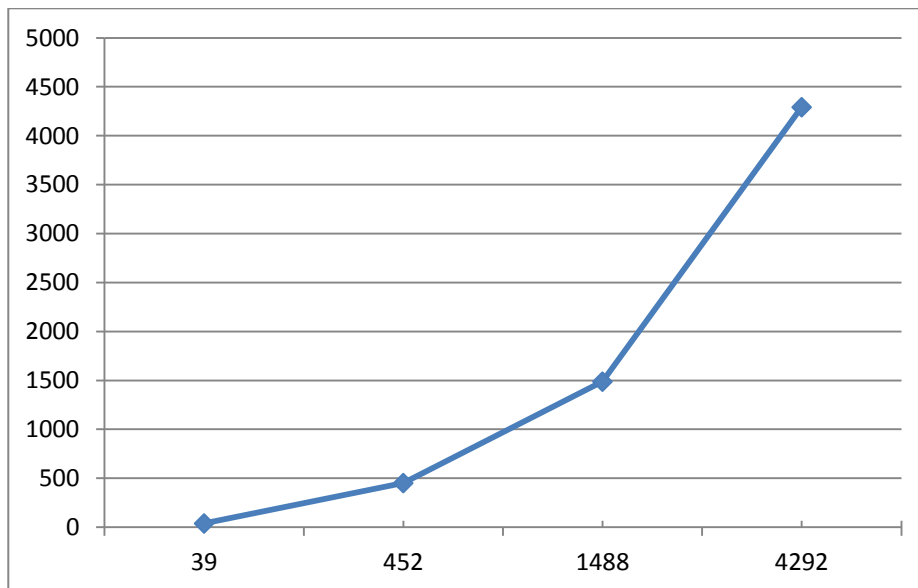


Рис. 2.22. График зависимости коэффициента сжатия от размера словаря

При кодировании коротких текстов коэффициент сжатия оказывается меньше 1, т.е. избыточность не удаляется, а вводится еще больше. При увеличении размера словаря коэффициент сжатия увеличивается. Большой размер буфера также ухудшает коэффициент сжатия. Если размер словаря кратен степеням двойки, то сжатие лучше. Однако, при

большом размере словаря снижается скорость кодирования. При увеличении размера текста коэффициент сжатия увеличивается.

2.4. Фрактальные методы кодирования изображений

Современные компьютеры весьма интенсивно применяют графику. Операционные системы с интерфейсом оконного типа используют картинки, например, для отображения директорий или папок. Некоторые совершаемые системой действия, например загрузку и пересылку файлов, также отображаются графически. Многие программы и приложения предлагают пользователю графический интерфейс (GUI), который значительно упрощает работу пользователя и позволяет легко интерпретировать полученные результаты. Компьютерная графика используется во многих областях повседневной деятельности при переводе сложных массивов данных в графическое представление.

Фрактал (лат. fractus — дроблёный, сломанный, разбитый) — геометрическая фигура, обладающая свойством самоподобия, то есть составленная из нескольких частей, каждая из которых подобна всей фигуре целиком. В математике под фракталами понимают множества точек в евклидовом пространстве, имеющие дробную метрическую размерность (в смысле Минковского или Хаусдорфа), либо метрическую размерность, отличную от топологической [14]. Слово «фрактал» может употребляться не только как математический термин. Фракталом в прессе и научно-популярной литературе могут называть фигуры, обладающие какими-либо из перечисленных ниже свойств: Обладает нетривиальной структурой на всех масштабах. В этом отличие от регулярных фигур (таких, как окружность, эллипс, график гладкой функции): если мы рассмотрим небольшой фрагмент регулярной фигуры в очень крупном масштабе, он будет похож на фрагмент прямой. Для фрактала увеличение масштаба не ведёт к упрощению структуры, на всех шкалах мы увидим одинаково сложную картину.

Является самоподобной или приближённо самоподобной. Обладает дробной метрической размерностью или метрической размерностью, превосходящей топологическую. Многие объекты в природе обладают фрактальными свойствами, например, побережья, облака, кроны деревьев, снежинки, кровеносная система и система альвеол человека или животных.

Фракталы, особенно на плоскости, популярны благодаря сочетанию красоты с простотой построения при помощи компьютера. Первые примеры самоподобных множеств с необычными свойствами появились в XIX веке (например, множество Кантора). Термин «фрактал» был введён Бенуа Мандельбротом в 1975 году и получил широкую популярность с выходом в 1977 году его книги «Фрактальная геометрия природы».

При фрактальном сжатии изображения для области меньшего размера подыскивается похожая на нее область большего размера того же изображения (рисунок 2.23).

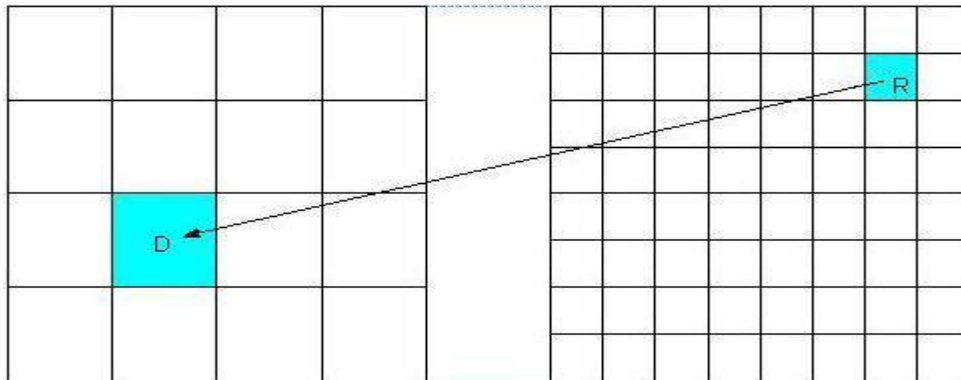


Рис. 2.23. Доменный и ранговый блоки

Область большого размера называется доменным блоком, а меньшего – ранговым блоком. Перевод доменного блока в ранговый осуществляется посредством аффинных преобразований, которые в случае изображения в оттенках серого цвета можно представить следующей системой уравнений (2.9):

$$\begin{bmatrix} x^* \\ y^* \\ z^* \end{bmatrix} = \begin{bmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & u \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} e \\ f \\ v \end{bmatrix}, \quad (2.9)$$

где x , y , z – соответственно координаты и яркость пикселя доменного блока; x^* , y^* , z^* – соответственно координаты и яркость пикселя рангового блока; a , b , c , d , e , f – коэффициенты преобразований координат на плоскости; u – коэффициент сжатия яркости; v – сдвиг по яркости.

В рассматриваемых алгоритмах фрактального сжатия и восстановления изображения введенные упрощения позволяют заменить матричные преобразования операциями изменения ориентации доменного блока и операциями расчета яркости пикселей. Изменения ориентации доменного блока осуществляется за счет его поворота на угол кратный 90° и зеркального отражения и поворота зеркального отражения.

Из всех возможных доменных блоков выбирается блок, ближайший (в выбранной метрике) к рассматриваемому ранговому блоку. Когда доменный блок найден, то запоминаются его номер и параметры преобразования в текущий ранговый блок, из которых и состоит решение задачи фрактального сжатия.

Ниже, на рисунке 2.32, представлен алгоритм фрактального сжатия.



Рис. 2.24. Алгоритм фрактального сжатия

Программное обеспечение в DELPHI 7

Приложение выполняет фрактальное сжатие / распаковку изображений с помощью классического алгоритма.

В нее можно загружать любые изображения, но размер должен быть не более 512x512 пикселей. Программа будет автоматически убирать цвет изображений.

Эти ограничения введены для того, чтобы существенно сократить время сжатия изображений на базе фрактального алгоритма.

Основные характеристики:

- Сжатие и декодирование изображений формата .bmp;
- Просмотр полученного результата;
- Просмотр размер полученного изображения.

На рисунке 2.35. представлено главное и единственное меню программы.

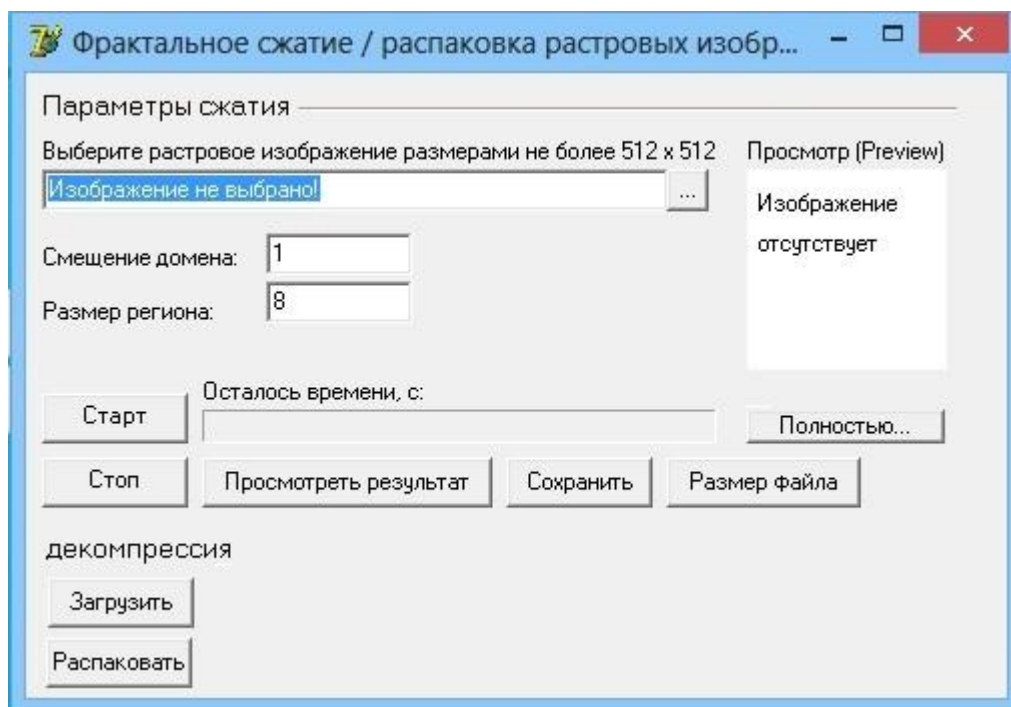


Рис. 2.25. Главное меню программы.

В программе доступны для изменения 2 следующих пункта: смещение домена и размер региона.

Смещение домена: Определяет шаг поиска участка в доменном изображении. Минимальный шаг равен 1. Чем больше шаг, тем быстрее выполняется поиск, но при этом могут быть пропущены важные детали изображения.

Размер региона: определяет размер области, на которое разбивается исходное изображение. При компрессии для каждой области осуществляется поиск подходящего домена с учетом трансформации (аффинных преобразований). Чем больше размер региона, тем хуже качество и при этом уменьшается размер IFS-изображения.

Методика работы в программе и проведение исследования основных технических характеристик системы

Для обеспечения фрактального сжатия была выбрана программа «Фрактальное сжатие / распаковка растровых изображений».

Для того чтобы начать работать с программой необходимо:

1. Загрузить изображение, которое удовлетворяет требованию программы;
2. Выставить значение параметров «Смещение домена» и «Размер региона»;
3. Нажать кнопку старт;

4. Дождаться окончания сжатия изображения;
5. Просмотреть полученный результат и, если он удовлетворил нас, сохранить его.

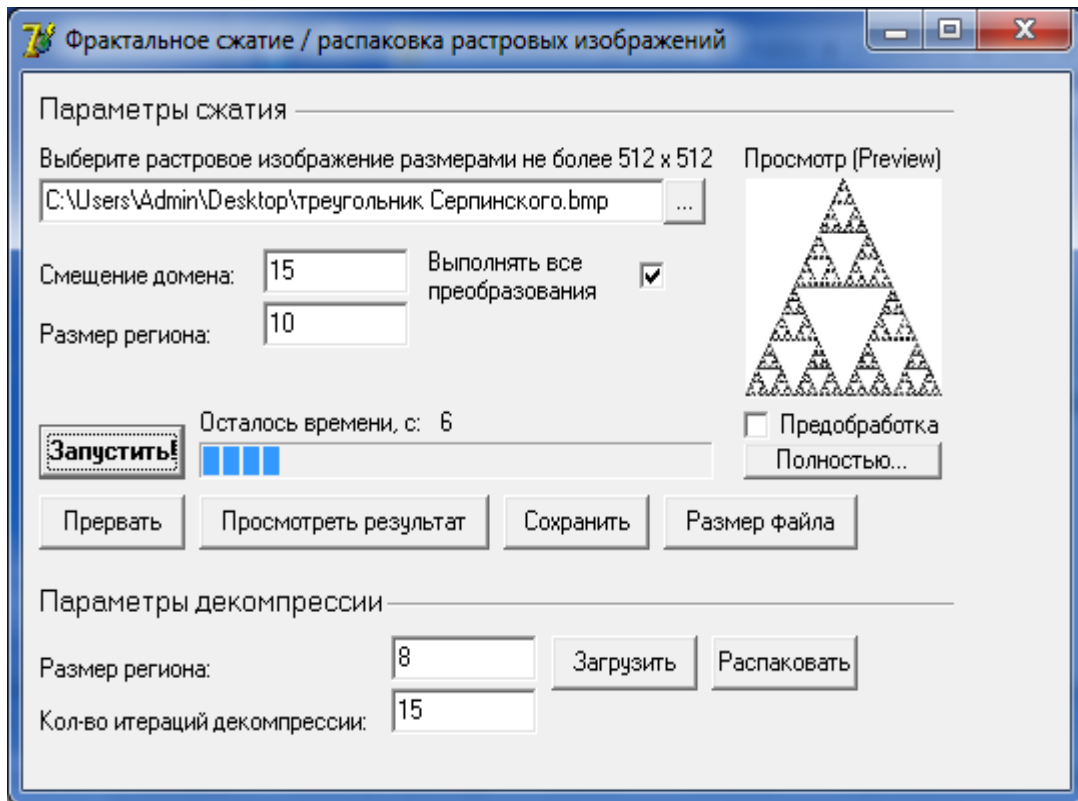


Рис. 2.26. Сжатие изображения в программе «Фрактальное сжатие / распаковка растровых изображений»

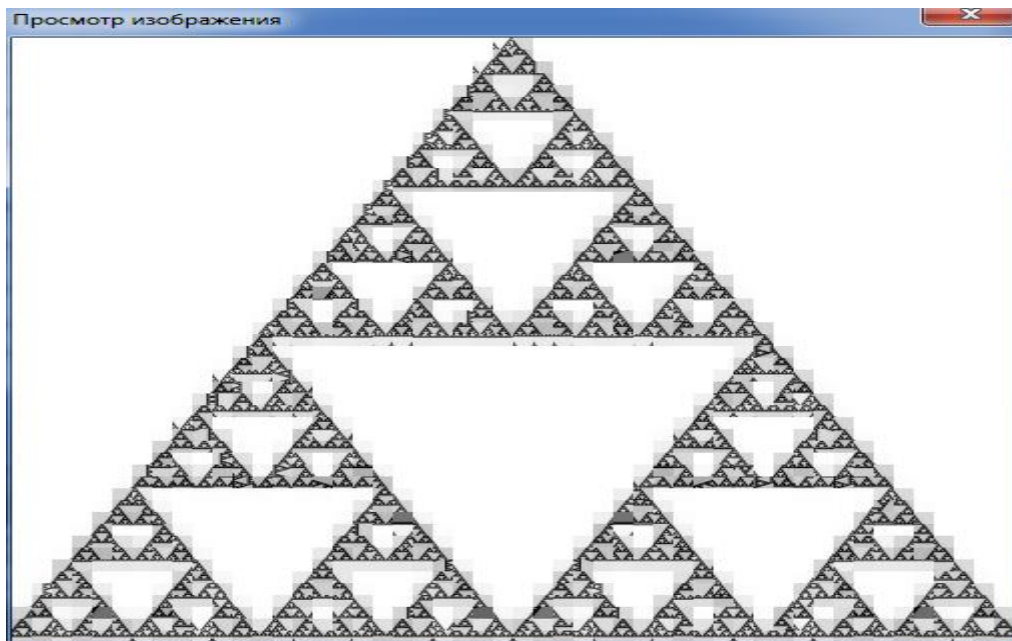


Рис. 2.27. Результат сжатия.

После сжатия изображение будет иметь формат .IFS.

Стандартные кодеки Windows не способны декодировать полученное изображение, поэтому в программе предусмотрена функция просмотра изображений формата .IFS.

Чтобы просмотреть нужное нам изображение необходимо:

1. В меню «декомпрессия» нажать на кнопку «Загрузить»;
2. Выбрать изображение формата .IFS;
3. В меню «декомпрессия» нажать на кнопку «Распаковать»

Далее рассмотрим сжатие изображений со спутника X-SAR Европейского космического агентства. На рисунке 2.36 представлено первоначальное изображение со спутника размером 435 Кб и разрешением 473x314 пикселей.



Рис. 2.28. Первоначальное изображение.

Далее, на рисунке 2.29, представлено это же изображение после обработки, при параметре смещение домена=1 и размер региона=8, время потраченное на сжатие $t=703$ с, размер файла =11,4Кб.



Рис. 2.29. Обработанное изображение.

Далее, на рисунке 2.30, представлено это же изображение после обработки, при параметре смещение домена=10 и размер региона=8, время потраченное на сжатие $t=6$ с, размер файла =11Кб.

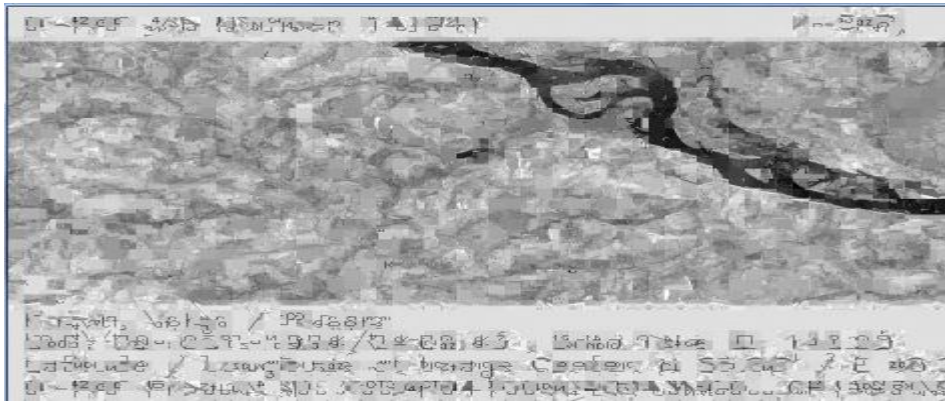


Рис. 2.30. Обработанное изображение.

Далее, на рисунке 2.31, представлено это же изображение после обработки, при параметре смещение домена=5 и размер региона=10, время потраченное на сжатие $t=27$ с, размер файла =7,12Кб.

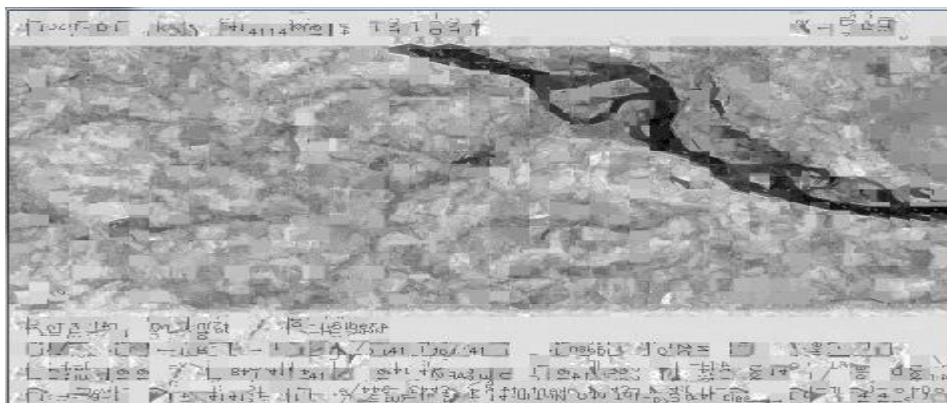


Рис. 2.31. Обработанное изображение.

Далее, на рисунке 2.32, представлено это же изображение после обработки, при параметре смещение домена=10 и размер региона=12, время потраченное на сжатие $t=5$ с, размер файла =4,95Кб.

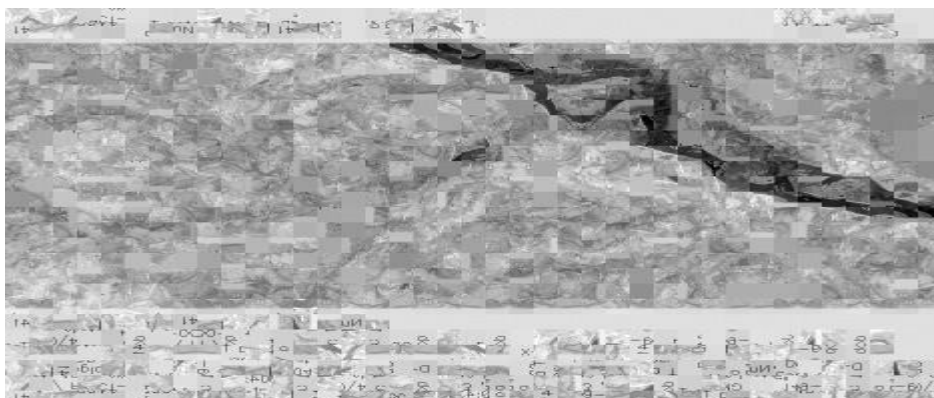


Рис. 2.32. Обработанное изображение.

Далее, на рисунке 2.33, представлено это же изображение после обработки, при параметре смещение домена=1 и размер региона=15, время потраченное на сжатие $t=566$ с, размер файла =3,04Кб.



Рис. 2.33. Обработанное изображение.

Далее, на рисунке 2.34, представлено это же изображение после обработки, при параметре смещение домена=10 и размер региона=14, время потраченное на сжатие $t=4$ с, размер файла =3,55Кб.

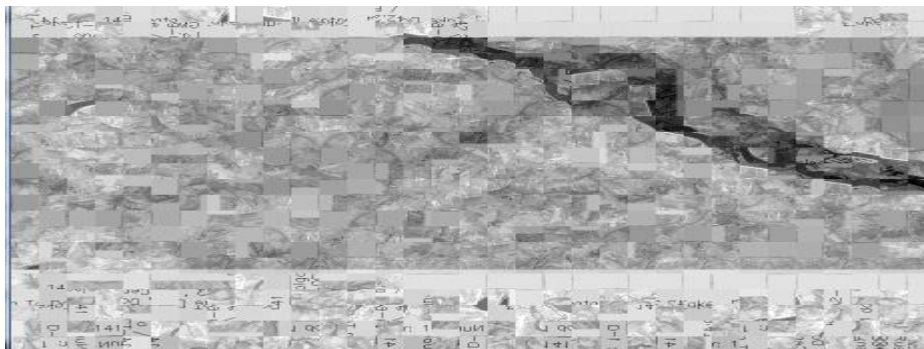


Рис. 2.34. Обработанное изображение.

Далее, на рисунке 2.35, представлено это же изображение после обработки, при параметре смещение домена=20 и размер региона=10, время потраченное на сжатие $t=1$ с, размер файла =4,12Кб.

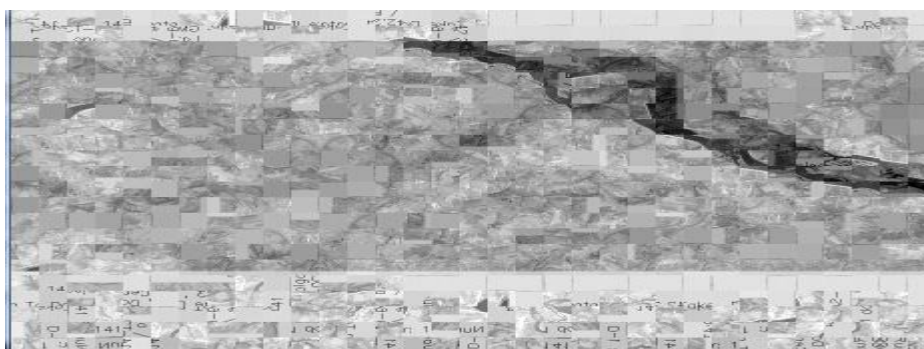


Рис. 2.35. Обработанное изображение.

Ниже, на рисунке 2.36 представлен график зависимости времени сжатия от размера изображения.

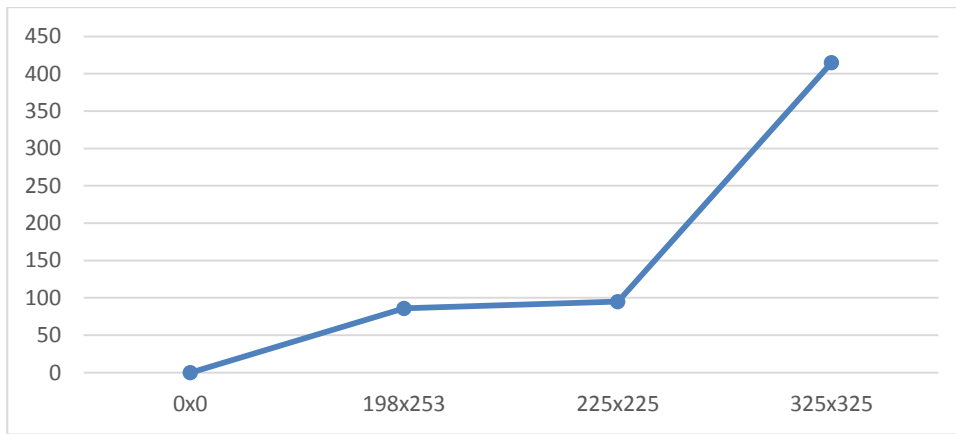


Рис. 2.36. График зависимости времени сжатия от размера изображения.

На рисунке 2.37 представлен график зависимости времени сжатия от параметра «смещение домена»

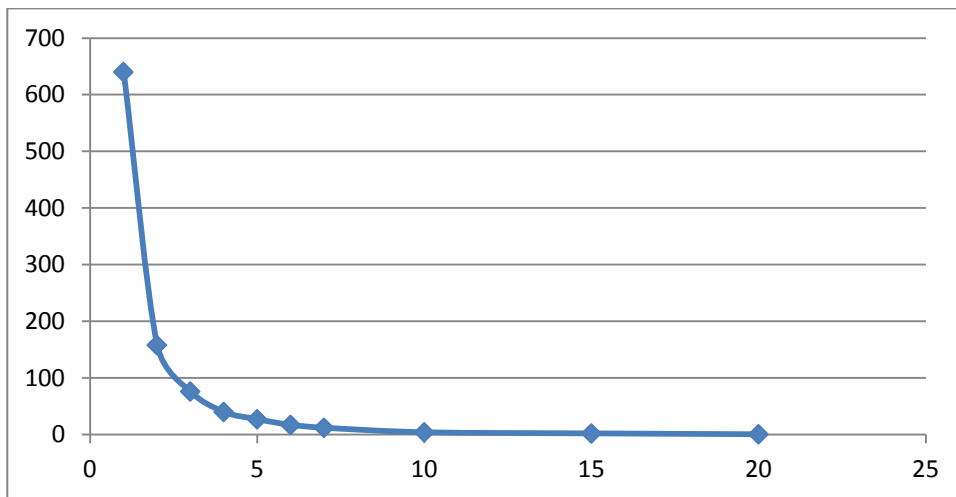


Рис. 2.37. График зависимости времени сжатия от параметра «смещение домена».

На рисунке 2.38 представлен график зависимости времени сжатия от параметра «размер региона».

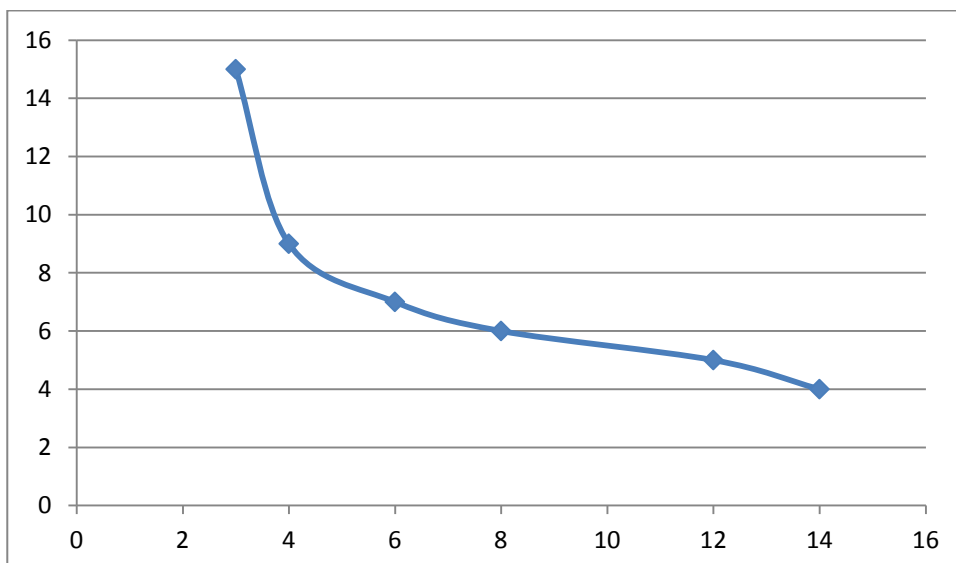


Рис. 2.38. График зависимости времени сжатия от параметра «размер региона».

На рисунке 2.39 представлен график зависимости размера изображения (в КБайт) от параметра «размер региона»

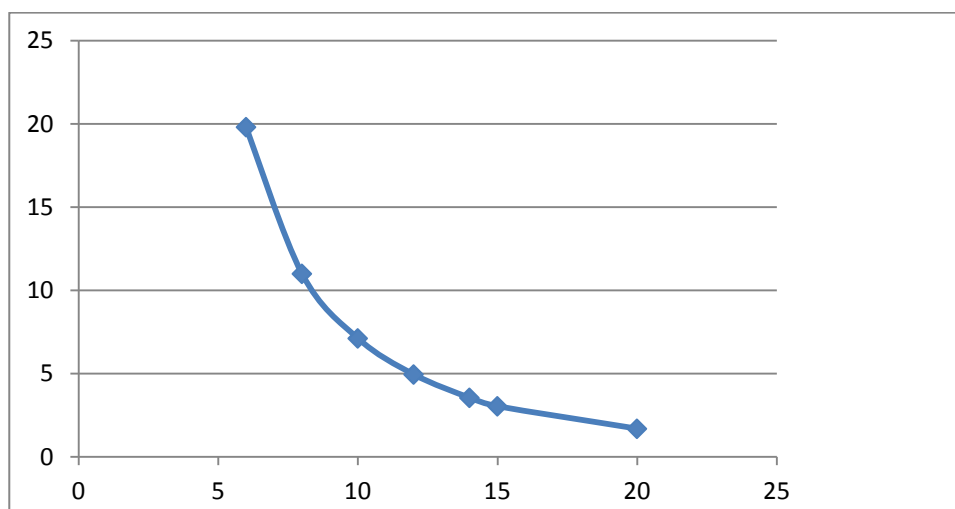


Рис. 2.39. График зависимости размера изображения (в КБайт) от параметра «размер региона»

Ниже, в таблице 2.6 представлена зависимость размера от коэффициента сжатия.

Таблица 2.6. Зависимость размера изображения от коэффициента сжатия.

Коэф-т сжатия	143	122	105	88	61	40
Размер изображения, Кб	3,04	3,55	4,12	4,95	7,12	11

Как видно из таблицы, при максимально достигнутом коэффициенте сжатия равному 143, размер изображения уменьшился с 435Кб до 3,04Кб.

Из представленных выше графиков можно сделать вывод, что:

1. Чем больше изображение, тем больше время сжатия;
2. Чем больше параметр «смещение домена», тем меньше время сжатия;
3. Чем меньше параметр «размер региона», тем больше время сжатия;
4. Чем меньше параметр «размер региона», тем больше размер изображения.

Рассмотрим таблицу 2.4, в которой сводятся воедино параметры различных алгоритмов сжатия изображений [13].

Таблица 2.7. Алгоритмы сжатия

Алгоритм	К-ты сжатия	На что ориентирован	Потери
RLE	32, 2, 0.5	3,4-х битные	Нет
LZW	1000, 4, 5/7	1-8 битные	Нет
Хаффмана	8, 1.5, 1	8 битные	Нет
СCITT-3	213(3), 5, 0.25	1-битные	Нет
JBIG	2-30 раз	1-битные	Нет
Lossless JPEG	2 раза	24-битные, серые	Нет
JPEG	2-20 раз	24-битные, серые	Да
Рекурсивное сжатие	2-200 раз	24-битные, серые	Да
Фрактальный	2-2000 раз	24-битные, серые	Да

Использование сжатия с потерями предоставляет возможность за счет потерь регулировать качество изображений. Коэффициенты сжатия у фрактальных алгоритмов варьируются в пределах 2-2000 раз. Причем большие коэффициенты достигаются на реальных изображениях, что нетипично для предшествующих алгоритмов. Ниже представлен график зависимости размера изображения от коэффициента сжатия

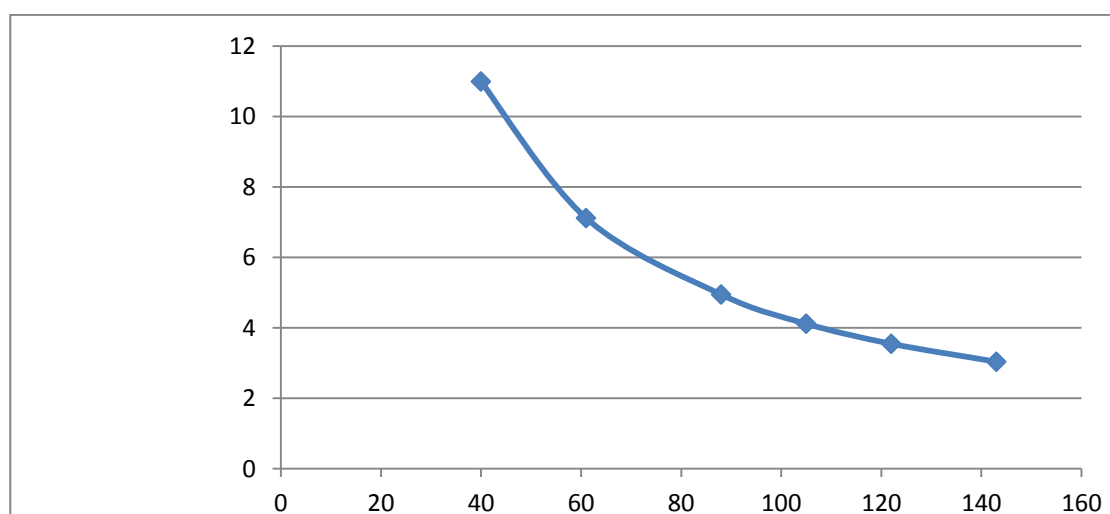


Рис. 2.40. Зависимость размера изображения по оси ординат, от коэффициента сжатия по оси абсцисс

Ниже, для наглядности, приведены 2 рисунка, первый – исходное изображение; второй – изображение с максимально достигнутом коэффициентом сжатия равным 143.

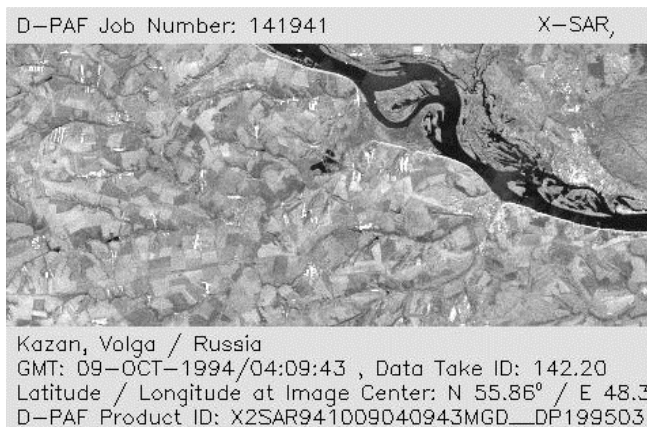


Рис. 2.41. Исходное изображение

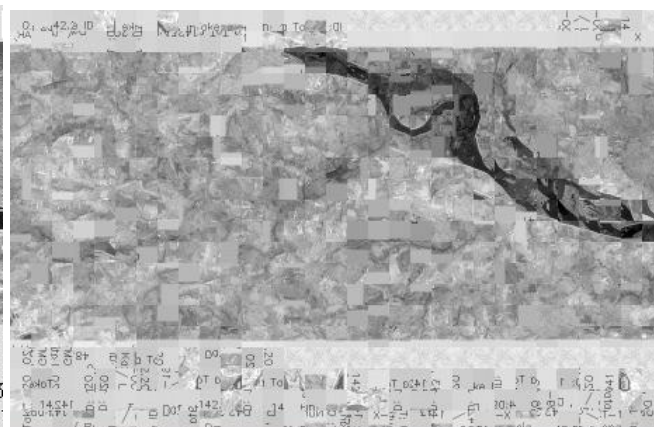


Рис. 2.42. Изображение с коэффициентом сжатия 143

Как видно из рисунков, при сжатии текстовой информации она становится нечитабельная, хотя как всё изображение в целом остаётся узнаваемым.

Недостатком этого алгоритма является потребность в больших вычислительных мощностях при архивации. Фактически это первый существенно несимметричный алгоритм. Причем, если у всех предшествующих алгоритмов коэффициент симметричности (отношение времени архивации ко времени разархивации) не превышает 3, то у фрактального алгоритма он колеблется от 1000 до 10000.

2.5. Вейвлет преобразования сигналов и изображений [15]

В настоящее время вейвлет-анализ является одним из наиболее мощных и при этом гибких средств исследования данных: помимо возможностей сжатия и фильтрации данных, анализ в базисе вейвлет-функций позволяет решать задачи идентификации, моделирования, аппроксимации стационарных и нестационарных процессов, исследовать вопросы наличия разрывов в производных, осуществлять поиск точек склеивания данных, удалять в данных тренд, отыскивать признаки фрактальности информации. Стоит отметить, что в основе подобных возможностей, обеспечивающих вейвлет-анализу весьма перспективное будущее, лежит природа его многомасштабности. Иначе говоря, гармонический анализ не способен конкурировать с вейвлет-анализом.

Вейвлет-преобразование широко используется для анализа сигналов. Помимо этого, оно находит большое применение в области сжатия данных. В дискретном вейвлет-преобразовании наиболее значимая информация в сигнале содержится при высоких амплитудах, а менее полезная — при низких. Сжатие данных может быть получено за счет отбрасывания низких амплитуд. Вейвлет-преобразование позволяет получить высокое

соотношение сжатия в сочетании с хорошим качеством восстановленного сигнала. Вейвлет-преобразование было выбрано для стандартов сжатия изображений JPEG2000 и ICER. Однако, при малых сжатиях вейвлет-преобразование уступает по качеству в сравнении с оконным Фурье-преобразованием, которое лежит в основе стандарта JPEG.

Выбор конкретного вида и типа вейвлетов во многом зависит от анализируемых сигналов и задач анализа. Для получения оптимальных алгоритмов преобразования разработаны определенные критерии, но их еще нельзя считать окончательными, так как они являются внутренними по отношению к самим алгоритмам преобразования и, как правило, не учитывают внешних критериев, связанных с сигналами и целями их преобразований. Отсюда следует, что при практическом использовании вейвлетов необходимо уделять достаточное внимание проверке их работоспособности и эффективности для поставленных целей по сравнению с известными методами обработки и анализа.

Изобретение вейвлетов было напрямую связано с необходимостью более глубокого анализа сигналов, чем анализ сигнала с помощью преобразования Фурье. Вейвлеты используют в тех случаях, когда результат анализа некоторого сигнала должен содержать не только простое перечисление его характерных частот (масштабов), но и сведения об определенных локальных координатах, при которых эти частоты проявляют себя. Анализ и обработка нестационарных (во времени) или неоднородных (в пространстве) сигналов разных типов представляют собой основное поле применения вейвлет-анализа.

Вейвлеты – новые системы базисных функций, используемых для представления, фильтрации, сжатия, хранения и т.д. любого из «сигналов»

$$f : R^n \rightarrow C. \quad (2.10)$$

Если $n = 1$, переменная t представляет собой время, и мы работаем с временными сигналами. $f : R \rightarrow C$. Случай $n=2$ относится к обработке изображений. То есть, вейвлет представляет собой функцию от времени, которая используется для анализа одномерного или двумерного сигнала.

В своих работах Малл определяет вейвлет как функцию ϕ с нулевым средним значением:

$$\int_{-\infty}^{\infty} \phi(t) dt = 0. \quad (2.11)$$

Вейвлетом называется волновое колебание с начальным значением амплитуды, равным нулю, затем увеличивающимся значением и затем снова уменьшающейся до нулевого значения. Это выражается в небольшом колебании исследуемого сигнала. Такое поведение этих функций позволяет записать и исследовать сейсмические волны или колебания сердца.

В общем случае, вейвлеты представляют собой функции, имеющие специфические свойства, позволяющие эффективно обрабатывать сигналы. Вейвлеты могут комбинироваться (с применением операций сдвига, умножения и суммирования), с выборками изучаемого сигнала для получения соответствующей информации.

Одна из основополагающих идей использования вейвлетов для представления сигналов заключается в разбивке приближения к сигналу на две составляющие: грубую (аппроксимирующую) и утонченную (детализирующую), с последующим их уточнением итерационным методом. Каждый шаг такого уточнения соответствует определенному уровню декомпозиции и восстановления сигнала. Это возможно как во временной, так и в частотной областях представления сигнала вейвлетами. Вейвлеты применяются во многих областях науки; главные области применения – обработка и анализ сигналов и сжатие изображений.

Для обработки сигнала с помощью вейвлетов сначала выбирается анализирующий (материнский) вейвлет. Обозначим его $x \mapsto \psi(x)$. Как правило, они определены на компактном носителе (на промежутке $[0, L]$). Под носителем функции понимается ее область определения.

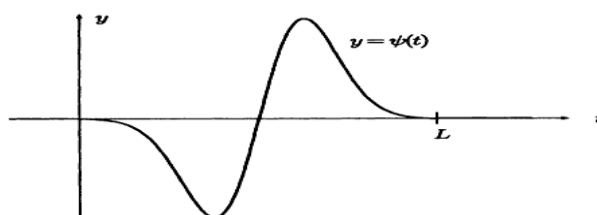


Рис. 2.43. Материнский вейвлет

Растянутые и сдвинутые копии вейвлета ψ называются вейвлетными функциями. Для них принято следующее обозначение:

$$\varphi_{a,b} : \mathbb{R} \rightarrow \mathbb{C}, t \mapsto \frac{1}{|a|^{1/2}} \psi\left(\frac{t-b}{a}\right). \quad (2.12)$$

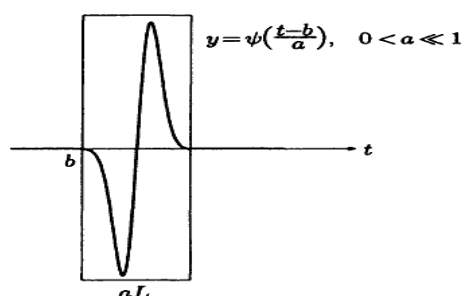


Рис. 2.44. Общий вид совокупности волновых функций (вейвлет-пакетов)

Параметр a называется масштабирующим параметром, а параметр b – параметром сдвига.

Результат выполнения вейвлет-преобразования (массив данных, получающийся в результате) полностью зависит от материнского вейвлета.

На практике используются следующие типы вейвлетов:

- Вейвлет Хаара.
- Мексиканская шляпа.
- Модулированная гауссова кривая.
- Производная гауссовой кривой.
- Вейвлеты Добеши, Грассмана, Мейера.
- Вейвлеты Бэттла-Лемарье.
- Симлеты.
- Койфлеты.

Математический аппарат вейвлета Хаара

Вейвлет Хаара представляет собой следующую функцию:

$$\psi(x) = \begin{cases} 1, (0 \leq x < \frac{1}{2}) \\ -1, (\frac{1}{2} \leq x < 1) \end{cases} \quad (2.13)$$

Используя в качестве материнской функции вейвлет Хаара, можно определить вейвлетные функции.

$$\psi_{r,k} = 2^{-r/2} \psi_{haar} \left(\frac{t - k \cdot 2^r}{2^r} \right), r, k \in \mathbb{Z} \quad (2.14)$$

В качестве носителя данная функция имеет интервал длины 2^r :

$$I_{r,k} = [k \cdot 2^r, (k+1) \cdot 2^r] \quad (2.15)$$

Большим значениям r соответствуют большие интервалы L , следовательно, вейвлетные функции имитируют большие волны. Одним из примеров, иллюстрирующих применение вейвлетов, является разложение сигнала с помощью выбранного преобразования [5].

Масштабирующая функция имеет вид:

$$\phi(x) = \begin{cases} 1, 0 \leq x < 1 \\ 0, x \notin [0,1) \end{cases} \quad (2.16)$$

Ее интеграл: $\int_{-\infty}^{\infty} \phi(x) dx = 1$.

Масштабирующая функция определяет аппроксимацию сигнала (позволяет исследовать числовые характеристики и качественные свойства объекта, сводя задачу к изучению более простых или более удобных объектов).

Построение масштабирующей функции и функции вейвлета Хаара представлено на рисунок 2.45.

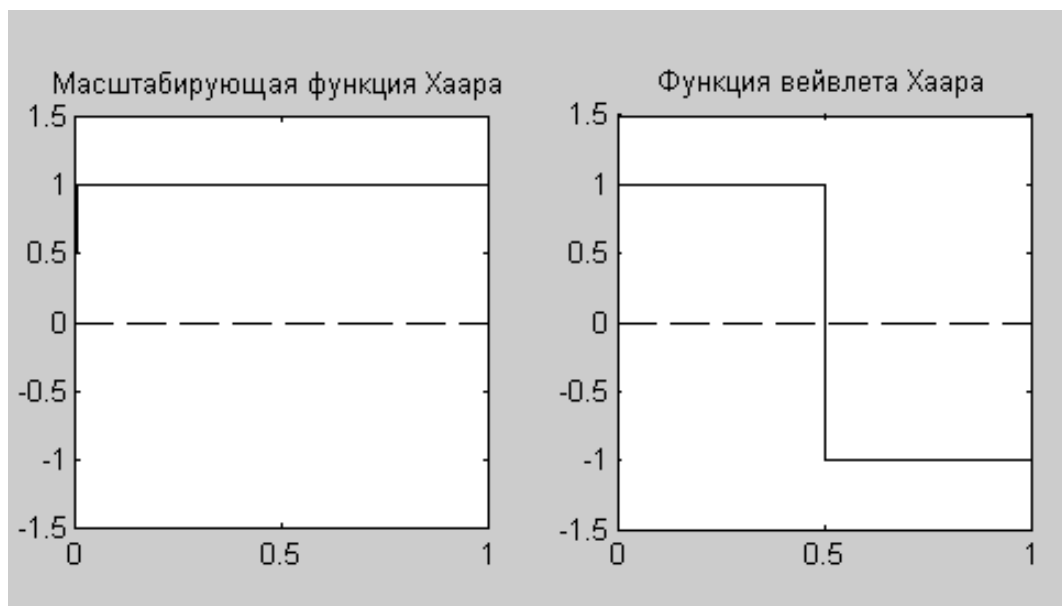


Рис. 2.45. Масштабирующая функция и функция вейвлета Хаара

Преобразование Хаара использует комбинацию этих функций. Чтобы понять, как выполняется преобразование, выполним сначала преобразование одномерного массива.

Пример. Выполним преобразование Хаара следующего массива:

$$A = (6, 7, 8, 9, 10, 6, 7, 8).$$

Вычислим сначала средние и полуразности элементов:

1 шаг:

$$\text{Средние: } (6+7)/2 = 13/2 = 6.5; (8+9)/2 = 8.5; (10+6)/2 = 8; (7+8)/2 = 7.5.$$

$$\text{Полуразности: } (6-7)/2 = -0.5; (8-9)/2 = -0.5; (10-6)/2 = 2; (7-8)/2 = -0.5$$

Таким образом, получим массив из восьми элементов:

$$A_1 = (6.5, 8.5, 8, 7.5, -0.5, -0.5, 2, 0.5)$$

Обратим внимание на последние четыре элемента – их значения гораздо меньше, чем до выполнения преобразования. Поэтому их оставим без изменений. Средние значения называются аппроксимирующими коэффициентами, в которых накапливается максимальная информация, а значения полуразностей позволяют осуществлять полное восстановление

исходного сигнала. На следующем шаге мы будем работать с первыми четырьмя элементами, а вторую половину массива оставим без изменений.

2 шаг:

Средние: $(6.5+8.5)/2 = 7.5$; $(8+7.5)/2 = 7.75$.

Полуразности: $(6.5-8.5)/2 = -1$; $(8-7.5)/2 = 0.25$.

Новый массив: $A_2 = (7.5, 7.75, -1, 0.25, -0.5, -0.5, 2, -0.5)$

3 шаг:

Средние: $(7.5+7.75)/2 = 7.625$; $(-1+0.25) = 0.625$.

Полуразности: $(7.5-7.75)/2 = -0.125$; $(-1-0.25)/2 = -0.625$.

Новый массив $A_3 = (7.625, 0.625, -0.125, -0.625, -0.5, -0.5, 2, -0.5)$

Из за постоянного вычисления полуразностей происходит постоянное уменьшение значений исходных элементов (пикселей изображения), поэтому после выполнения преобразования, можно к получившемуся массиву дополнительно применить один из алгоритмов сжатия (например, Хаффман или арифметическое кодирование). Если необходимо использовать сжатие с потерями, то здесь используется операция квантования или удаления наименьших значений полуразностей: например, может быть записано: $A_3 = (7.625, 0.6250, 0, 0, 0, 2, 0)$. Тогда исходные данные восстановятся с искажениями.



Рис. 2.46. Результат применения вейвлета Хаара

Математический аппарат вейвлета Добеши

В настоящее время вейвлет Добеши 4-ого порядка является наиболее используемым. Функция данного вейвлета и его масштабирующая функция представлены на рисунке 2.47.

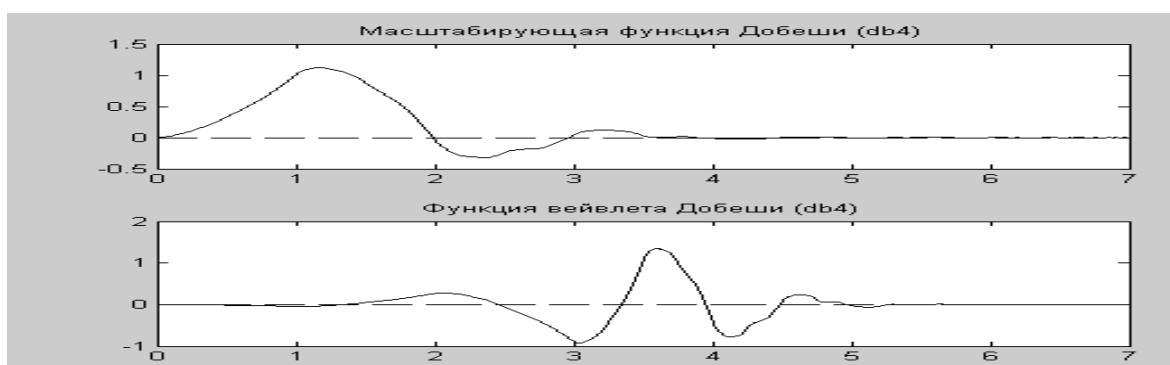


Рис. 2.47. Масштабирующая функция и функция вейвлета Добеши 4-го порядка

Масштабирующая функция и функция вейвлета Добеши задаются рекурсивно. Кроме того, функции удовлетворяют следующему требованию: масштабирующая функция имеет компактный носитель – отрезок $[0, 3]$, а также должна быть равна нулю вне этого отрезка. Эту функцию, как и вейвлет, можно задать рекурсивно, определив начальные условия.

Начальные значения:

$$\varphi(0) = 0, \varphi(1) = \frac{1 + \sqrt{3}}{2}, \varphi(2) = \frac{1 - \sqrt{3}}{2}, \varphi(3) = 0 \quad (2.17)$$

И рекурсивное соотношение:

$$\begin{aligned} \varphi(r) = & \frac{1 + \sqrt{3}}{4} \varphi(2r) + \frac{3 + \sqrt{3}}{4} \varphi(2r - 1) + \frac{3 - \sqrt{3}}{4} \varphi(2r - 2) + \\ & \frac{1 - \sqrt{3}}{4} \varphi(2r - 3) = \end{aligned} \quad (2.18)$$

$$\begin{aligned} h_0 \varphi(2r) + h_1 \varphi(2r - 1) + h_2 \varphi(2r - 2) + h_3 \varphi(2r - 3) = \\ (h_0, h_1, h_2, h_3) \cdot (\varphi(2r), \varphi(2r - 1), \varphi(2r - 2), \varphi(2r - 3)). \end{aligned}$$

Сумма начальных значений равна единице.

$$\varphi(0) + \varphi(1) + \varphi(2) + \varphi(3) = 0 + \frac{1 + \sqrt{3}}{2} + \frac{1 - \sqrt{3}}{2} + 0 = 1 \quad (2.19)$$

Далее функция вычисляется по шагам. Для этого на шаге 1 применяются условие компактного носителя, начальные значения и рекурсивные соотношения. На последующих шагах также применяется рекурсия. Для построения вейвлета также используется рекурсивное соотношение.

$$\psi(r) = -\frac{1 + \sqrt{3}}{4} \varphi(2r - 1) + \frac{3 + \sqrt{3}}{4} \varphi(2r) - \frac{3 - \sqrt{3}}{4} \varphi(2r + 1) + \frac{1 - \sqrt{3}}{4} \varphi(2r + 2) \quad (2.20)$$

Использование рекурсивных соотношений при построении вейвлетов и масштабирующих функций способствует тому, что построенные кривые являются

самоподобными, то есть обладают фрактальными свойствами. Это делает возможным применение вейвлетов к анализу фрактальных последовательностей.

Фильтры $\{h_n\}_{n \in \mathbb{Z}}$ и $\{g_n\}_{n \in \mathbb{Z}}$ вейвлетов $\phi(x)$ и $\psi(x)$ есть коэффициенты масштабирующих уравнений $\phi(x) = \sqrt{2} \sum_n h_n \phi(2x - n)$ и $\psi(x) = \sqrt{2} \sum_n g_n \psi(2x - n)$, где $g_n = (-1)^n \overline{h_{N-n}}$ (черта обозначает комплексное сопряжение). Фильтр h называется низкочастотным, а фильтр g – высокочастотным. Для восстановления используются фильтры $\{h_n\}$ и $\{g_n\}$ вейвлетов ϕ и ψ , а для восстановления с помощью фильтров – сопряженные (транспонированные) фильтры.

Математический аппарат непрерывного вейвлет-преобразования

Запись непрерывного вейвлет-преобразования (НВП) выглядит следующим образом:

$$Wf(a, b) = \int_{-\infty}^{\infty} f(t) \frac{1}{\sqrt{s}} \psi^* \left(\frac{t-b}{a} \right) dt \quad (2.21)$$

Областью определения вейвлет-преобразования Wf является множество:

$$\mathbb{R}^2_- = \{(a, b) \mid a \in \mathbb{R}^*, b \in \mathbb{R}\} \quad (2.22)$$

Параметр a определен в комплексной плоскости, параметр b – в области действительных чисел. В теории вейвлетов ось a масштабируется вертикально, а ось b – горизонтально. Часто область определения Wf ограничивается положительными значениями параметра a .

Главной задачей является нахождение вейвлет-коэффициентов. Использование свойств вейвлет-преобразования позволяет осуществить анализ сигналов. Например, слишком большие амплитуды значений вейвлет-коэффициентов указывают на положение перепадов, которые имеют резкие изменения интенсивности изображения. Различные масштабы описывают контуры структуры изображения меняющихся размеров. Такое выделение перепадов эффективно для распознавания образов при компьютерной визуализации [4].

Представляется необходимым продемонстрировать действие НВП на примере вейвлета Хаара.

Вейвлет Хаара может быть записан как:

$$\psi \left(\frac{t-b}{a} \right) = \begin{cases} 1, & \left(b \leq t \leq b + \frac{a}{2} \right) \\ -1, & \left(b + \frac{a}{2} \leq t < b + a \right) \\ 0 & \end{cases} \quad (2.23)$$

Вейвлет-преобразование определяется как:

$$Wf(a,b) = \frac{1}{\sqrt{a}} \left(\int_b^{b+a/2} f(t)dt - \int_{b+a/2}^{b+a} f(t)dt \right) = \frac{\sqrt{a}}{2} \left(\frac{2}{a} \int_b^{b+a/2} f(t)dt - \frac{2}{a} \int_{b+a/2}^{b+a} f(t)dt \right). \tag{2.24}$$

Это означает, что за исключением нормирующего множителя значение $Wf(a,b)$ представляет собой разность между двумя средними значениями функции f , эти средние значения берутся по двум соседним интервалам длины $a/2$ в окрестности точки b .

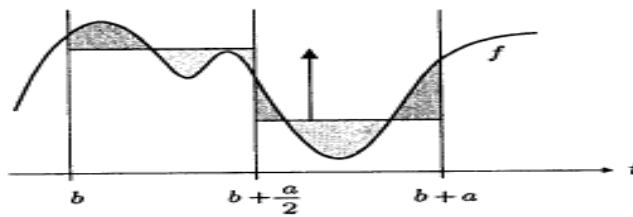


Рис. 2.48. Результат НВП в случае вейвлета Хаара

Величину $Wf(a,b)$ можно рассматривать с другой точки зрения:

$$Wf(a,b) = \frac{1}{\sqrt{a}} \int_a^{b+a/2} \left(f(t) - f\left(t + \frac{a}{2}\right) \right) dt = -\frac{1}{\sqrt{a}} \int_b^{b+a/2} \left(\int_t^{t+a/2} f'(r) dr \right) dt = \dots = -\frac{1}{\sqrt{a}} \int_{a/2}^{a/2} \left(\frac{a}{2} - |r| \right) f'\left(b + \frac{a}{2} + r\right) dr \tag{2.25}$$

В этой форме величина $Wf(a,b)$ проявляется как взвешенное среднее производной f' на интервале $[b, b+a]$.

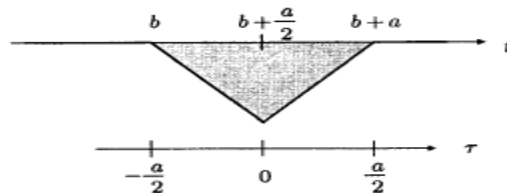


Рис. 2.49. Другая интерпретация величины Wf

Подставляя вместо $f(t)$ конкретную функцию, можно вычислить ее НВП.

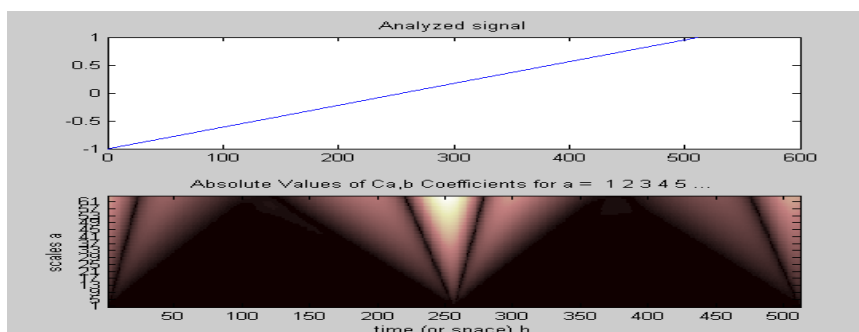


Рис. 2.50. Результат НПВ сигнала

По полученной спектрограмме можно судить о характере сигнала, например, наблюдать максимумы «всплесков энергии» в различные промежутки времени (показатель b) и при различных масштабах (показатель a).

Математический аппарат дискретного вейвлет-преобразования

Для обработки и просмотра сигналов и изображений при различных разрешениях также используется дискретное вейвлет-преобразование.

В общем виде прямое дискретное преобразование изображения $f(x,y)$ размерности $M \times N$ можно выразить в общем виде следующим образом:

$$T(u, v, \dots) = \sum_{x, y} f(x, y) g_{u, v, \dots}(x, y) \quad (2.26)$$

Здесь x и y – пространственные переменные (координаты пикселей изображения), а u, v – переменные в частотной области, характеризующие закономерность распределения пикселей изображения. Зная $T(u, v, \dots)$, можно восстановить функцию $f(x, y)$ с помощью обратного преобразования:

$$f(x, y) = \sum_{u, v, \dots} T(u, v, \dots) h_{u, v, \dots}(x, y). \quad (2.27)$$

Члены последовательностей g и h называются прямыми и обратными ядрами преобразования, определяющими природу, сложность и эффективность пары преобразований. Дискретное вейвлет-преобразование, в отличие от преобразования Фурье, обозначает целый класс преобразований, которые различаются не только своими ядрами (коэффициентами), но и свойствами этих коэффициентов (*с каким классом вейвлетов мы будем работать*) и способом применения этих коэффициентов (*сколько требуется вычислить различных решений*).

Можно охарактеризовать каждое ДВП с помощью ядра преобразования или, основываясь на множестве параметров, которые однозначно определяют пару ядер. Все преобразования являются родственными (их функции разложения представляют собой «маленькие волны» или «вейвлеты»), которые имеют переменную частоту колебаний и ограниченную длительность.

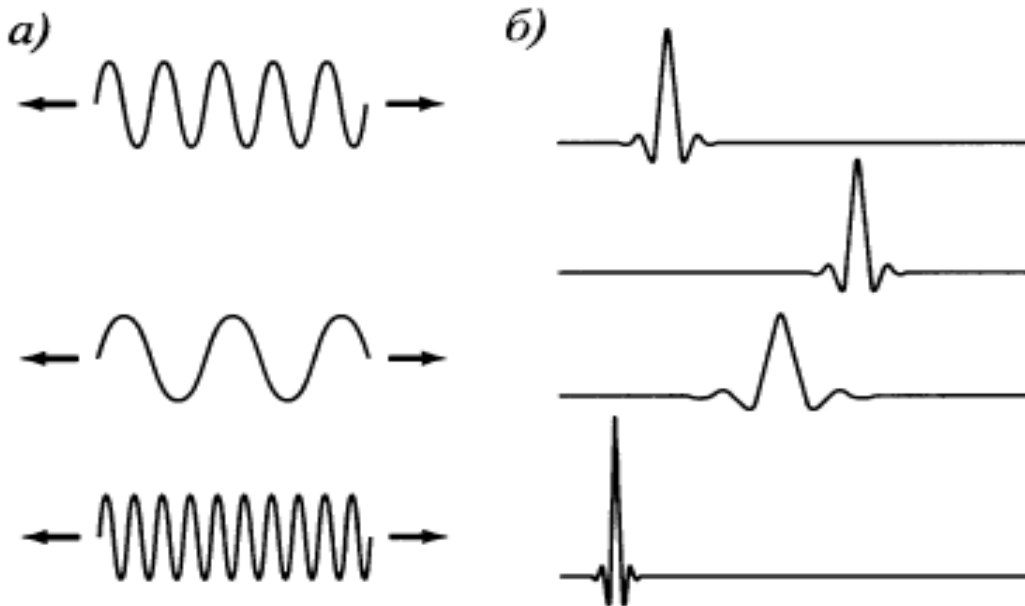


Рис. 2.51. а) Множество функций разложения Фурье – синусоиды с переменной частотой и бесконечной длительностью б) Функции разложения ДВП – «маленькие волны» с переменной частотой и конечной длительностью

Свойства «ядер», являющиеся общими для всех типов вейвлетов.

Свойство 1. *Разделимость, масштабируемость и переносимость.* Ядра можно представить в виде трех разделимых двумерных вейвлетов:

$$\begin{aligned}\psi^H(x, y) &= \psi(x)\varphi(y), \\ \psi^V(x, y) &= \varphi(x)\psi(y), \\ \psi^D(x, y) &= \psi(x)\psi(y).\end{aligned}\tag{2.28}$$

Здесь $\psi^H(x, y), \psi^V(x, y), \psi^D(x, y)$ называются горизонтальными, вертикальными и диагональными вейвлетами, а двумерная функция:

$$\varphi(x, y) = \varphi(x)\varphi(y)\tag{2.29}$$

Называется масштабирующей функцией. Каждая из этих двумерных функций является произведением двух одномерных масштабирующих функций [5].

$$\begin{aligned}\varphi_{j,k}(x) &= 2^{j/2}\varphi(2^j x - k), \\ \psi_{j,k}(x) &= 2^{j/2}\psi(2^j x - k).\end{aligned}\tag{2.30}$$

Параметры i и j являются целыми числами. Число k определяет положение одномерных функций на оси x , масштаб j – ширину по оси x , множитель $2^{j/2}$ отвечает за высоту и амплитуду.

Свойство 2. *Кратномасштабная совместимость.* Свойство оговаривает требования к выбранной масштабирующей функции.

а) Каждая функция $\varphi_{i,j}(x)$ ортогональная любому своему целочисленному сдвигу.

б) Множество функций, которые можно представить в виде рядов разложения по $\varphi_{j,k}(x)$ при малых масштабах (j мало), принадлежит множеству функций, представимых в более высоких значениях.

в) Каждую масштабирующую функцию можно представить с произвольной точностью при $j \rightarrow \infty$.

Свойство 3. Ортогональность. Функции разложения $\{\varphi_{i,j}(x)\}$ образуют ортогональный или биортогональный базис (отсюда и название еще одного типа вейвлетов – ортогональные и биортонональные) в пространстве одномерных измеримых, суммируемых в квадрате функций. Базис обладает тем свойством, что каждая представимая в этом базисе функция имеет единственный набор коэффициентов. Для вещественных ортогональных ядер имеет место равенство $g = h$. В биортогональном случае

$$\langle h_r, g_s \rangle = \delta_{rs} = \begin{cases} 1, & r = s. \\ 0 & \end{cases} \quad (2.31)$$

Для практического применения важно следствие из приведенных свойств – функцию вейвлета и масштабирующую функцию можно представить в виде линейной комбинации своих же копий с удвоенным разрешением:

$$\varphi(x) = \sum_n h_\varphi(n) \sqrt{2} \varphi(2x - n), \quad \psi(x) = \sum_n h_\psi(n) \sqrt{2} \psi(2x - n) \quad (2.32)$$

В среде MATLAB дискретное вейвлет-преобразование реализуется с помощью процедур `dwt` и `fwt` (быстрое вейвлет-преобразование). Можно выполнять разложение как двумерных, так и одномерных сигналов.



Рис. 2.52. Результат дискретного вейвлет-преобразования

Алгоритм быстрого вейвлет-преобразования

Обобщая преобразование на двумерный и многомерный уровень, можно представить схему преобразования (быстрого вейвлет-преобразования).

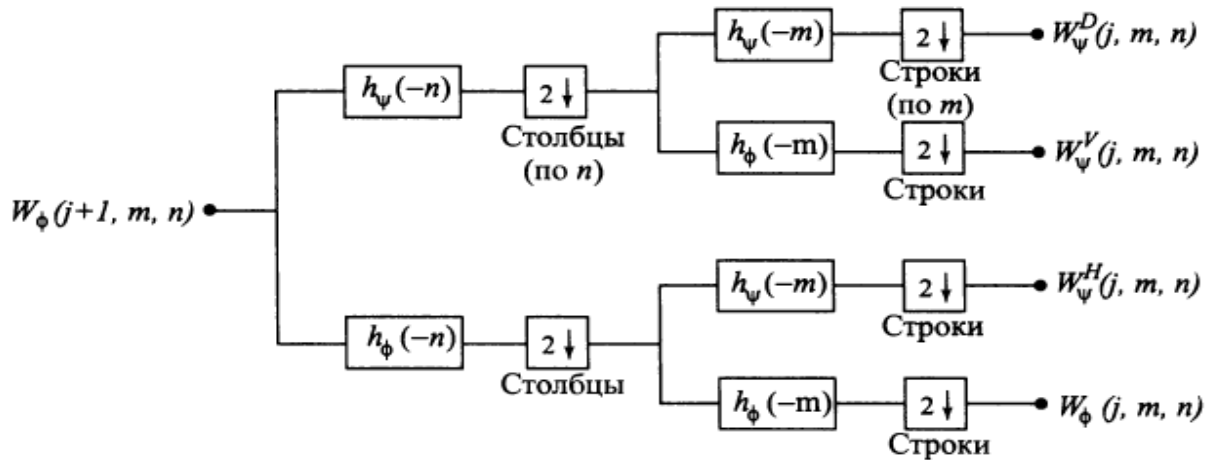


Рис. 2.53. Схема быстрого вейвлет-преобразования

Сначала обрабатываются столбцы, а затем строки. Используются при этом фильтры низких и высоких частот.

Каждый проход через блок фильтров на рис. 10 разлагает входные данные на четыре компоненты меньшего диапазона (или масштаба). Коэффициенты W_φ преобразования получаются двумя проходами низкочастотной фильтрации (т. е. с фильтром h_φ), и поэтому они называются коэффициентами приближения; коэффициенты $W_\varphi^i, i = H, V, D$ называются, соответственно, коэффициентами горизонтальных, вертикальных и диагональных деталей. Поскольку само изображение $f(x, y)$ представлено в наивысшем разрешении, то оно становится входом $W_\varphi(j+1, m, n)$ первой итерации процедуры. Кроме того, здесь имеется три переменные преобразованного пространства: масштаб j , горизонтальная трансляция p и вертикальная трансляция t . Эти переменные соответствуют обозначениям u, v, \dots в первых двух уравнениях (8), (9). Далее посмотрим примеры применения двумерного вейвлет-преобразования (dwt2, fwt) при обработке изображений.

При выполнении этапов преобразования используется операция свертки сигналов. Слово «свертка» обозначает совместное сворачивание двух величин или функций [2]. Дискретная свертка векторов f и g обозначается $f * g$ и вычисляется с помощью соотношения:

$$(f * g)_i = \sum_j f_j g_{i-j} \quad (2.33)$$

В результате выполнения операции получается один вектор, охватывающий диапазон входного сигнала. Поэтому вейвлет-преобразования называют *поддиапазонными*. Дискретная свертка двух числовых последовательностей $f(i)$ и $g(i)$ задается неравенством:

$$h(i) = f(i) * g(i) = \sum_j f(i)g(i - j)$$

Свертка сигналов применяется, например, для удаления шума или сглаживания сигналов.

Использование приложения Wavelet Toolbox Matlab для сжатия информации путем вейвлет-преобразований [12].

Пакет MATLAB Wavelet Toolbox представляет собой совокупность программ, позволяющих выполнять вейвлет-анализ и обработку сигналов средствами графического интерфейса пользователя. Пакет позволяет продемонстрировать практические приложения теории вейвлетов и вейвлет-преобразований. Остановимся на следующих пунктах:

1. Просмотр вейвлетов.
2. Одномерный дискретный вейвлет-анализ.
 - 2.1. Вейвлет-разложение сигнала.
 - 2.2. Интерпретация статистических характеристик сигнала.
 - 2.3. Гистограммы коэффициентов аппроксимации и детализации.
3. Сжатие одномерных сигналов.
4. Удаление шума из сигналов.

Главное меню пакета вызывается из MATLAB с помощью команды `wavemenu`.

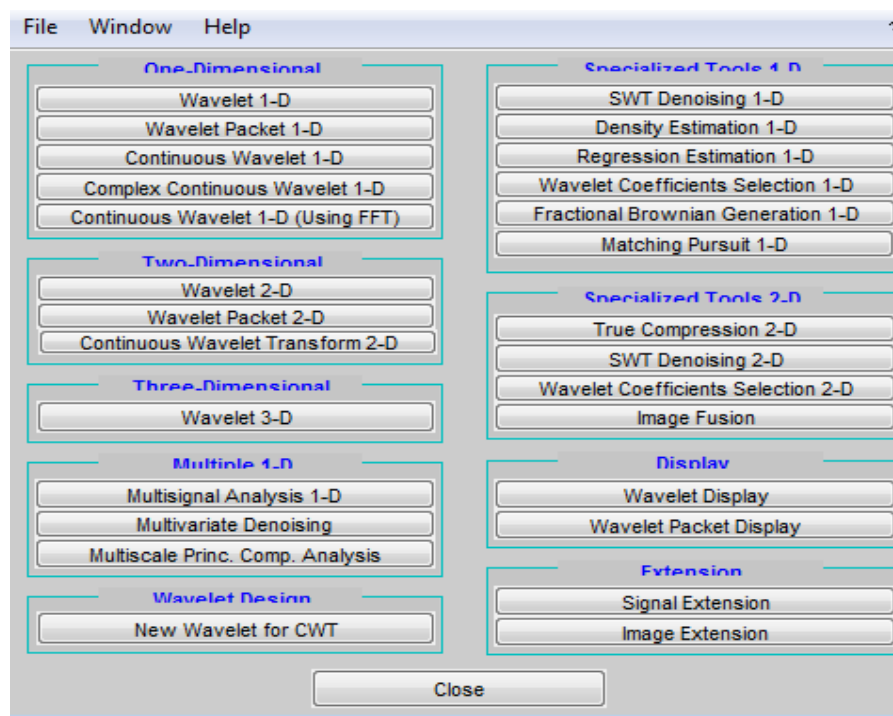


Рис. 2.54. Главное окно Wavelet Toolbox

Способы просмотра вейвлетов в Wavelet Toolbox

Способы просмотра вейвлетов – группа Display (кнопка Wavelet Display). Откроется следующее окно:

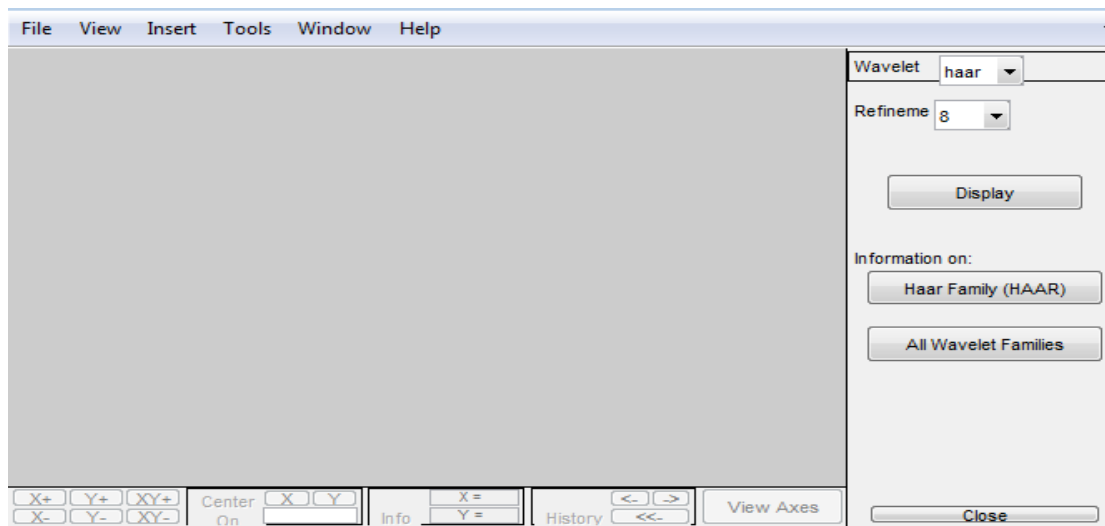


Рис. 2.55. Окно раздела Wavelet Display после запуска

Вверху расположено обычное меню, справа и внизу – меню, позволяющие управлять отображением нужной информации.

В правом меню можно из раскрывающегося списка тип вейвлета, задать уровень разрешения (параметр *Refinement*)

– определить шаг вычисления значений вейвлетов (этот параметр равен $1/2^N$, N по умолчанию равно 8).

Также можно посмотреть информацию по выбранному типу вейвлета.

Выбрав из списка тип вейвлета 'haar', следует нажать кнопку *Display* ($Refinement = 8$).

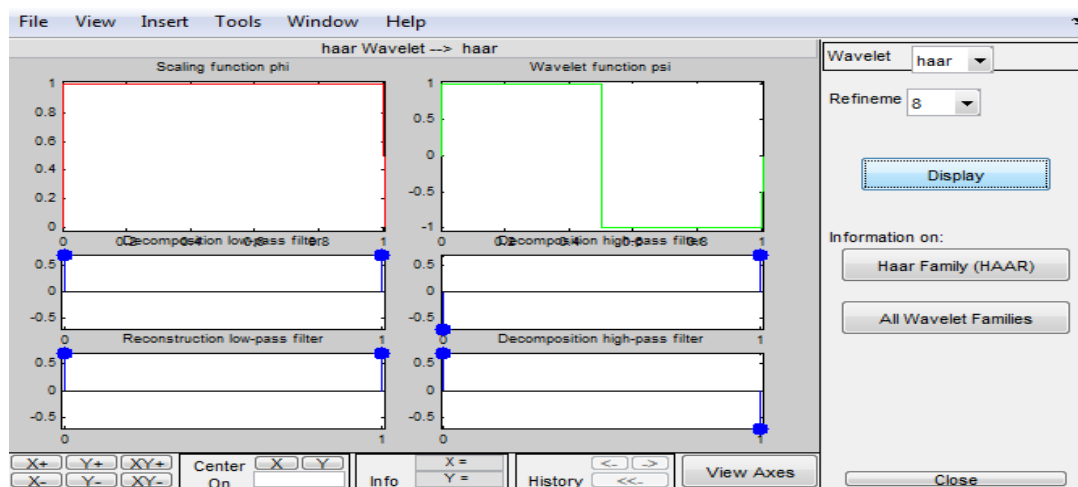


Рис. 2.56. Вейвлет и масштабирующая функция Хаара, фильтры восстановления и разложения

Вверху расположены графики масштабирующей функции и функции вейвлета. Ниже изображены фильтры восстановления и разложения вейвлета (низкочастотные и высокочастотные). Для восстановления используются фильтры (коэффициенты масштабирующих уравнений) h и g функций $\varphi(x)$ и $\psi(x)$, а для вейвлет-разложения – сопряженные (транспонированные) фильтры. Это хорошо видно на графиках фильтров. Для вейвлета Хаара фильтры простые, так как базовыми операциями являются нахождение средних и полуразностей. У других вейвлетов фильтры более сложные по своей структуре.

Например, если выбрать вейвлет Добеши 7-ого порядка, то получится следующее изображение:

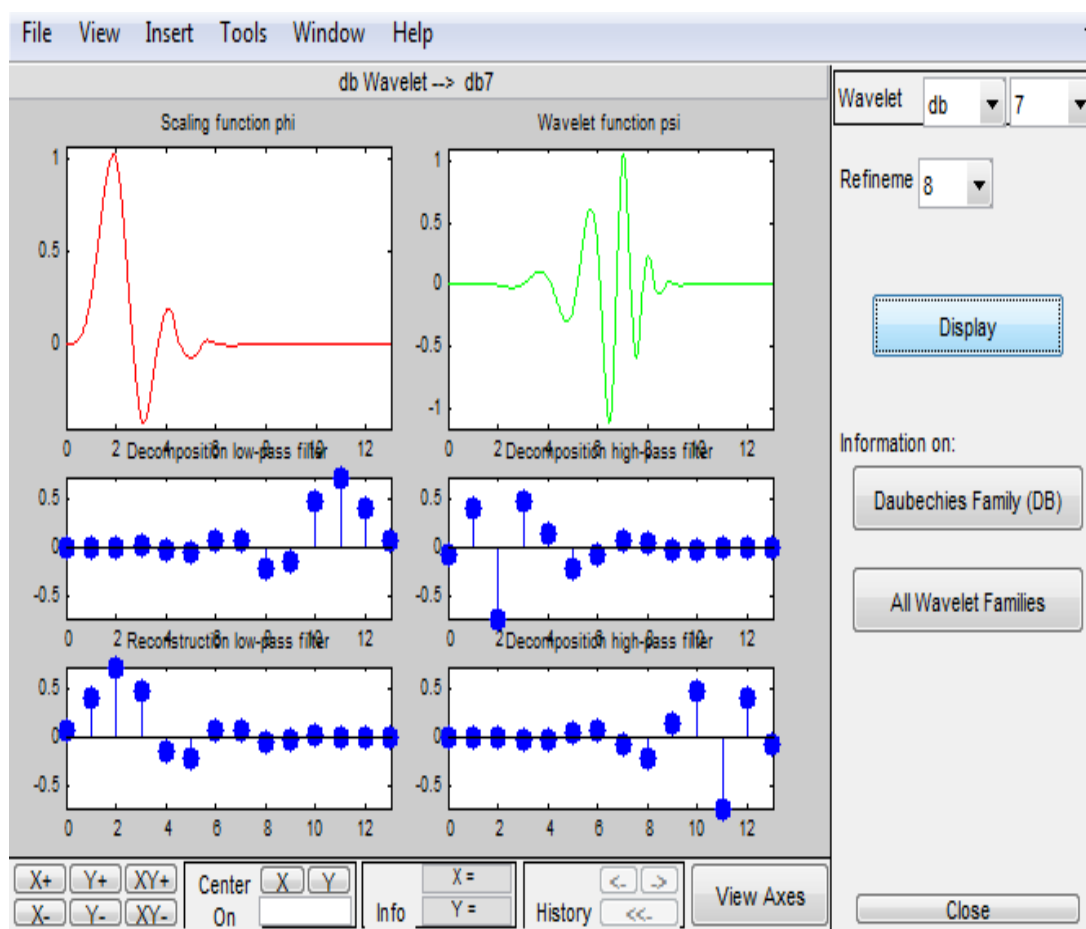


Рис. 2.57. Вейвлет 7-го порядка и масштабирующая функция Добеши, фильтры восстановления и разложения

Сжатие сигналов при помощи вейвлет-преобразований в Wavelet Toolbox

В главном окне необходимо нажать кнопку Wavelet 1-D, появится диалоговое окно, интерфейс которого позволяет осуществлять обработку одномерных сигналов.

Загрузка сигнала Frequency Breakdown (Load → Example Analysis → Basic Signals → Frequency Breakdown):

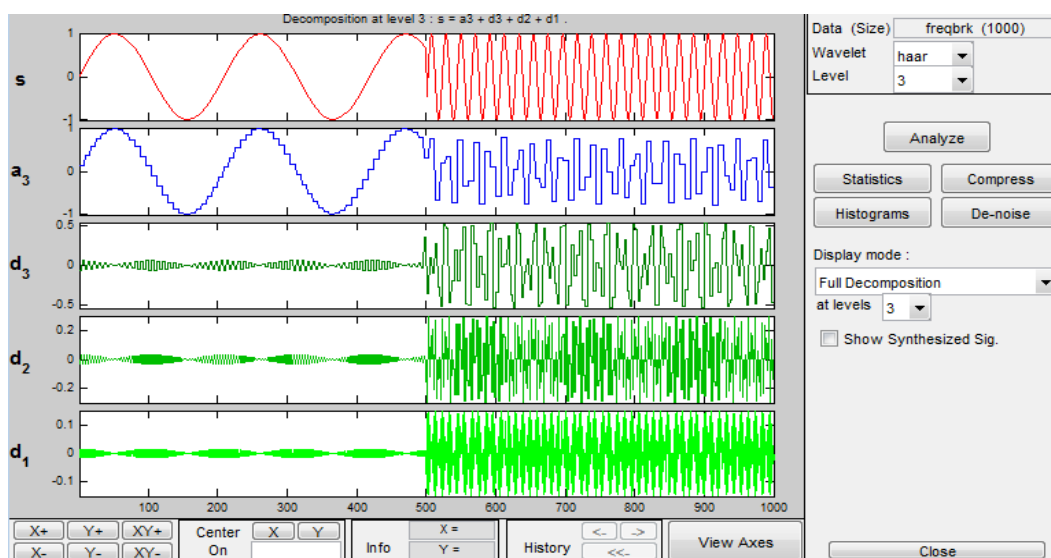
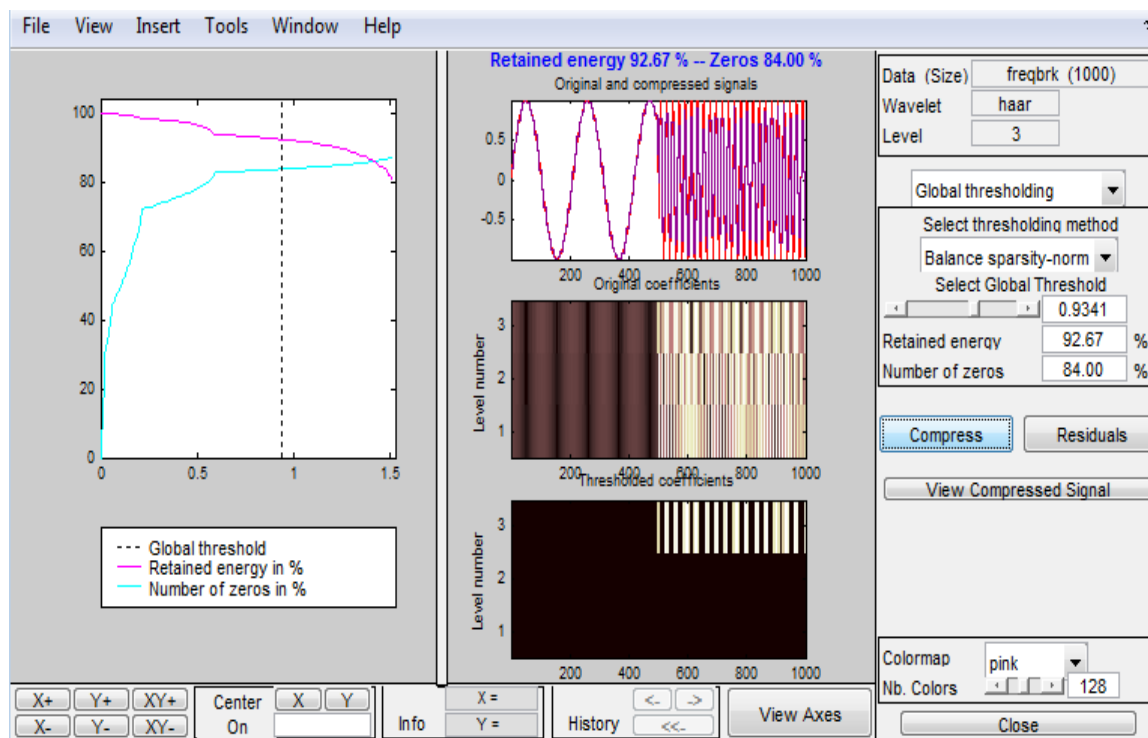


Рис. 2.58. Разложение сигнала

Кнопка **compress** позволяет осуществить сжатие одномерного сигнала.

Рис. 2.59. Интерфейс окна, вызываемого нажатием кнопки **Compress**

При сжатии одномерных сигналов используется метод глобальной пороговой обработки детализирующих коэффициентов. Управлять сжатием можно группой компонентов справа, устанавливая ползунком значение. Задав порог (процент сохраняемых коэффициентов), в результате получается сжатый сигнал.

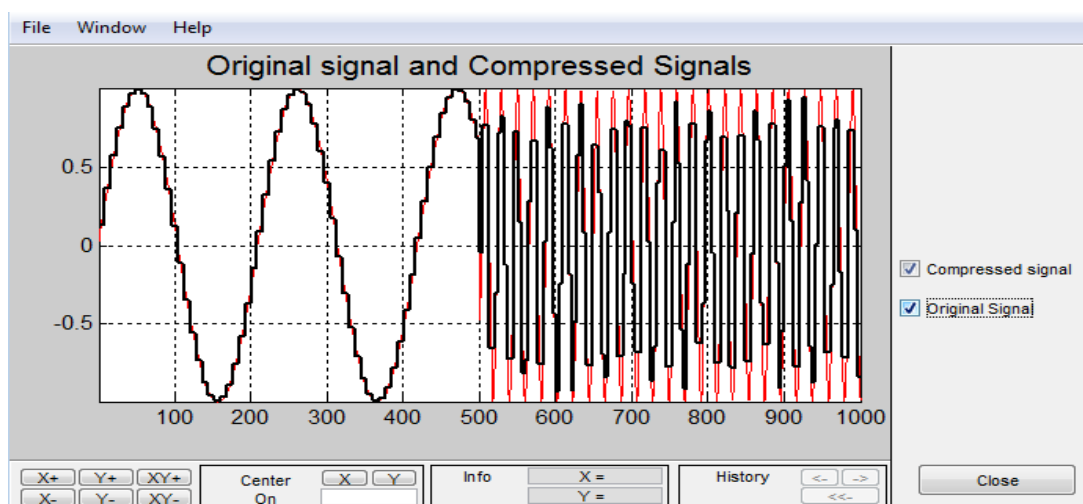


Рис. 2.60. Исходный и сжатый сигнал при помощи вейвлет-преобразования Хаара третьего уровня

Сжатие одномерного сигнала при помощи различных вейвлет-преобразований

На сегодняшний день существует большое количество математических алгоритмов вейвлет-преобразований, данный раздел содержит рассмотрение трех основных и сравнительный анализ их характеристик.

Вейвлет Хаара пятого уровня. Разложение исследуемого сигнала при помощи вейвлета Хаара пятого уровня представлено на рисунке 2.61

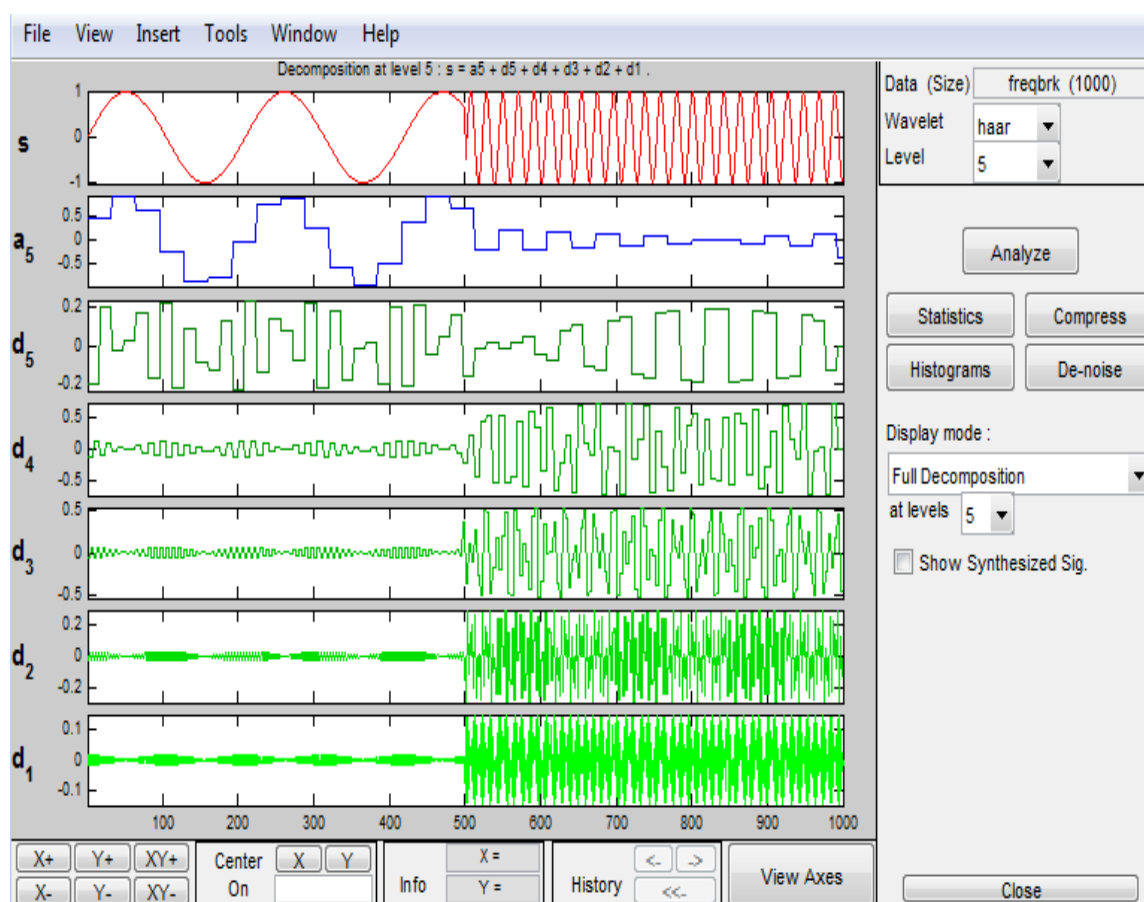


Рис. 2.61. Разложение сигнала при помощи вейвлета Хаара пятого уровня

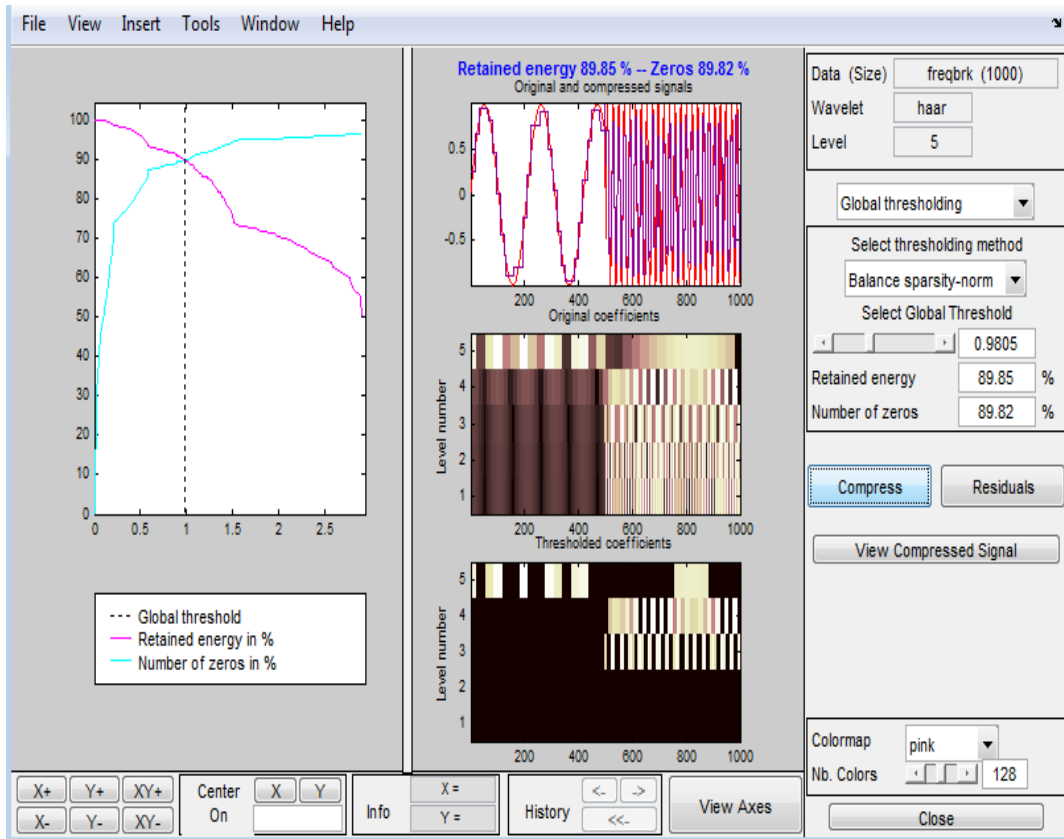


Рис. 2.62. Сжатие сигнала при помощи вейвлета Хаара пятого уровня

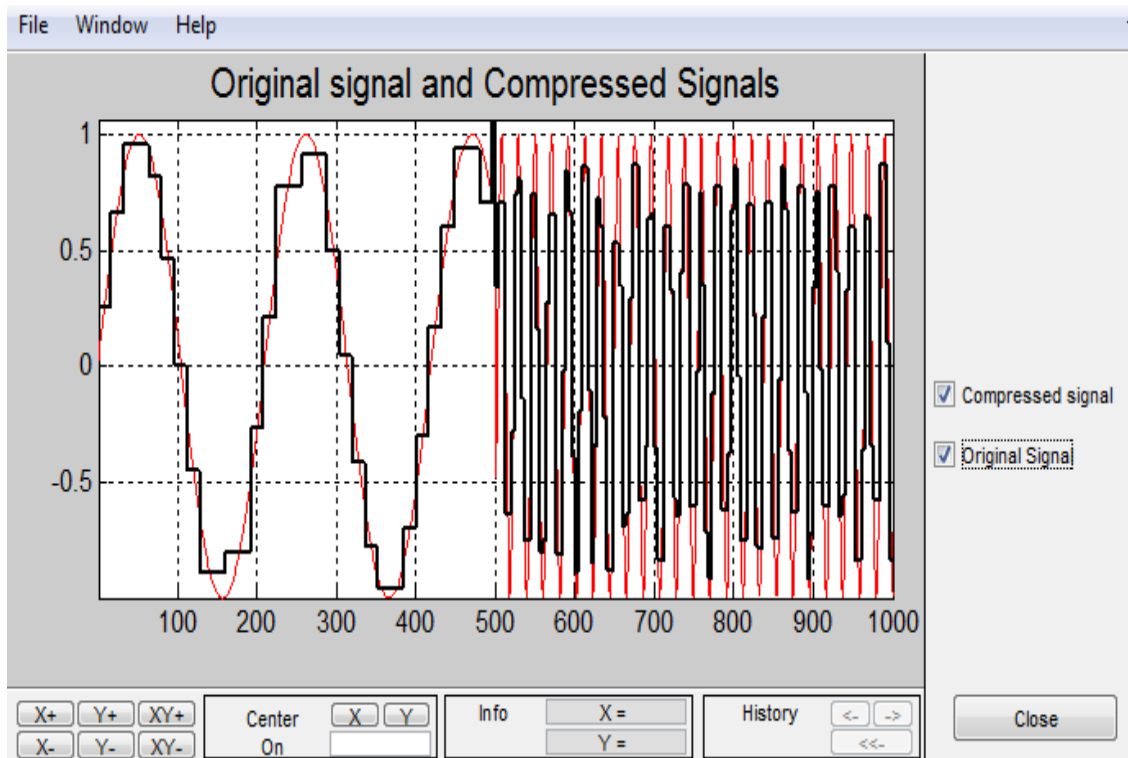


Рис. 2.63. Исходный и сжатый сигнал при помощи вейвлет-преобразования Хаара пятого уровня при коэффициенте сжатия 0.98

Далее необходимо проанализировать зависимость количества сохраненной энергии сигнала от коэффициента сжатия, результаты представлены в таблице 2.7.

Таблица 2.7. Зависимость количества сохраненной энергии от коэффициента сжатия для вейвлет-преобразования Хаара пятого порядка

$K_{сж}$	0,5	1	1,5	2	2,5
$Q_{сохр}$	95,93	89,47	75,37	70,36	65,32

Данную зависимость удобно представить в виде графика (рисунок 2.64):

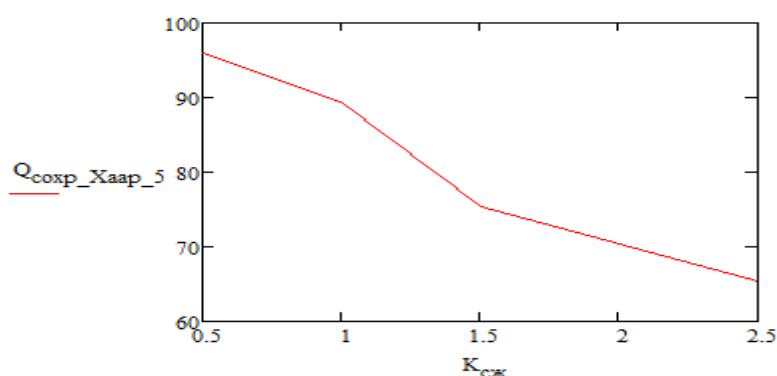


Рис. 2.64. Зависимость количества сохраненной энергии от коэффициента сжатия для вейвлет-преобразования Хаара пятого уровня

Вейвлет Добеши пятого уровня и пятого порядка. Разложение исследуемого сигнала при помощи вейвлета Добеши пятого уровня и пятого порядка представлено на рисунке 2.65.

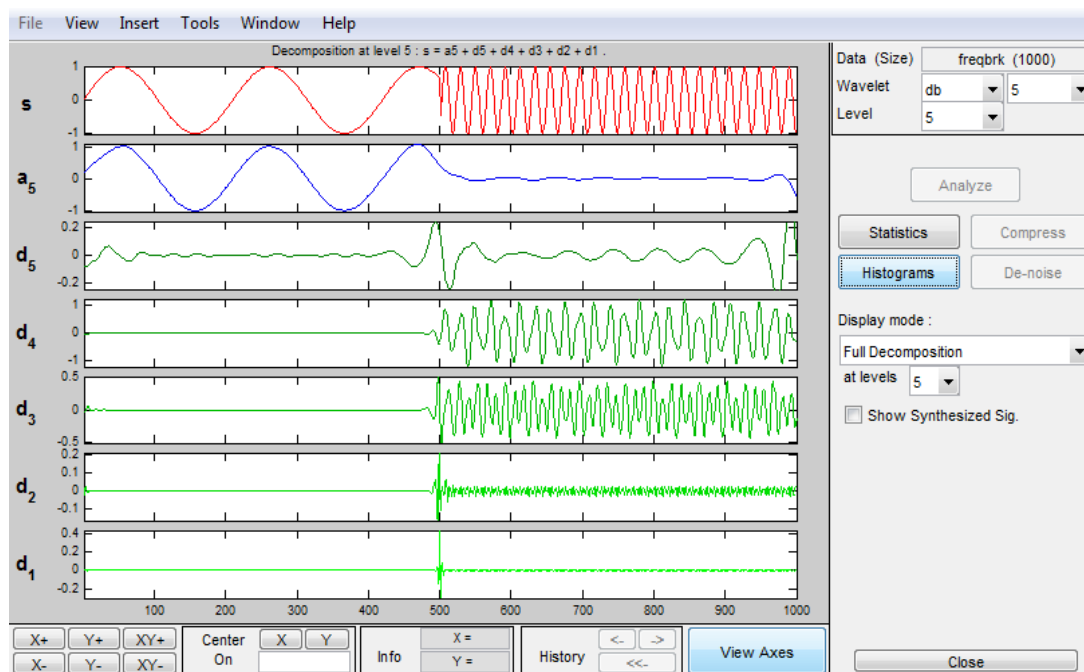


Рис. 2.65. Разложение сигнала при помощи вейвлета Добеши пятого уровня и пятого порядка

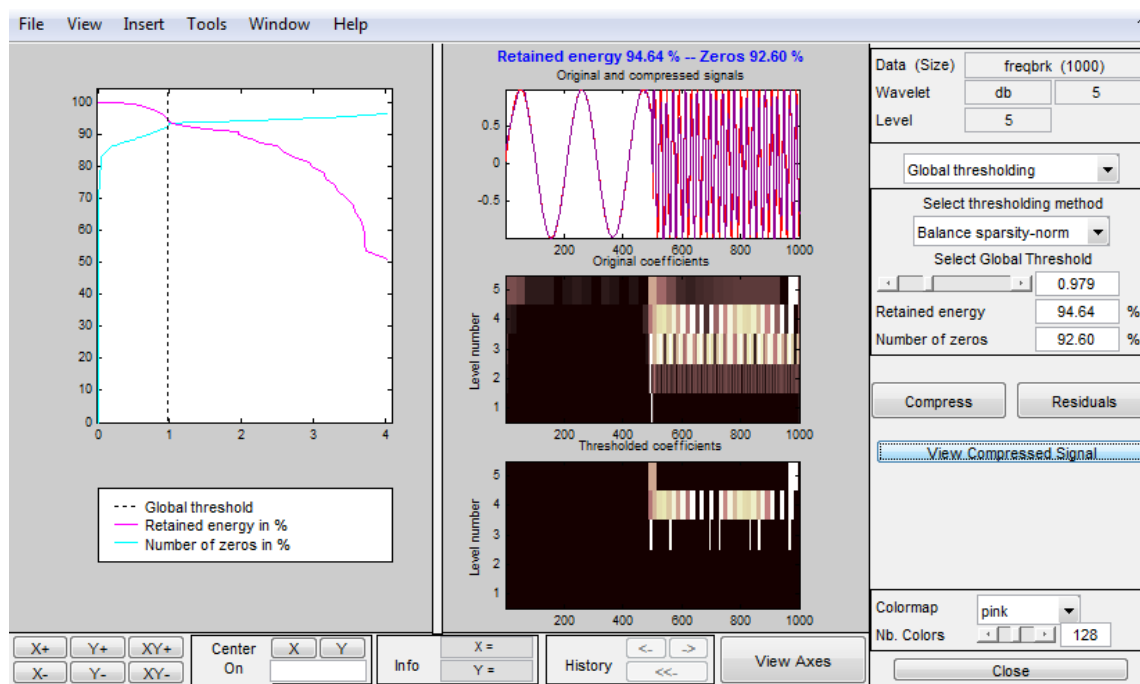


Рис. 2.66. Сжатие сигнала при помощи вейвлета Добеши пятого уровня и пятого порядка

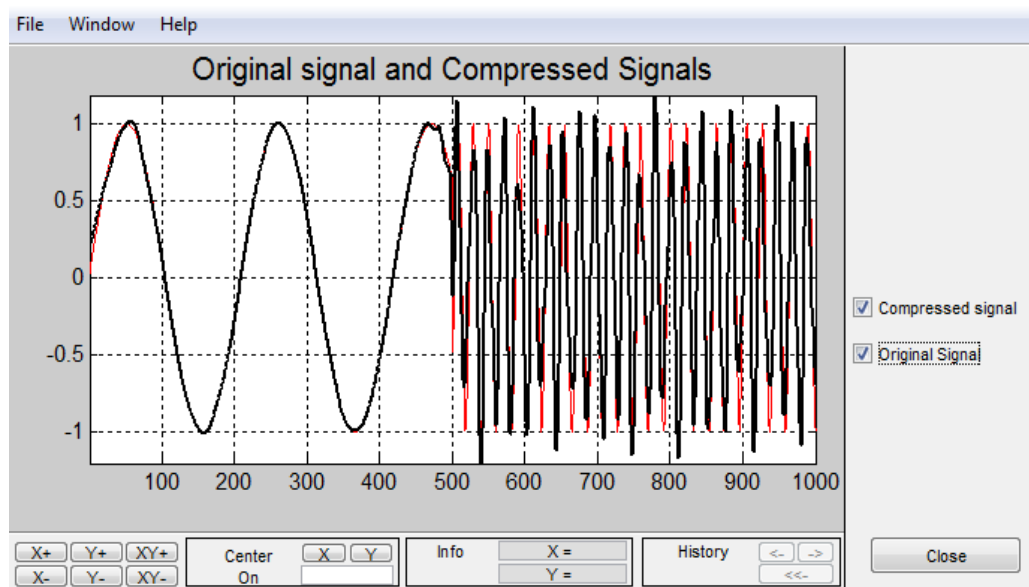


Рис. 2.67. Исходный и сжатый сигнал при помощи вейвлет-преобразования Добеши пятого уровня и пятого порядка при коэффициенте сжатия 0.98

Далее необходимо проанализировать зависимость количества сохраненной энергии сигнала от коэффициента сжатия, результаты представлены в таблице 2.7

Таблица 2.7 – Зависимость количества сохраненной энергии от коэффициента сжатия для вейвлет-преобразования Добеши пятого порядка

$K_{сж}$	0,5	1	1,5	2	2,5
$Q_{сохр}$	99,35	93,61	91,81	89,95	86,41

Данную зависимость удобно представить в виде графика (рисунок 2.68)

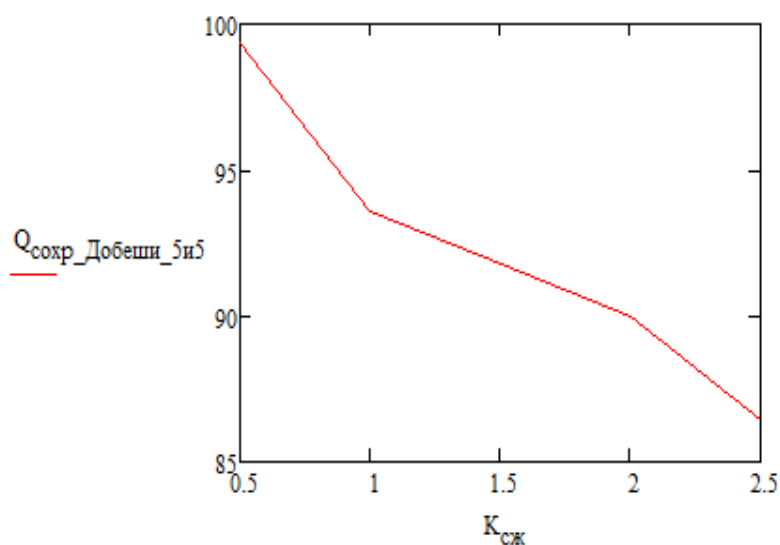


Рис. 2.68. Зависимость количества сохраненной энергии от коэффициента сжатия для вейвлет-преобразования Добеши пятого уровня и пятого порядка

Биортогональное преобразование пятого уровня и порядка 5,5. Разложение исследуемого сигнала при помощи биортогонального преобразования пятого уровня и порядка 5,5 представлено на рисунке 2.70.

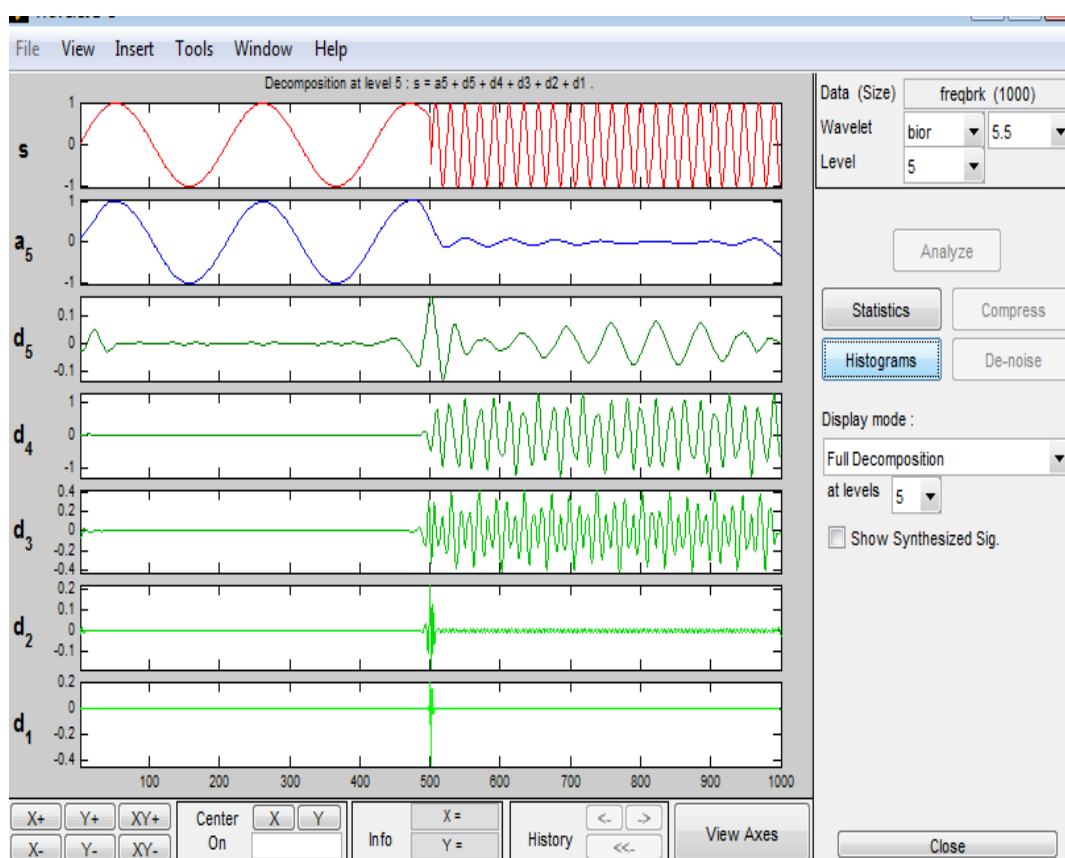


Рис. 2.70. Разложение сигнала при помощи биортогонального преобразования пятого уровня и порядка 5,5

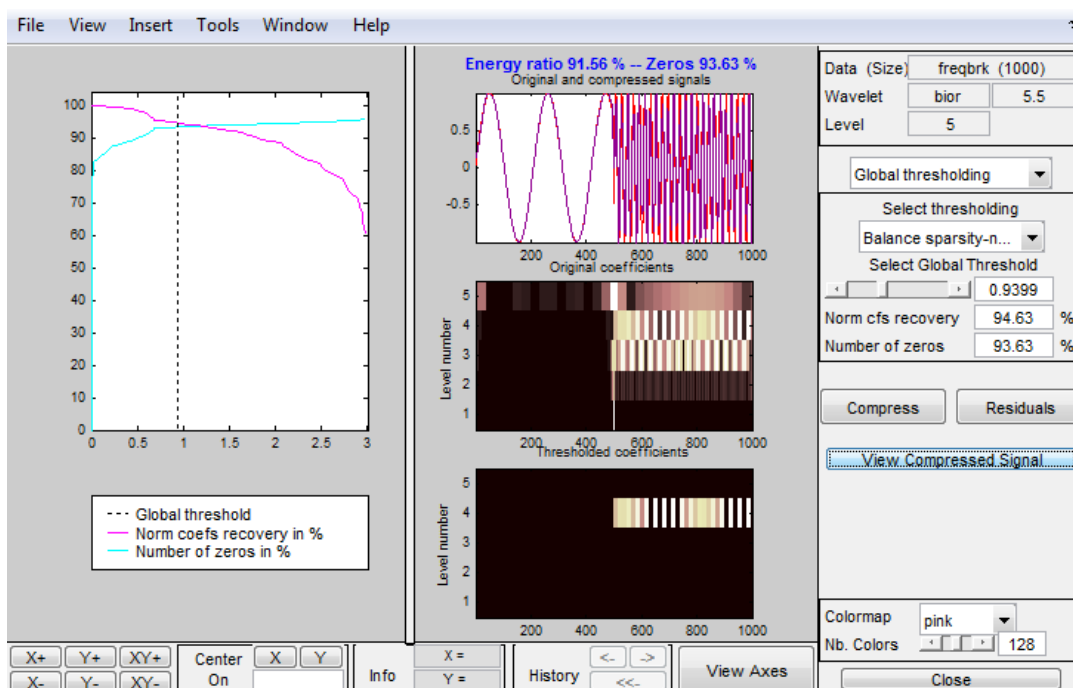


Рис. 2.71. Сжатие сигнала при биортогональном вейвлет-преобразовании пятого уровня и порядка 5,5

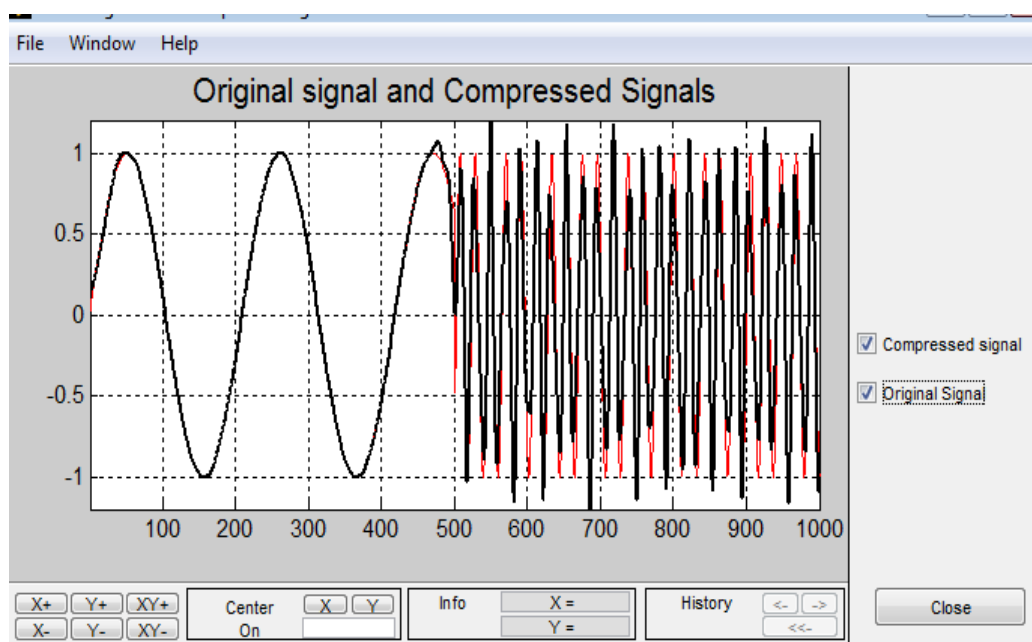


Рис. 2.72. Исходный и сжатый сигнал при помощи биортогонального вейвлет-преобразования пятого уровня и 5,5 порядка при коэффициенте сжатия 0.98

Далее необходимо проанализировать зависимость количества сохраненной энергии сигнала от коэффициента сжатия, результаты представлены в таблице 2.8.

Таблица 2.8. Зависимость количества сохраненной энергии от коэффициента сжатия для вейвлет-преобразования Хаара пятого порядка

$K_{сж}$	0,5	1	1,5	2	2,5
$Q_{сохр}$	98,68	94,63	92,71	88,65	82,28

Данную зависимость удобно представить в виде графика (рисунок 2.73):

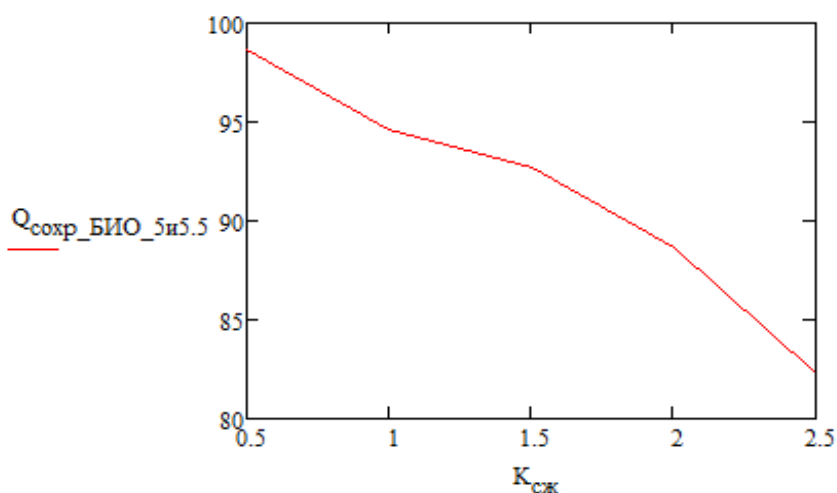


Рис. 2.73. Зависимость количества сохраненной энергии от коэффициента сжатия для биортогонального пятого уровня и 5,5 порядка

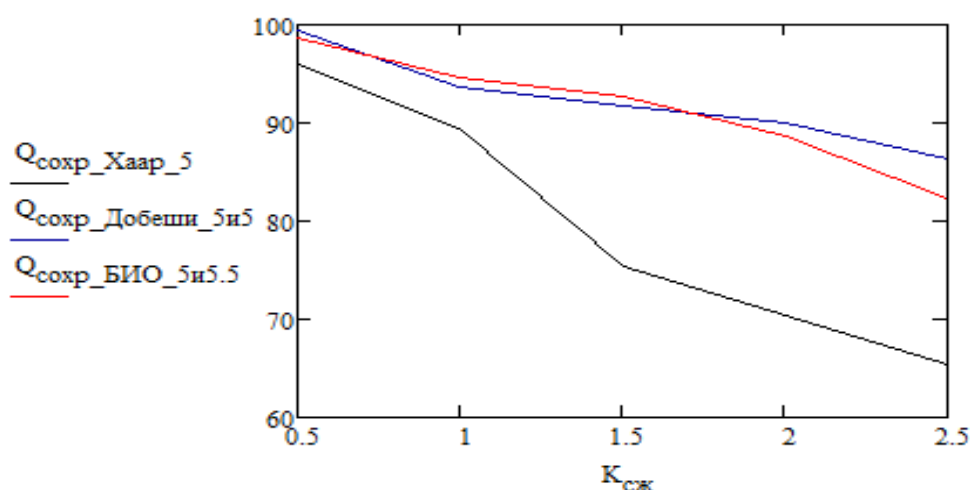


Рис. 2.74. Сравнение энергетических характеристик сжатых сигналов при различных методах вейвлет-преобразований

В результате выполнения данного индивидуального задания был проведен теоретический обзор современных алгоритмов вейвлет-преобразований для сжатия одномерных сигналов.

Было установлено, что использование вейвлет-преобразований является эффективным средством для преобразования информации для передачи по каналам связи, так в современном телекоммуникационном оборудовании предусмотрено использование помехоустойчивых алгоритмов кодирования с использованием избыточности.

Так же, были приобретены практические навыки работы с приложением Wavelet Toolbox Matlab, которое позволяет наглядно продемонстрировать принцип действия сжатия информации путем вейвлет-преобразований.

Были построены графики зависимостей количества сохраненной энергии сигнала от коэффициента сжатия для различных алгоритмов вейвлет-преобразований.

В том числе, было установлено, что использование вейвлет-преобразований Добеши и Хаара эффективнее использования биортогонального вейвлет-преобразования с точки зрения сохранной энергии сигнала.

Данная работа имеет практическое значения, так как была разработана методика для выполнения лабораторной работы по сжатию информации с использованием вейвлет-преобразований при помощи программного обеспечения Matlab 2015.

Таким образом, можно сделать обоснованный вывод о том, что вейвлет-анализ и обработка сигналов на сегодняшний день быстро развивающаяся область прикладной науки, которая в будущем может стать ведущим способом обработки информации. На основе вейвлет-преобразования могут быть построены высокоэффективные алгоритмы сжатия информации, что актуально в наши дни.

ГЛАВА 3. ПОМЕХОУСТОЙЧИВОЕ КОДИРОВАНИЕ В ТЕЛЕКОММУНИКАЦИОННЫХ СИСТЕМАХ

Корректирующие коды в телекоммуникационных системах

В теории современных телекоммуникационных систем значительное внимание уделяется изучению методов кодирования информации. В технике электрической связи широко используют результаты теории кодирования.

Кодирование - операция отождествления символов или групп символов одного кода с символами или группами символов другого кода. Необходимость кодирования возникает, прежде всего, из потребности приспособить форму сообщения к данному каналу связи или к какому-либо другому устройству, предназначенному для преобразования или хранения информации.

Типичная структурная схема системы передачи дискретной информации (СПДИ) приведена на рис. 3.1. Источник вырабатывает сообщения, которые необходимо передавать по каналу СПДИ. Это могут быть последовательности дискретных сообщений (данные, телеграфные сообщения и т.д.) либо непрерывные сообщения (речь, телевидение, результаты телеизмерений и др.), преобразованные в цифровую форму.



Рис. 3.1. Структурная схема системы передачи дискретной информации

Реальные сообщения содержат избыточность и для согласования источника с каналом передачи информации используют кодер источника. Совместно с декодером они образуют кодек источника. Методы кодирования источников изучались в модуле 2. Основная задача любой СДПИ - передача информации с заданными верностью и скоростью передачи. Эти требования находятся в противоречии, причем, повышение скорости передачи информации в реальных СПДИ приводит к снижению помехоустойчивости и верности передачи.

Согласно известных теорем К. Шеннона, в принципе возможно сколь угодно большое повышение верности передачи информации, если скорость передачи по каналу $R_{\text{кан}}$ не превышает пропускной способности канала $C_{\text{к}}$. Достигается это применением достаточно длинных корректирующих кодов (КК).

Корректирующие коды - это коды, позволяющие обнаруживать или исправлять ошибки, возникающие при передаче сообщений по каналам связи.

С этой целью в структуру КК вводится **избыточность**.

Кодек КК (кодер и декодер канала) приведены на рис.3.1. В реальных условиях длина кода ограничена допустимой сложностью устройств кодирования и, прежде всего, декодирования, поэтому эффект от применения корректирующих кодов зависит от параметров кода и ограничений на реализацию кодека канала.

Современная теория предлагает широкий набор корректирующих кодов, различных по структуре, принципам построения и корректирующей способности. В последующих разделах рассмотрены классы кодов, для которых разрабо-

таны достаточно простые и эффективные алгоритмы кодирования/декодирования и которые наиболее перспективны для использования в каналах телекоммуникационных систем.

Классификация корректирующих кодов

В теории и технике помехоустойчивого кодирования известно множество корректирующих кодов, которые могут быть классифицированы по различным признакам. Классификация кодов приведена на рис. 3.2.

По способу формирования КК подразделяются на блочные и непрерывные. Формирование *блочных кодов* предусматривает разбиение передаваемых цифровых последовательностей на отдельные блоки, которые подаются на вход кодера. Каждому такому блоку на выходе кодера соответствует блок кодовых символов, работа кодера определяется правилом, или *алгоритмом кодирования*. Формирование *непрерывных кодов* осуществляется непрерывно во времени, без разделения на блоки, что и определяет наименование этого класса кодов. Блочные коды исторически были предложены и изучены ранее, на заре развития теории кодирования.

В классе непрерывных кодов следует отметить *сверточные коды*, которые по характеристикам превосходят блочные коды, и, по этой причине, находят широкое применение в телекоммуникационных системах. Многие коды носят имена ученых, которые их предложили и исследовали. Таким примером является непрерывный код Финка-Хагельбаргера, предложенный советским ученым Л.М. Финком и немецким специалистом Р. Хагельбаргером. Длительное время этот код служил в литературе показательным примером непрерывного кода с простым алгоритмом кодирования/декодирования, но после открытия сверточных кодов уступил им место.

Для описания процедур кодирования/декодирования как блочных, так и сверточных кодов используют адекватный математический аппарат. Для описания *линейных кодов* используется хорошо разработанный аппарат линейной

алгебры. Формирование нелинейных кодов производится с применением нелинейных процедур. Такой подход позволяет в некоторых случаях получить *нелинейные коды* с рядом специальных свойств. В теории и технике кодирования важной является проблема сложности реализации процедур кодирования/декодирования и, в особенности, процедур декодирования. Поэтому некоторые классы кодов (коды Хемминга, циклические коды Боуза-Чоудхури-Хоквингема, Рида-Соломона, Файра и др.) были разработаны совместно с алгоритмами декодирования, связанными со структурными свойствами этих кодов. И, наоборот, разработка новых алгоритмов декодирования сверточных кодов (алгоритм А. Витерби, последовательное декодирование, пороговое декодирование) инициировала поиски соответствующих сверточных кодов. Отличительные преимущества корректирующих кодов (как блочных, так и сверточных) побуждали поиски новых подходов к реализации путей

повышения помехоустойчивости и эффективности телекоммуникационных систем. На рис. 3.2 отмечены, соответственно, *новые методы кодирования*: сигнально-кодовые конструкции, турбокоды, пространственно-временные коды и т.п.

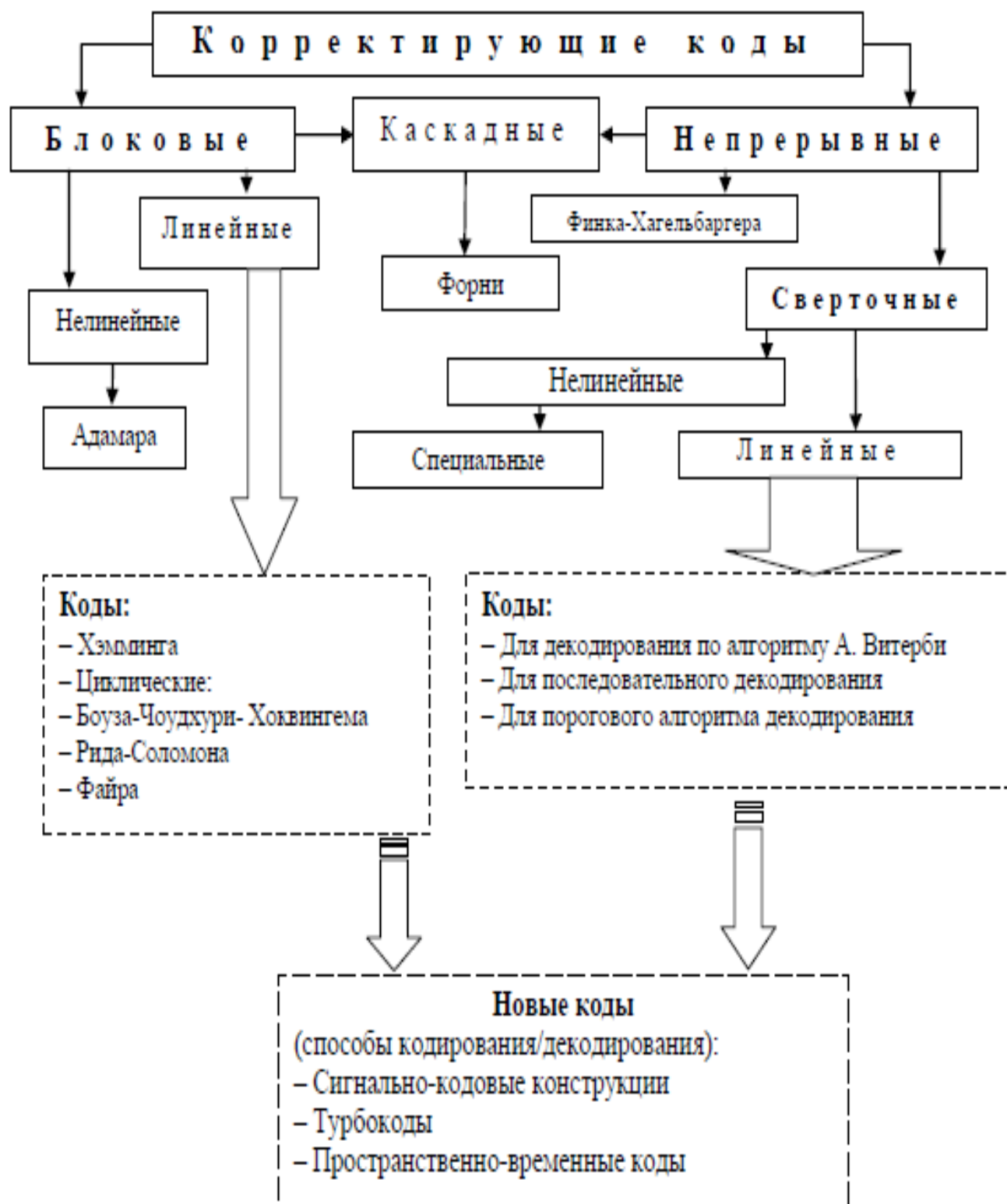


Рис. 3.2. Классификация корректирующих кодов

3.1 Исследование кодов Хемминга, БЧХ (Боуза-Чоудхури-Хоквенгема), Рида-Соломона на базе MATLAB 2015

Код Хэмминга

Наиболее известным линейным кодом является **блоковый код Хэмминга**. В данном пункте, будет рассмотрен (7,4)-код Хэмминга. Линейный блочный код, для которого выполняется неравенство Хэмминга, является кодом Хэмминга, то есть код (7,4) не единственный код, который является данным кодом.

Введем некоторые обозначения, которые составляют основу помехоустойчивого кодирования:

k – количество информационных бит

n – количество бит на выходе кодера

r – количество проверочных(избыточных) бит

Вместо k бит информационного вектора в канал передается n бит кодового вектора. В этом случае говорят об избыточном кодировании со скоростью: $R = \frac{n}{k}$.

Чем ниже скорость, тем больше избыточность кода, и тем большими возможностями для защиты от ошибок он обладает. Однако, следует учитывать, что с увеличением избыточности затраты на передачу информацию также возрастают.

В кодировании и декодировании участвуют генераторная и проверочная матрицы, структура которых следующая:

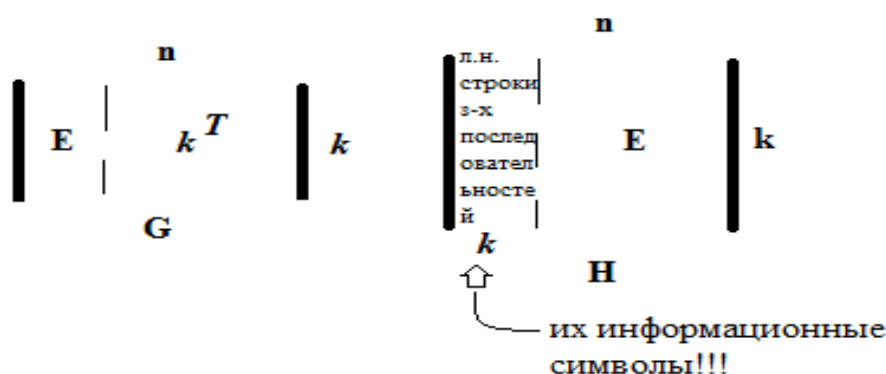


Рис. 3.3. Генераторная и проверочная матрицы

Также, код Хэмминга имеет минимальное кодовое расстояние в пространстве равное трем, из этого следует, что данный код умеет обнаруживать любые двукратные ошибки и исправлять любые однократные.

Рассмотрим код Хэмминга (7,4) в теории.

Для этого кода генераторная и проверочная матрицы уже заготовлены и имеют следующий вид:

$$G_{4 \times 7} = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{pmatrix} \quad H_{3 \times 7} = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{pmatrix}$$

Рис. 3.4. Структура генераторной и проверочной матрицы для кода (7,4)

Кодирование осуществляется следующим образом:

Кодовое слово n и информационное слово k связаны соотношением:

$$n = v = k * G$$

где G — порождающая матрица, структура которой была описана выше.

Например, информационный вектор $k = (1010)$ отобразится в кодовый вектор следующим образом:

$$v = (1010) \times \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{pmatrix} = (0011010)$$

Рис. 3.5. Кодирование кода (7,4)

Легко заметить, что последние четыре разряда кодового вектора совпадают с информационным вектором. Это свойство называется систематичностью кода.

Здесь первые три бита последовательности являются проверочными (избыточными).

Декодирование осуществляется на получении **синдрома** и производится это все следующим образом:

Система проверочных уравнений выглядит:

$$s = n = v * H^T$$

Вектор s принято называть **синдромом**. Таким образом, ошибка будет обнаружена, если хотя бы одна из компонент s не равна нулю.

При передаче информационного слова $a = (1010)$ по каналу без шумавыходной вектор был $n = (0011010)$. Можем убедиться, что в этом случае синдром равен 0.

$$s = (0011010) \times \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} = (000)$$

Рис. 3.6. Получение синдрома для кода (7,4)

Если, например, в кодовом слове произошла одиночная ошибка на четвертой позиции ($r = (0010010)$), то синдромом является четвертая строка транспонированной проверочной матрицы.

$$s = (0010010) \circ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} = (110)$$

Рис. 3.7. Нахождение ошибочного бита для кода (7,4)

Перебрав все возможные позиции одиночной ошибки, получим полную таблицу синдромов однократных ошибок - таблицу соответствий номера ошибочного разряда получающемуся при этом синдрому. Таким образом, производится декодирование. Если ошибки найдены, то они соответственно исправляются.

Проведение эксперимента и обработка результатов в MATLAB 2015

Задание:

1. Собрать схему
2. Подготовить схемы для реализации кодов Хэмминга (7,4) и (15,11) основываясь на примере, представленном в отчете.

При коде (7,4) в параметрах кодера и декодера MessageLengthK, от M-degreeprimitivepolynomial: устанавливается gfprimfd(3,'min'), а при коде (15,11) gfprimfd(4,'min').

3. Для полученных кодов изменять вероятность ошибки в пределах от 0 до 1 (не менее 4-х точек) и снимать с дисплеев полученные входные последовательности, закодированные последовательности, декодированные последовательности и ошибки.

4. Построить графики зависимости числа обнаруженных ошибок от вероятности ошибки для кодов (7,4) и (15,11).

В рабочем поле необходимо собрать схему для работы кода Хэмминга (7,4). Схема представлена на рисунке 3.8.

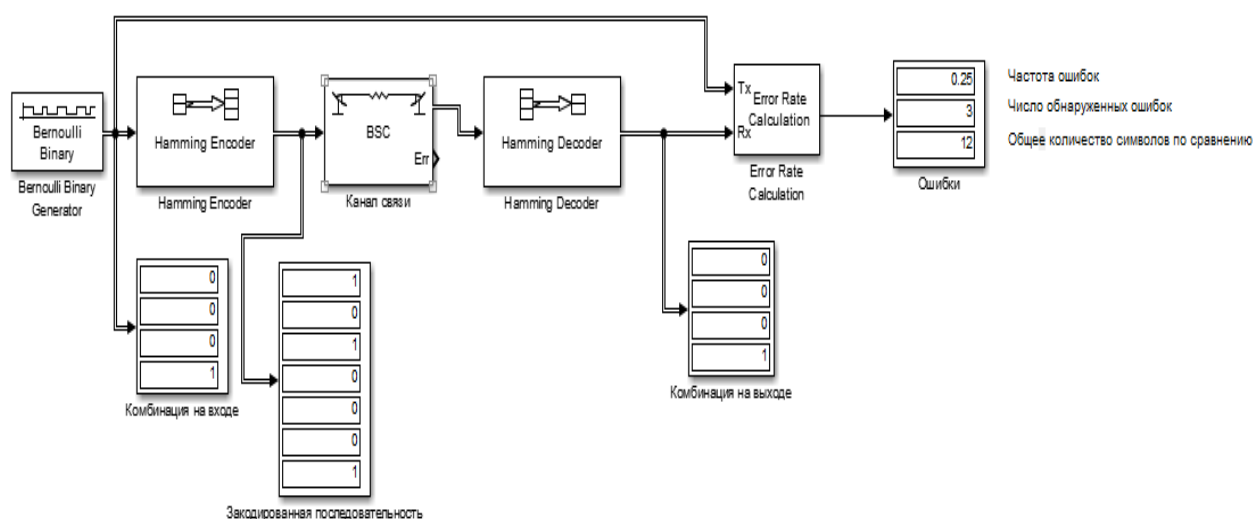


Рис. 3.8. Линия передачи с применением кода Хэмминга

В состав линии с кодированием входят:

1. BernoulliBinaryGenerator
2. HammingEncoder
3. Binary Symmetric Channel (каналпередачи)
4. HammingDecoder
5. Error RateCalculation (анализаторошибок)
6. Display

Устанавливаем характеристики блоков для кода (7,4)

Bernoulli Binary Generator

Generate a Bernoulli random binary number.
To generate a vector output, specify the probability as a vector.

Parameters

Probability of a zero:

Initial seed:

Sample time:

Frame-based outputs

Samples per frame:

Output data type:

Рис. 3.9. Параметры Bernoulli Binary Generator

Hamming Encoder (mask) (link)

Create a Hamming code with message length K and codeword length N . The number N must have the form $2^M - 1$, where M is an integer greater than or equal to 3. K must equal $N - M$.

The input must contain exactly K elements. If it is frame-based, then it must be a column vector.

Parameters

Codeword length N :

Message length K , or M -degree primitive polynomial:

Рис. 3.10. Параметры Hamming Encoder

Binary Symmetric Channel (mask) (link)

Add binary errors to the input signal.

Parameters

Error probability:

Initial seed:

Output error vector

Output data type:

Рис. 3.11. Параметры Binary Symmetric Channel

Hamming Decoder (mask) (link)

Recover a binary message vector from a binary Hamming codeword vector. The message is of length K and the codeword is of length N, where N has the form $2^M - 1$, for some integer M greater than or equal to 3. K must equal N-M.

The input must contain exactly N elements. If it is frame-based, then it must be a column vector.

Parameters

Codeword length N:

Message length K, or M-degree primitive polynomial:

Рис. 3.12. Параметры Hamming Decoder

Error Rate Calculation (mask) (link)

Compute the error rate of the received data by comparing it to a delayed version of the transmitted data. The block output is a three-element vector consisting of the error rate, followed by the number of errors detected and the total number of symbols compared. This vector can be sent to either the workspace or an output port.

The delays are specified in number of samples, regardless of whether the input is a scalar or a vector. The inputs to the 'Tx' and 'Rx' ports must be scalars or column vectors.

The 'Stop simulation' option stops the simulation upon detecting a target number of errors or a maximum number of symbols, whichever comes first.

Parameters

Receive delay:

Computation delay:

Computation mode:

Output data:

Рис. 3.13 Параметры Error Rate Calculation



Рис. 3.14. Комбинация на входе, Закодированная последовательность, Комбинация на выходе, Ошибки (вероятность ошибок равна 0,2)

Анализируя рисунок выше, можно сделать вывод, что комбинация на входе совпадает с комбинацией на выходе, таким образом, передача осуществилась удачно. Что касается

ошибок, то их частота равна 0,25, число обнаруженных ошибок равно 3, общее количество символов по сравнению равно 12. Кодирование и декодирование здесь осуществляется методом описанном в выше.

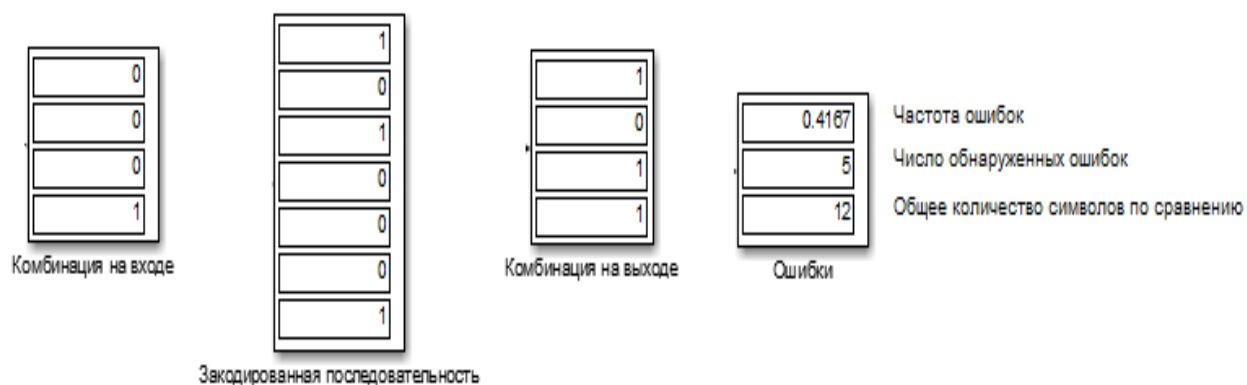


Рис. 3.15. Комбинация на входе, Закодированная последовательность, Комбинация на выходе, Ошибки (вероятность ошибок равна 0,4)

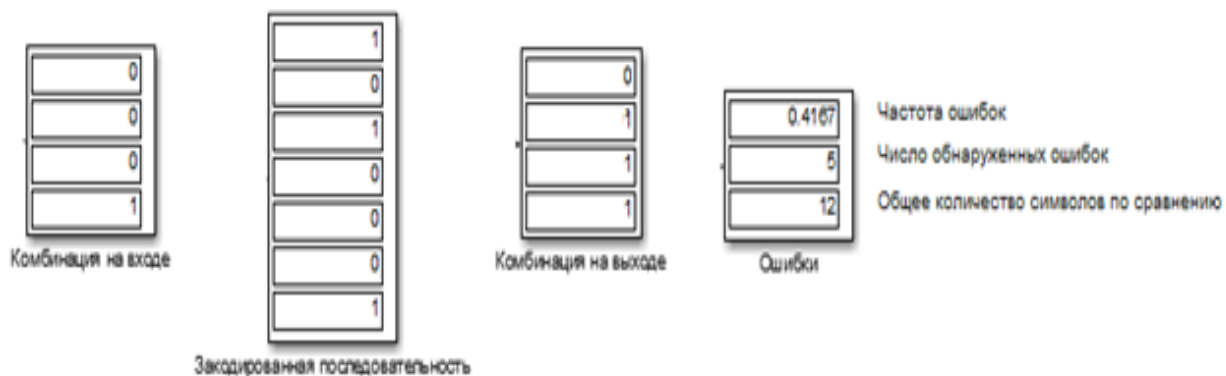


Рис. 3.16. Комбинация на входе, Закодированная последовательность, Комбинация на выходе, ошибки (вероятность ошибок равна 0,6)

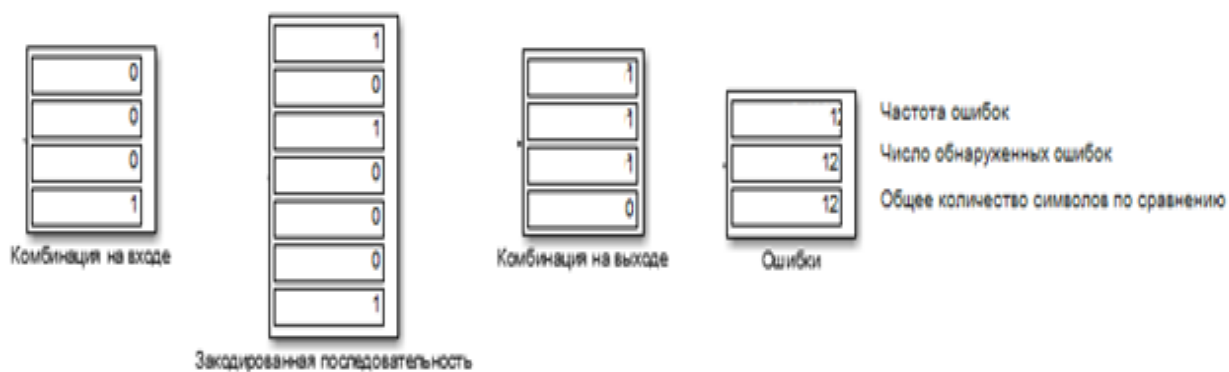


Рис. 3.17. Комбинация на входе, Закодированная последовательность, Комбинация на выходе, Ошибки (вероятность ошибок равна 0,8)

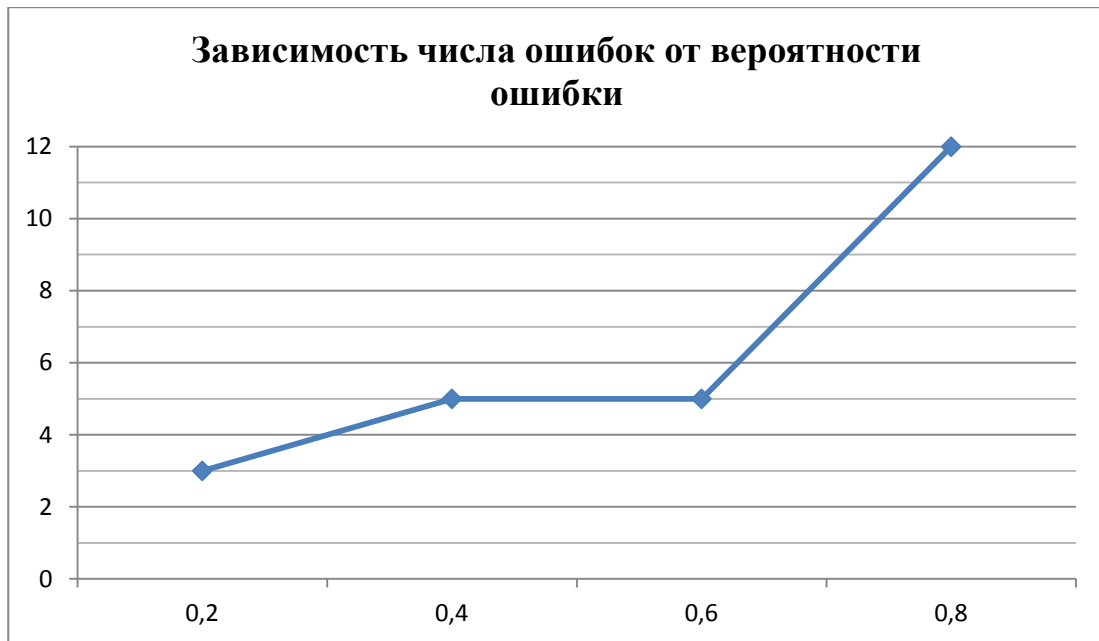


Рис. 3.18. График зависимости числа ошибок (OY) от вероятности ошибки (OX) для кода Хэмминга (7,4)

Устанавливаем характеристики блоков для кода (15,11)

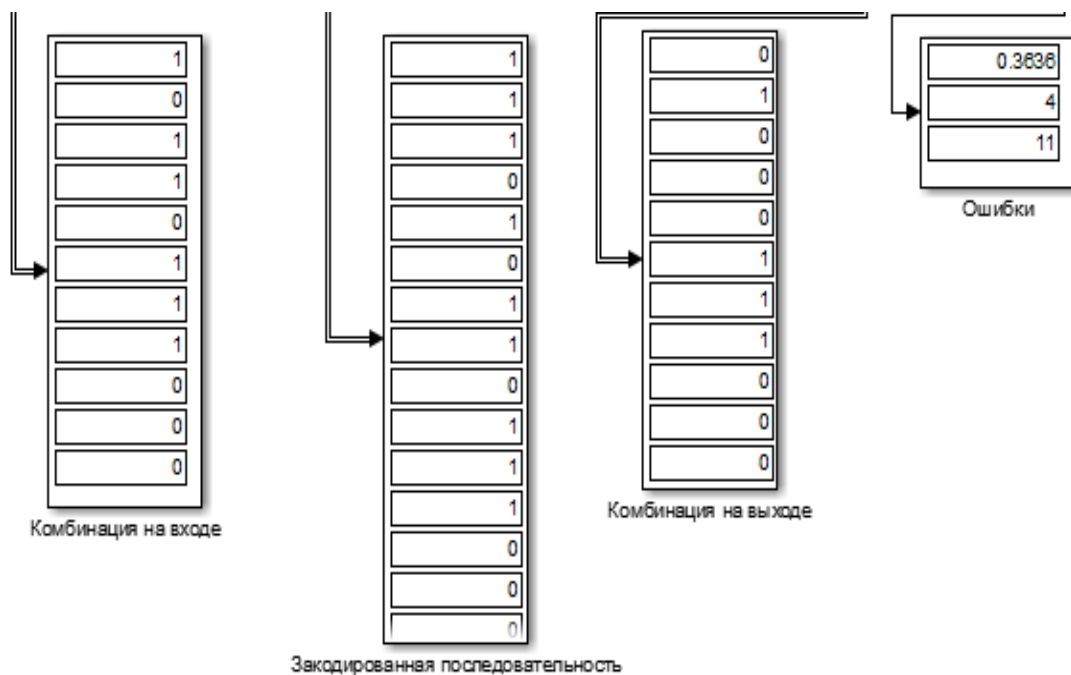


Рис. 3.19. Комбинация на входе, Закодированная последовательность, Комбинация на выходе, Ошибки (вероятность ошибок равна 0,2)

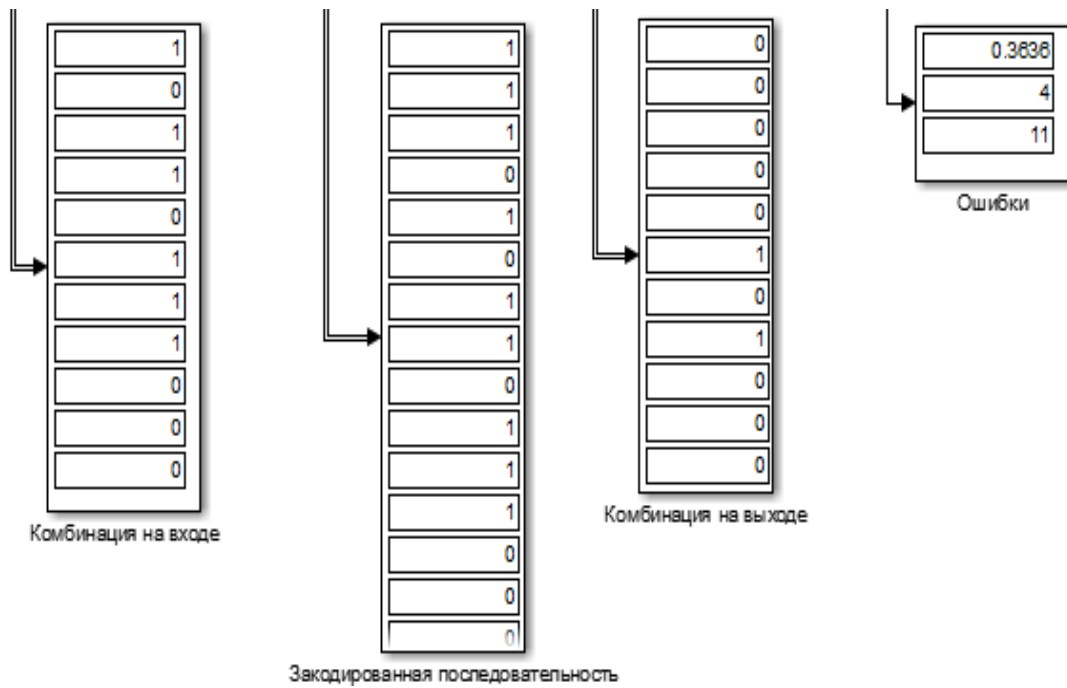


Рис. 3.20. Комбинация на входе, Закодированная последовательность, Комбинация на выходе, Ошибки (вероятность ошибок равна 0,4)

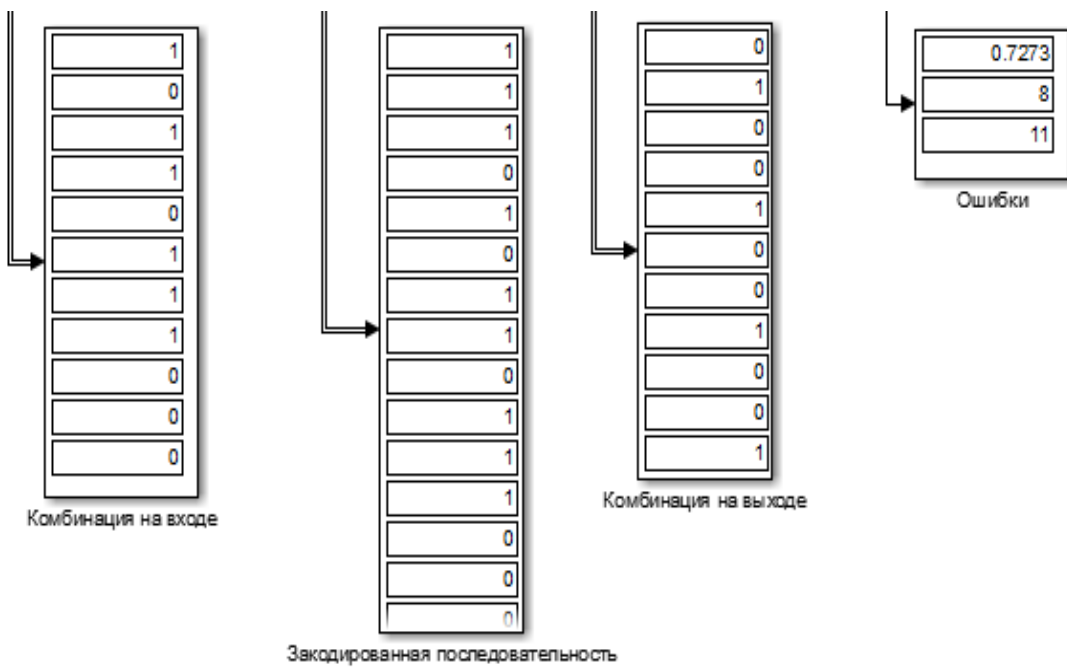


Рис. 3.21. Комбинация на входе, Закодированная последовательность, Комбинация на выходе, Ошибки (вероятность ошибок равна 0,6)

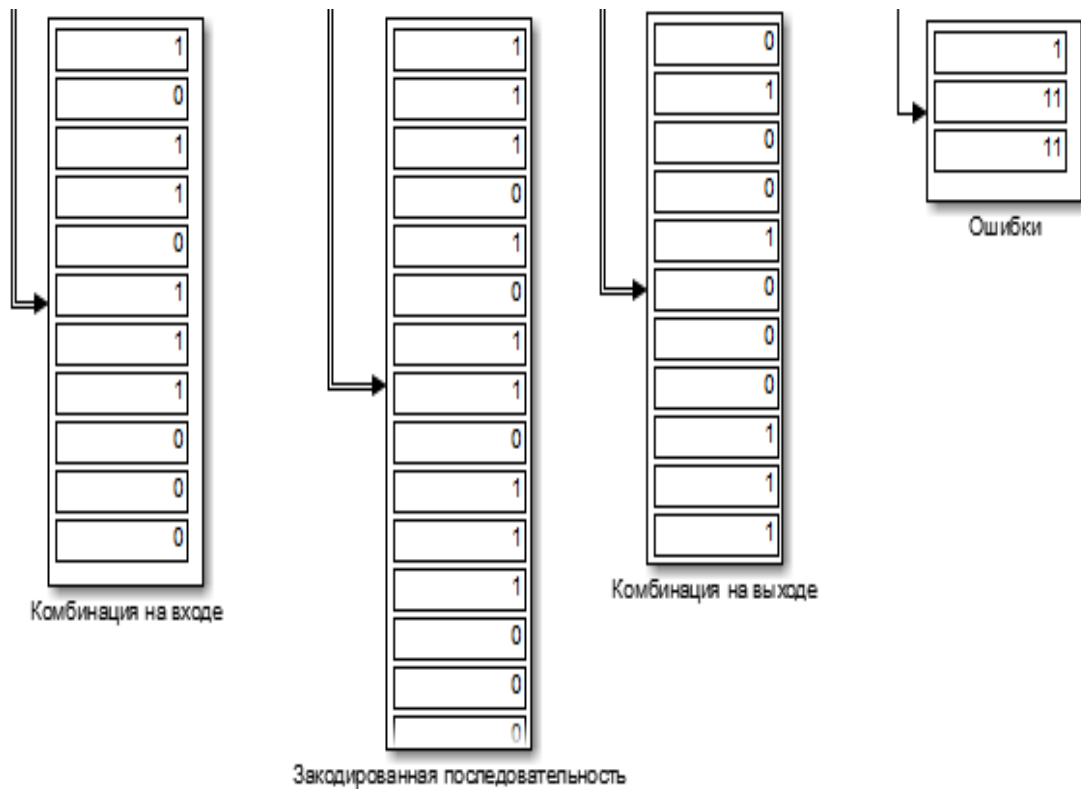


Рис. 3.22. Комбинация на входе, Закодированная последовательность, Комбинация на выходе, Ошибки (вероятность ошибок равна 0,8)



Рис. 3.23. График зависимости числа ошибок (OY) от вероятности ошибки (OX) для кода Хэмминга (15,11)

В результате проверки построена схема линии передачи с кодированием Хэмминга в среде Simulink. Построены графики зависимостей числа ошибок на выходе декодера от вероятности ошибки в канале связи для кодов (7,4) и (15,11).

Из графиков (рисунок 3.18 и рисунок 3.23) видно, что число ошибок увеличивается с ростом вероятности ошибок.

Код БЧХ (Боуза-Чоудхури-Хоквенгема)

Код БЧХ является циклическим кодом. С циклическим кодом Хэмминга у них много чего общего, а именно алгоритм кодирования, который отличается только нахождением генераторного полинома, а процесс декодирования полностью схож. Необходимо начать с небольшого введения в код БЧХ. Многочлен $g(x)$ степени называется *примитивным*, если $x^j + 1$ делится на $g(x)$ без остатка для $j = 2^k - 1$ и не делится ни для какого меньшего значения (где k – количество информационных бит). Например, многочлен $1 + x^2 + x^3$ примитивен: он делит $x^7 + 1$, но не делит $x^j + 1$ при $j < 7$. Примитивен также многочлен $1 + x^3 + x^4$ – он делит $x^{15} + 1$, но не делит $x^j + 1$ при $j < 15$ (для кода (7,15)).

Кодирующий многочлен $g(x)$ для БЧХ-кода, длина кодовых слов которого n , строится так. Находится примитивный многочлен минимальной степени такой, что $n \leq 2^q - 1$ или $n \geq \log_2(n + 1)$. Пусть α – корень этого многочлена, тогда рассмотрим кодирующий многочлен $g(x) = \text{НОК}(m_1(x), \dots, m_{d-1}(x))$, где $m_1(x), \dots, m_{d-1}(x)$ – многочлены минимальной степени, имеющие корнями соответственно $\alpha, \alpha^2, \dots, \alpha^{d-1}$.

Построенный кодирующий многочлен производит код с минимальным расстоянием между кодовыми словами, не меньшим d , и длиной кодовых слов n .

Кодирование:

Например, нужно построить БЧХ-код с длиной кодовых слов $n=15$ и минимальным расстоянием между кодовыми словами $d=5$. Степень примитивного многочлена равна $q = \log_2(n + 1) = 4$ и сам он равен $1 + x^3 + x^4$ (примитивный многочлен 4-ой степени для кода (7,15)). Пусть α – его корень, тогда α^2 и α^4 – также его корни. Минимальным многочленом для α^3 будет $1 + x + x^2 + x^3 + x^4$. Следовательно,

$$\begin{aligned} g(x) &= \text{НОК}(1 + x^3 + x^4, 1 + x + x^2 + x^3 + x^4) = (1 + x^3 + x^4)(1 + x + x^2 + x^3 + x^4) \\ &= 1 + x + x^2 + x^4 + x^8 \end{aligned}$$

Степень полученного многочлена равна 8, построенный BCH-код будет (7,15) кодом. Слово 1000100 или $a(x) = x^4 + 1$ будет закодировано кодовым словом $a(x)g(x) = x^{12} + x^6 + x^5 + x^2 + x + 1$ или 111001100000100. На практике будет рассмотрен код BCH (15,7) и BCH (15,11).

Декодирование:

Декодирование производится путем деления закодированной последовательности на генераторный полином, который использовался при кодировании. Полученная последовательность при делении и будет декодированной последовательностью. В лучшем случае, остатка при делении не будет, это значит, что ошибок не выявлено.

Если же остаток есть, называется он не иначе как *синдром*, в таком случае ошибки присутствуют в закодированной последовательности. В этом случае поступают так. Закодированную последовательность складывают с вектором ошибок по модулю два, и исправляют ошибочный бит. Вектор ошибок формируется с помощью специальных схем, которые анализируя закодированную последовательность, формируют данный вектор.

Еще одним способ исправления ошибок является следующий метод. На основании полученного генераторного полинома строится схема и на каждом такте ее работы определяется *синдром*. В данном случае необходимо получить два синдрома, один из них это остаток, от деления полученный при декодировании, а второй это комбинация 100. Далее производится вычитание номера такта комбинации 100 и номера такта остатка. Полученная разность и является номером бита в закодированной комбинации.

Проведение эксперимента и обработка результатов

Задание:

1. Собрать схему
2. Подготовить схемы для реализации кодов BCH (15,7) и (15,11) основываясь на примере, представленном в отчете.
3. Для полученных кодов изменять вероятность ошибки в пределах от 0 до 1 (не менее 4-х точек) и снимать с дисплеев полученные входные последовательности, закодированные последовательности, декодированные последовательности и ошибки.
4. Построить графики зависимости числа обнаруженных ошибок от вероятности ошибки для кодов (15,7) и (15,11).
5. Все поэтапное исследование представить в отчете.

В рабочем поле необходимо собрать схему для работы кода BCH (15,7). Схема представлена на рисунке 3.24.

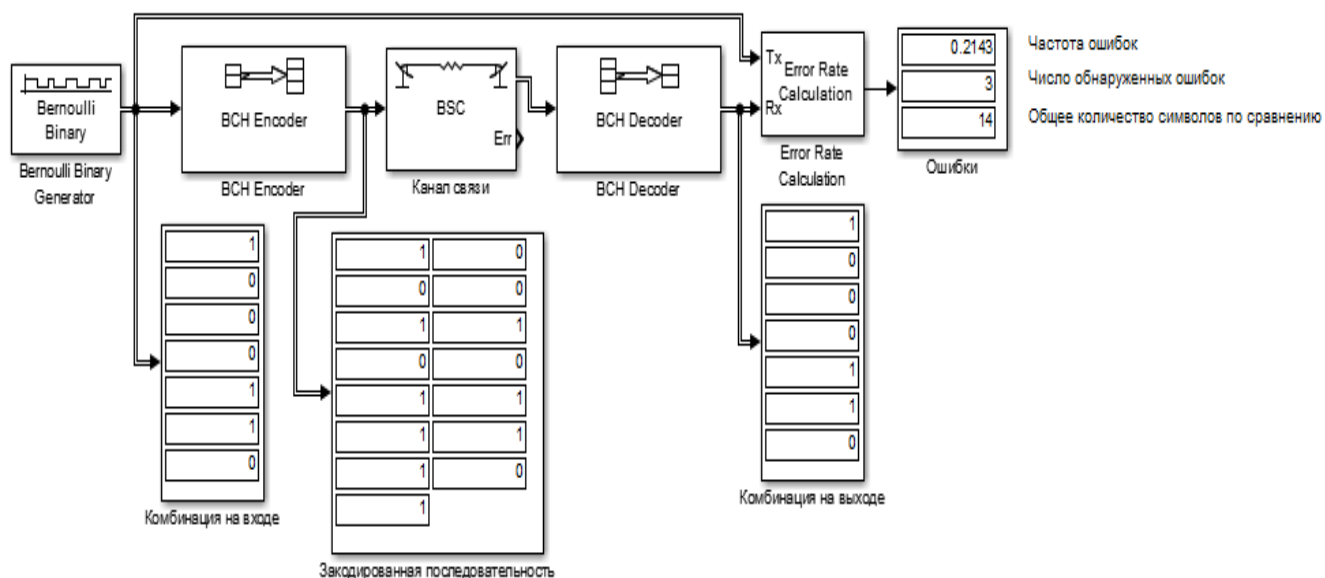


Рис. 3.24. Линия передачи с применением кода БЧХ (15,7)

В состав линии с кодированием входят:

1. BernoulliBinaryGenerator
2. BCHEncoder
3. Binary Symmetric Channel (каналпередачи)
4. BCHDecoder
5. Error RateCalculation (анализаторошибок)
6. Display

Устанавливаем характеристики блоков для кода (15,7)

Bernoulli Binary Generator

Generate a Bernoulli random binary number.
To generate a vector output, specify the probability as a vector.

Parameters

Probability of a zero:

Initial seed:

Sample time:

Frame-based outputs

Samples per frame:

Output data type:

Рис. 3.25. Параметры Bernoulli Binary Generator

BCH Encoder (mask) (link)

Encode the message in the input vector using an (N,K) BCH encoder with the narrow-sense generator polynomial. This block accepts a column vector input signal with an integer multiple of K elements. Each group of K input elements represents one message word to be encoded. The values of N and K must produce a valid narrow-sense BCH code.

Shorten the code by setting the shortened message length parameter S. In this case, use full length N and K values to specify the (N, K) code that is shortened to an (N - K + S, S) code.

Parameters

Codeword length, N:
15

Message length, K:
7

Рис. 3.26. Параметры BCH Encoder

Binary Symmetric Channel (mask) (link)

Add binary errors to the input signal.

Parameters

Error probability:
0.2

Initial seed:
71

Output error vector

Output data type: boolean

Рис. 3.27. Параметры Binary Symmetric Channel

BCH Decoder (mask) (link)

Decode the message in the input vector using an (N,K) BCH decoder with the narrow-sense generator polynomial. This block accepts a column vector input signal with an integer multiple of N elements. Each group of N input elements represents one codeword to be decoded. The values of N and K must produce a valid narrow-sense BCH code.

Shorten the code by setting the shortened message length parameter S. In this case, use full length N and K values to specify the (N, K) code that is shortened to an (N - K + S, S) code.

Parameters

Codeword length, N:
15

Message length, K:
7

Рис. 3.28 Параметры BCH Decoder

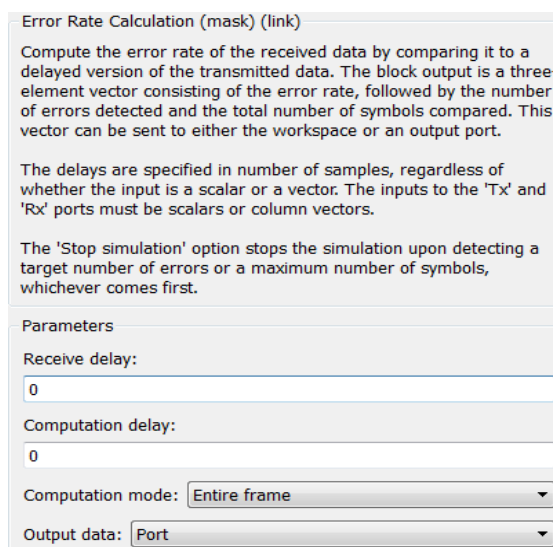


Рис. 3.29. Параметры Error Rate Calculation

Представим полученные результаты

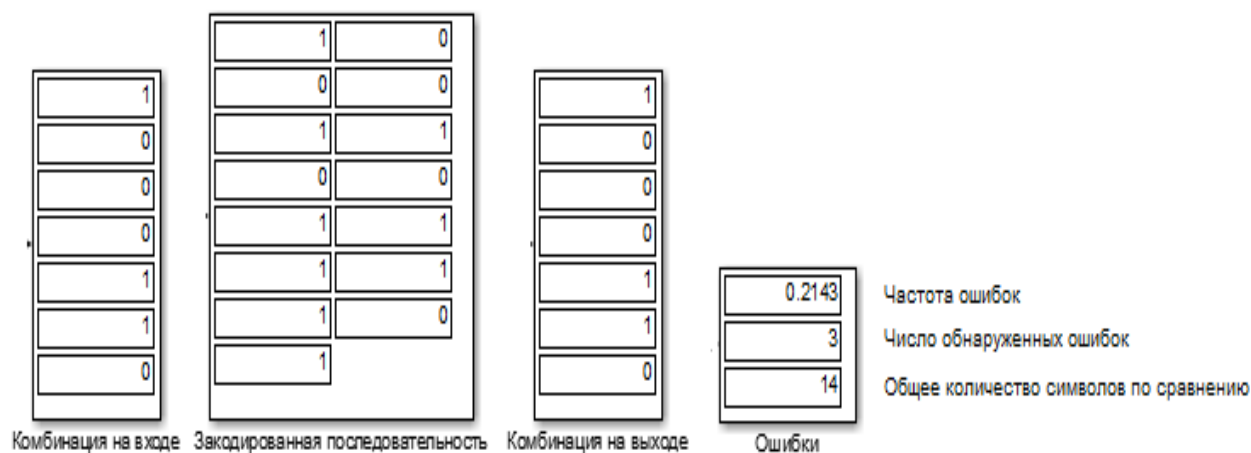


Рис. 3.30. Комбинация на входе, Закодированная последовательность, Комбинация на выходе, Ошибки (вероятность ошибок равна 0,2)

Можно сделать вывод, что комбинация на входе совпадает с комбинацией на выходе, таким образом, передача осуществилась удачно. Что касается ошибок, то их частота равна 0,2143, число обнаруженных ошибок равно 3, общее количество символов по сравнению равно 14. Кодирование и декодирование здесь осуществляется методом описанном выше.

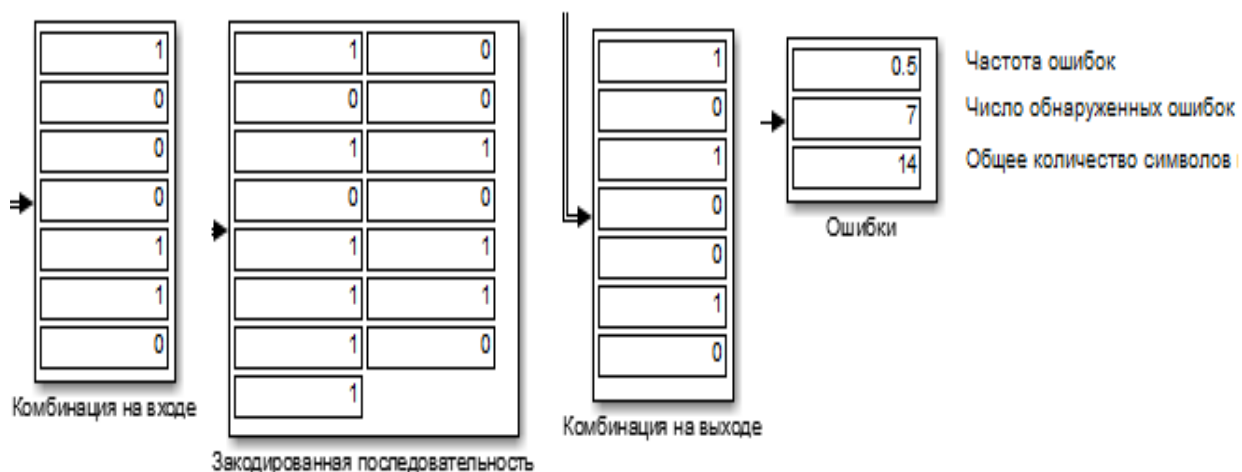


Рис. 3.31. Комбинация на входе, Закодированная последовательность, Комбинация на выходе, Ошибки (вероятность ошибок равна 0,4)

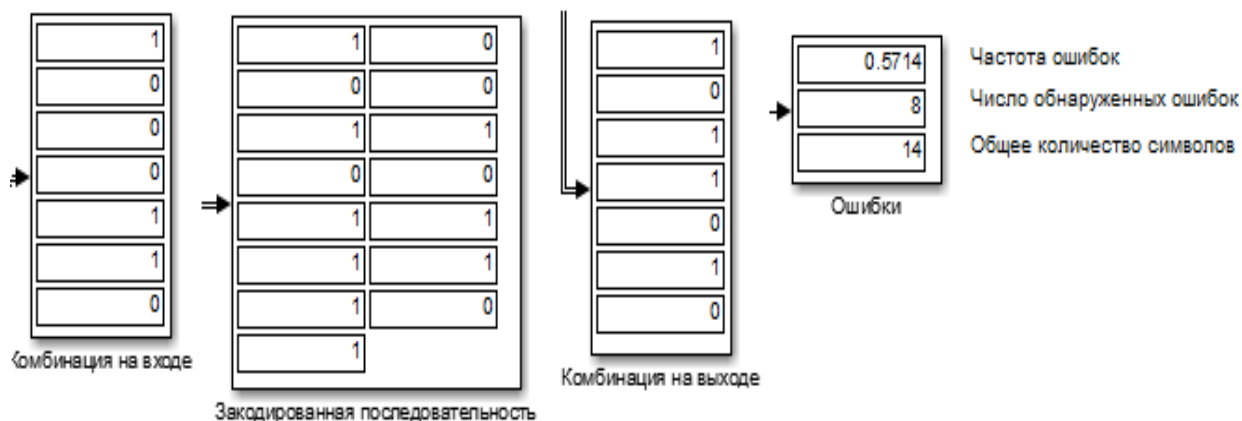


Рис. 3.32. Комбинация на входе, Закодированная последовательность, Комбинация на выходе, Ошибки (вероятность ошибок равна 0,6)

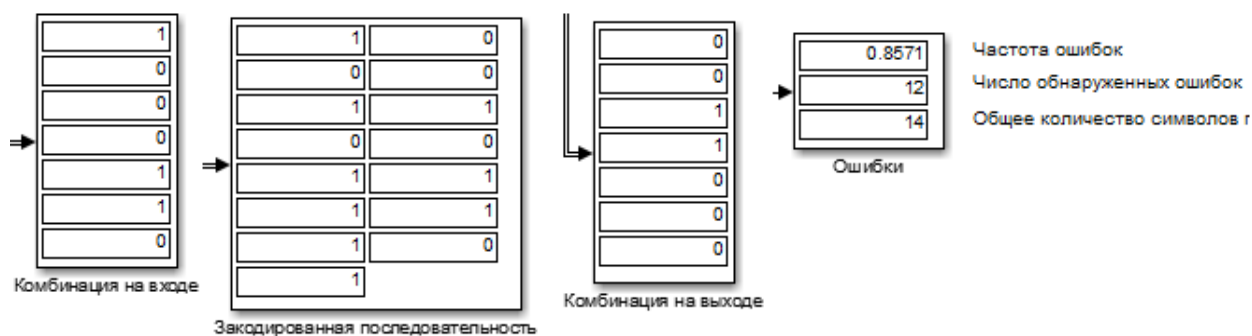


Рис. 3.33. Комбинация на входе, Закодированная последовательность, Комбинация на выходе, Ошибки (вероятность ошибок равна 0,8)

Устанавливаем характеристики блоков для кода (15,11)

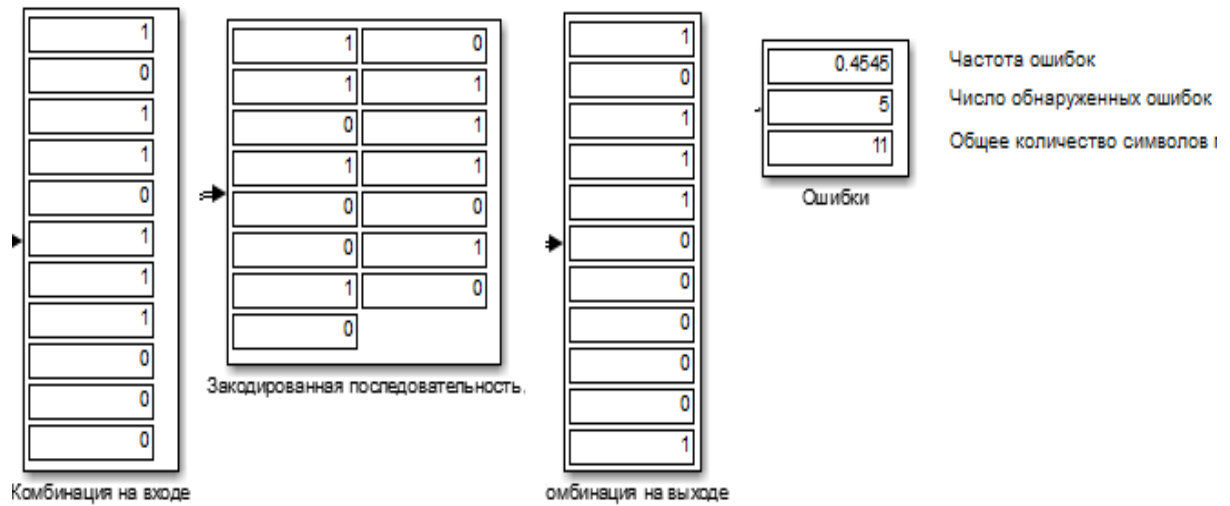


Рис. 3.34. Комбинация на входе, Закодированная последовательность, Комбинация на выходе, Ошибки (вероятность ошибок равна 0,2)

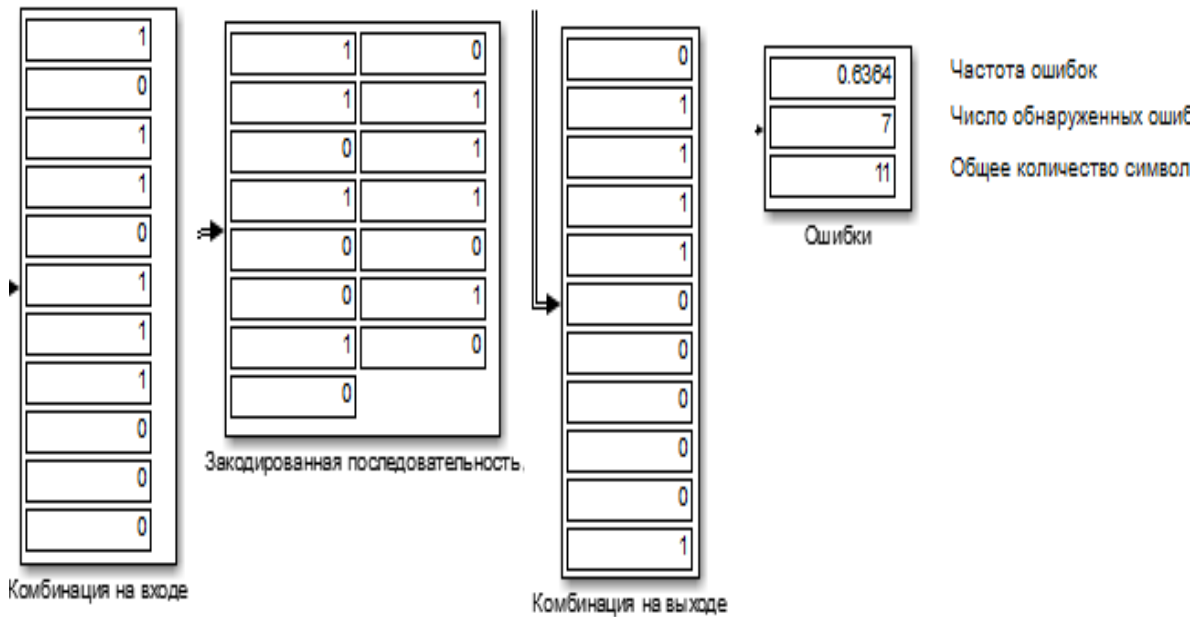


Рис. 3.35. Комбинация на входе, Закодированная последовательность, Комбинация на выходе, Ошибки (вероятность ошибок равна 0,4)

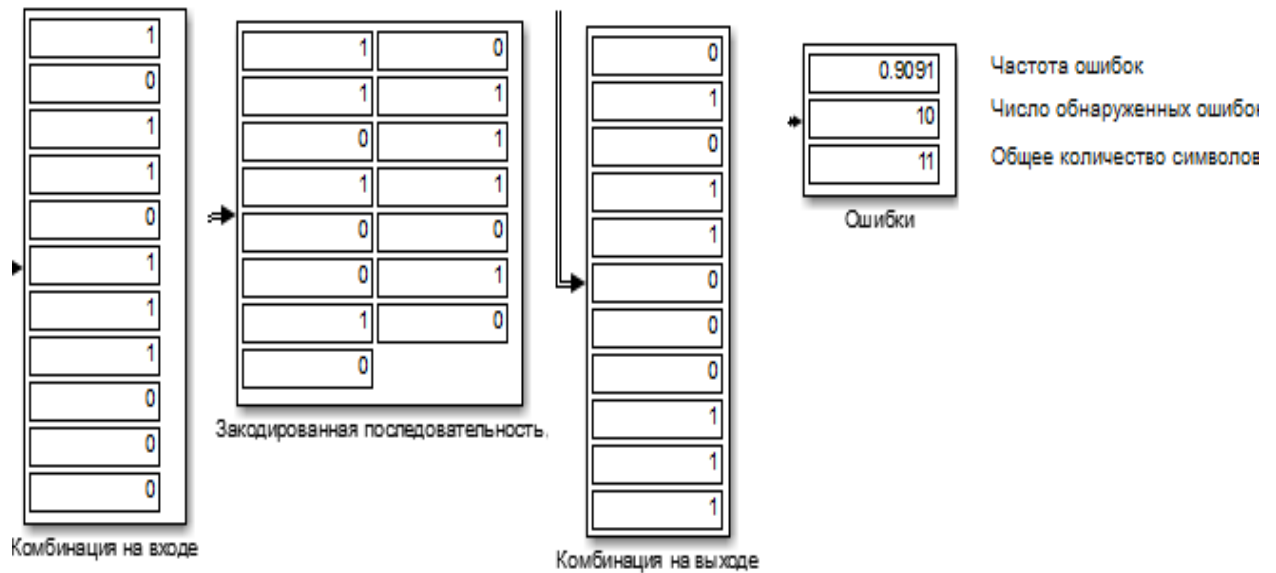


Рис. 3.36 Комбинация на входе, Закодированная последовательность, Комбинация на выходе, Ошибки (вероятность ошибок равна 0,6)

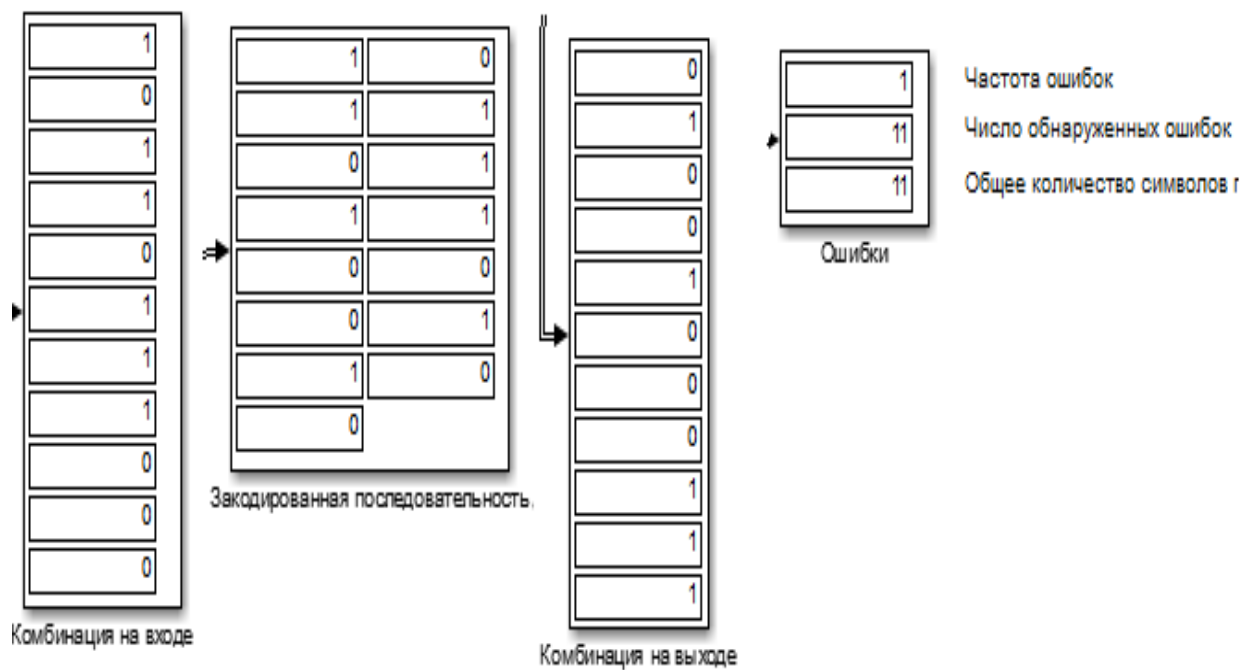


Рис. 3.37. Комбинация на входе, Закодированная последовательность, Комбинация на выходе, Ошибки (вероятность ошибок равна 0,8)

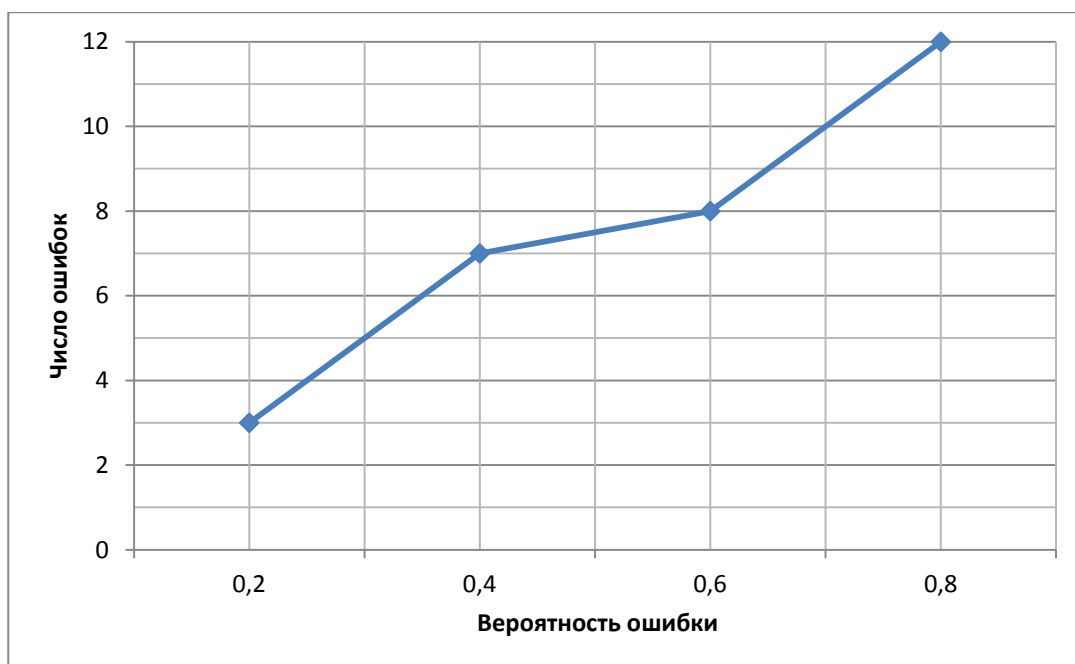


Рис. 3.38. График зависимости числа ошибок от вероятности ошибки для кода БЧХ (15,7)

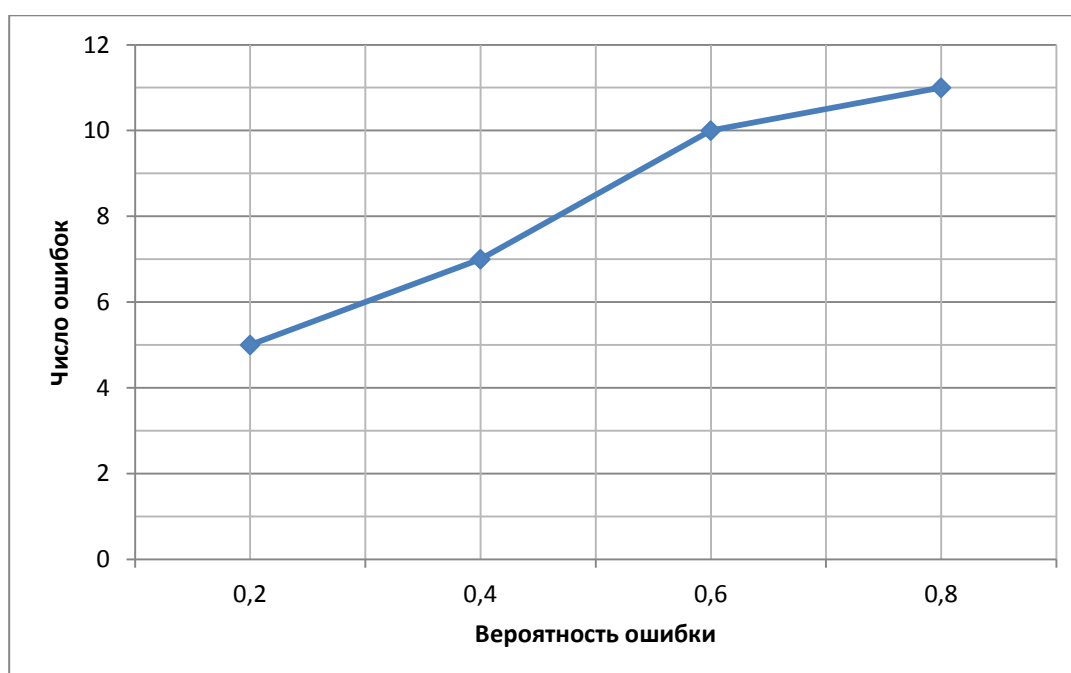


Рис. 3.39. График зависимости числа ошибок от вероятности ошибки для кода БЧХ (15,11)

В результате практической работы построена схема линии передачи с кодированием БЧХ в среде Simulink. Построены графики зависимостей числа ошибок на выходе декодера от вероятности ошибки в канале связи для кодов (15,7) и (15,11). Из графиков (рисунок 3.38 и

рисунок 3.39) видно, что число ошибок увеличивается с ростом вероятности ошибок. При этом код (15,7) оказался лучше по способности обнаружения ошибок.

Код Рида-Соломона

Кодировщик Рида-Соломона берет блок цифровых данных и добавляет дополнительные "избыточные" биты. Ошибки происходят при передаче по каналам связи или по разным причинам при запоминании (например, из-за шума или наводок, царапин на CD и т.д.). Декодер Рида-Соломона обрабатывает каждый блок, пытается исправить ошибки и восстановить исходные данные. Число и типы ошибок, которые могут быть исправлены, зависят от характеристик кода Рида-Соломона.

Свойства кодов Рида-Соломона

Коды Рида-Соломона являются подмножеством кодов БЧХ и представляют собой линейные блочные коды. Код Рида-Соломона специфицируется как RS (n,k) s - битных символов.

Это означает, что кодировщик воспринимает k информационных символов по s битов каждый и добавляет символы четности для формирования n символьного кодового слова. Декодер Рида-Соломона может корректировать до t символов, которые содержат ошибки в кодовом слове, где $2t = n - k$.

Диаграмма, представленная ниже, показывает типовое кодовое слово Рида-Соломона:

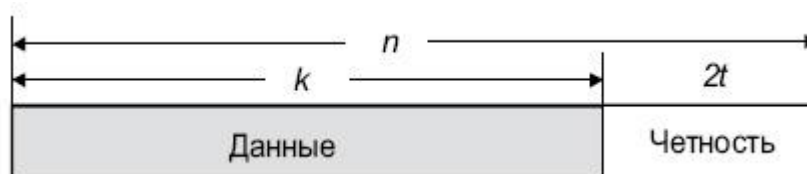


Рис. 3.40. Кодовое слово для кода Р-С

Популярным кодом Рида-Соломона является RS (255, 223) с 8-битными символами. Каждое кодовое слово содержит 255 байт, из которых 223 являются информационными и 32 байтами четности. Для этого кода $n = 255$, $k = 223$, $s = 8$, $2t = 32$, $t = 16$.

Декодер может исправить любые 16 символов с ошибками в кодовом слове: то есть ошибки могут быть исправлены, если число искаженных байт не превышает 16.

При размере символа s , максимальная длина кодового слова n для кода Рида-Соломона равна $n = 2^s - 1$.

Например, максимальная длина кода с 8-битными символами ($s=8$) равна 255 байтам.

Коды Рида-Соломона могут быть в принципе укорочены путем обнуления некоторого числа информационных символов на входе кодировщика (передать их в этом случае не нужно). При передаче данных декодеру эти нули снова вводятся в массив.

Код (255, 223), описанный выше, может быть укорочен до (200, 168). Кодировщик будет работать с блоком данных 168 байт, добавит 55 нулевых байт, сформирует кодовое слово (255, 223) и передаст только 168 информационных байт и 32 байта четности.

Объем вычислительной мощности, необходимой для кодирования и декодирования кодов Рида-Соломона, зависит от числа символов четности. Большое значение t означает, что большее число ошибок может быть исправлено, но это потребует большей вычислительной мощности по сравнению с вариантом при меньшем t .

Ошибки в символах

Одна ошибка в символе происходит, когда 1 бит символа оказывается неверным или когда все биты неверны.

Код RS(255,223) может исправить до 16 ошибок в символах. В худшем случае, могут иметь место 16 битовых ошибок в разных символах (байтах). В лучшем случае, корректируются 16 полностью неверных байт, при этом исправляется $16 * 8 = 128$ битовых ошибок.

Коды Рида-Соломона особенно хорошо подходят для корректировки кластеров ошибок (когда неверными оказываются большие группы бит кодового слова, следующие подряд).

Декодирование

Алгебраические процедуры декодирования Рида-Соломона могут исправлять ошибки и потери. Потерей считается случай, когда положение неверного символа известно. Декодер может исправить до t ошибок или до $2t$ потерь. Данные о потере (стирании) могут быть получены от демодулятора цифровой коммуникационной системы, т.е. демодулятор помечает полученные символы, которые вероятно содержат ошибки.

Когда кодовое слово декодируется, возможны три варианта.

1. Если $2s + r < 2t$ (s ошибок, r потеря), тогда исходное переданное кодовое слово всегда будет восстановлено. В противном случае:
2. Декодер детектирует ситуацию, когда он не может восстановить исходное кодовое слово. Или же:
3. Декодер некорректно декодирует и неверно восстановит кодовое слово без какого-либо указания на этот факт.

Вероятность каждого из этих вариантов зависит от типа используемого кода Рида-Соломона, а также от числа и распределения ошибок.

Архитектура кодирования и декодирования кодов Рида-Соломона

Кодирование и декодирование Рида-Соломона может быть выполнено аппаратно или программно.

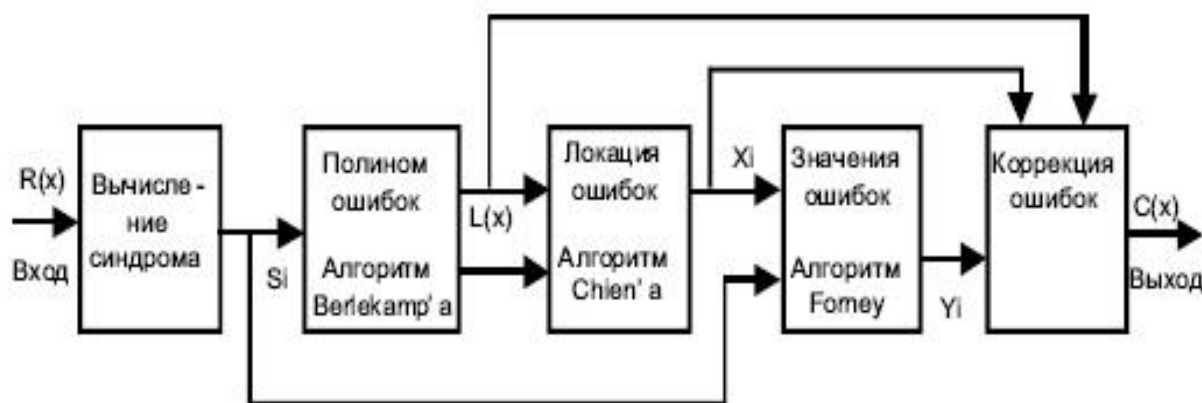


Рис. 3.42. Схема работы с кодами Рида-Соломона

Обозначения:

- $r(x)$ – Полученное кодовое слово
- S_i – Синдромы
- $L(x)$ – Полином локации ошибок
- X_i – Положения ошибок
- Y_i – Значения ошибок
- $c(x)$ – Восстановленное кодовое слово
- v – Число ошибок

Полученное кодовое слово $r(x)$ представляет собой исходное (переданное) кодовое слово $c(x)$ плюс ошибки:

$$r(x) = c(x) + e(x)$$

Декодер Рида-Соломона пытается определить позицию и значение ошибки для t ошибок (или $2t$ потерь) и исправить ошибки и потери.

Вычисление синдрома

Вычисление синдрома похоже на вычисление четности. Кодовое слово Рида-Соломона имеет $2t$ синдромов, это зависит только от ошибок (а не передаваемых кодовых слов). Синдромы могут быть вычислены путем подстановки $2t$ корней образующего полинома $g(x)$ в $r(x)$.

Нахождение позиций символьных ошибок

Это делается путем решения системы уравнений с t неизвестными. Существует несколько быстрых алгоритмов для решения этой задачи. Эти алгоритмы используют особенности структуры матрицы кодов Рида-Соломона и сильно сокращают необходимую вычислительную мощность. Делается это в два этапа.

1. Определение полинома локации ошибок.

Это может быть сделано с помощью алгоритма Berlekamp-Massey или алгоритма Эвклида. Алгоритм Эвклида используется чаще на практике, так как его легче реализовать,

однако алгоритм Berlekamp-Massey позволяет получить более эффективную реализацию оборудования и программ.

2. Нахождение корней этого полинома. Это делается с привлечением алгоритма поиска Chien.

Нахождение значений символьных ошибок

Здесь также нужно решить систему уравнений с t неизвестными. Для решения используется быстрый алгоритм Forney.

Реализация кодировщика и декодера Рида-Соломона. Аппаратная реализация

Существует несколько коммерческих аппаратных реализаций. Имеется много разработанных интегральных схем, предназначенных для кодирования и декодирования кодов Рида-Соломона. Эти ИС допускают определенный уровень программирования (например, RS (255, k), где t может принимать значения от 1 до 16).

Программная реализация

До недавнего времени программные реализации в "реальном времени" требовали слишком большой вычислительной мощности практически для всех кодов Рида-Соломона. Главной трудностью в программной реализации кодов Рида-Соломона являлось то, что процессоры общего назначения не поддерживают арифметические операции для поля Галуа. Однако оптимальное составление программ в сочетании с возросшей вычислительной мощностью позволяют получить вполне приемлемые результаты для относительно высоких скоростей передачи данных.

Практическая часть исследования

На данном этапе необходимо реализовать и проанализировать описанные выше алгоритмы кодирования на практике в программной среде Matlab 15. Необходимо построить соответствующие схемы, которые будут в себя включать источник сообщения, кодер, декодер, канал передачи, анализатор ошибок и соответственно устройства визуализации полученных результатов. Так же необходимо сравнить ошибки передачи с линией, которая не будет никак преобразовывать передаваемую информацию, а просто передавать ее по каналу связи и при увеличении вероятности ошибки в канале до 0,3. Коды, которые подлежат реализации это: код Хэмминга (7,4), код BCH (15,7), и код Рида-Соломона (255,223). Всю задачу выполнить в среде Matlab 15. Представить проделанную работу в отчете.

Для начала необходимо запустить саму программу. Происходит это следующим образом:

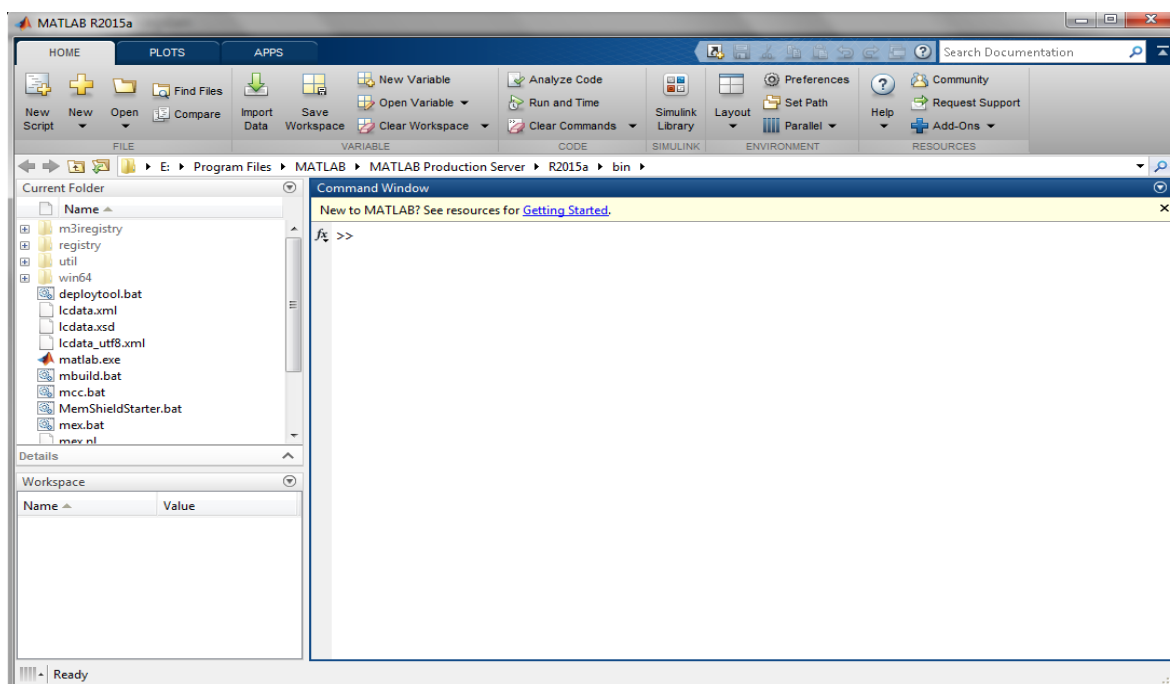


Рис. 3.43. Окно Matlab 15

Далее необходимо нажать на вкладку «Simulink Library», которая находится на панели инструментов в открывшемся окне. Откроется следующее окно:

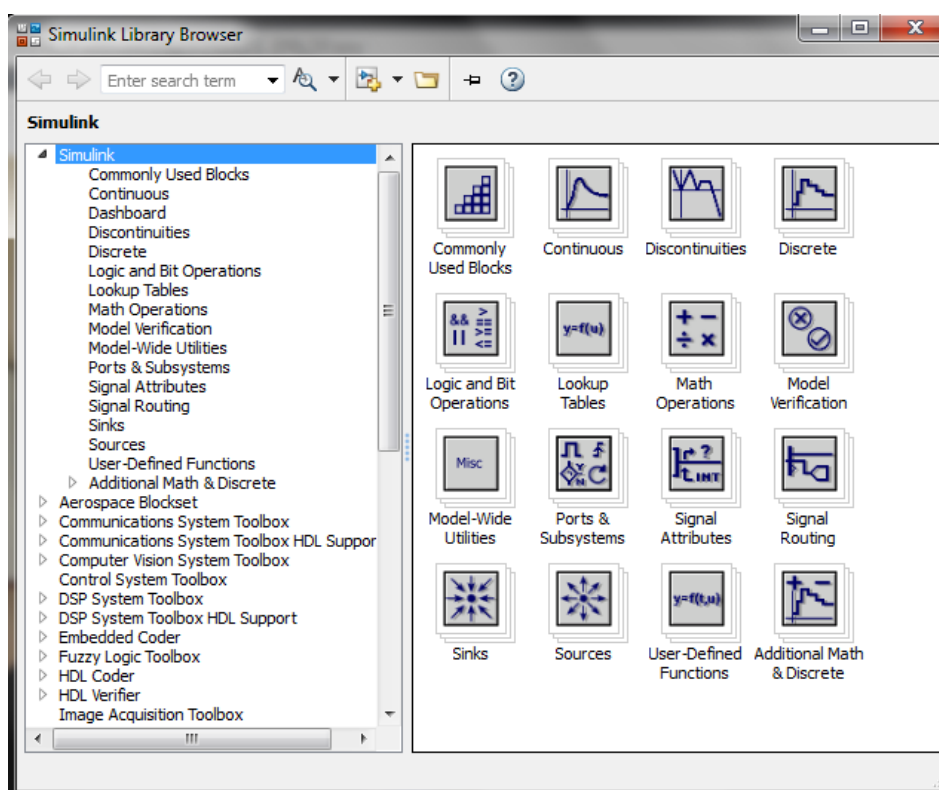


Рис. 3.44. Окно «Simulink Library»

Далее создаем новую модель (вкладка на панели инструментов) и приступаем к реализации описанных выше алгоритмов.

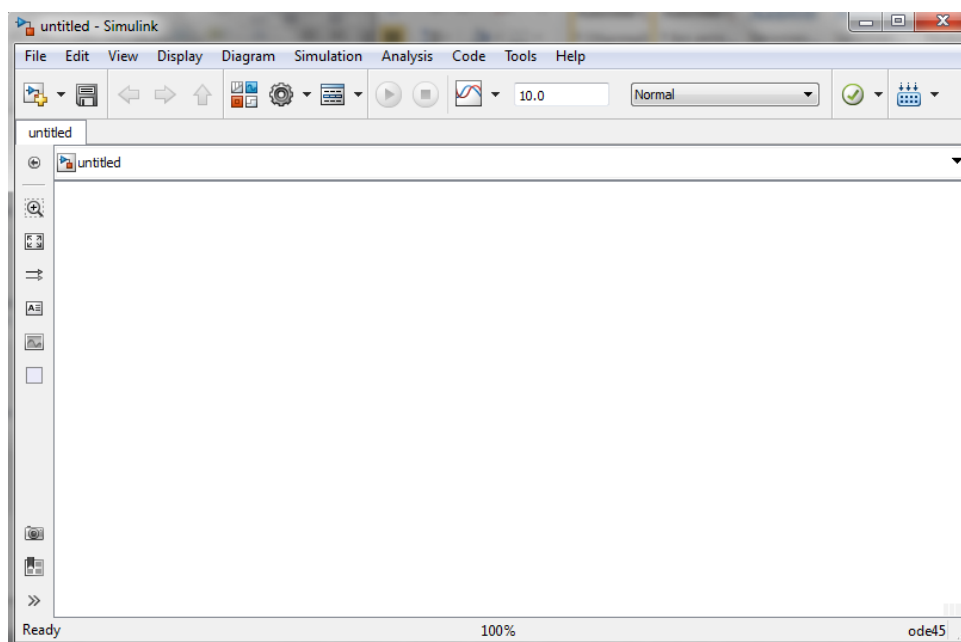


Рис. 3.45. Рабочее поле

Окно, представленное на рисунке 3.44, необходимо для поиска необходимых компонентов для реализации данной работы (кодеры, декодеры и т.д.)

Реализуем канал передачи такой же кодовой комбинации без преобразований.

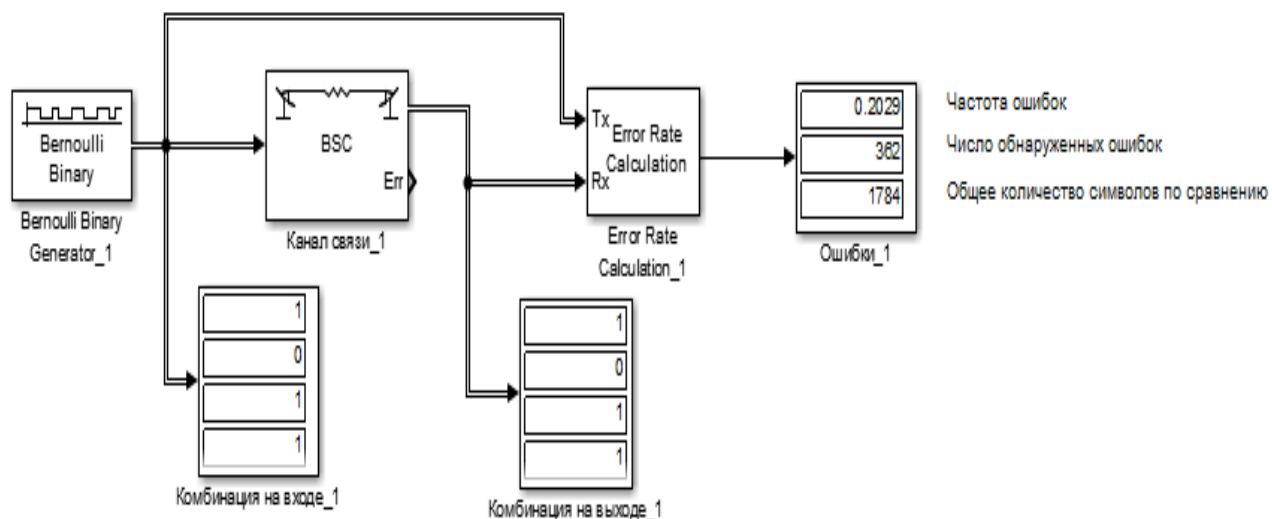


Рис. 3.46. Линия передачи без преобразований

В состав линии без преобразований входят:

1. BernoulliBinaryGenerator
2. Binary Symmetric Channel (каналпередачи)
3. Error RateCalculation (анализаторошибок)
4. Display

Характеристики блоков выставить следующие:

Bernoulli Binary Generator

Generate a Bernoulli random binary number.
To generate a vector output, specify the probability as a vector.

Parameters

Probability of a zero:

Initial seed:

Sample time:

Frame-based outputs

Samples per frame:

Output data type:

Рис. 3.47. Bernoulli Binary Generator

Binary Symmetric Channel (mask) (link)

Add binary errors to the input signal.

Parameters

Error probability:

Initial seed:

Output error vector

Output data type:

Рис. 3.48. Binary Symmetric Channel

Error Rate Calculation (mask) (link)

Compute the error rate of the received data by comparing it to a delayed version of the transmitted data. The block output is a three-element vector consisting of the error rate, followed by the number of errors detected and the total number of symbols compared. This vector can be sent to either the workspace or an output port.

The delays are specified in number of samples, regardless of whether the input is a scalar or a vector. The inputs to the 'Tx' and 'Rx' ports must be scalars or column vectors.

The 'Stop simulation' option stops the simulation upon detecting a target number of errors or a maximum number of symbols, whichever comes first.

Parameters

Receive delay:

Computation delay:

Computation mode:

Output data:

Рис. 3.49. ErrorRateCalculation (анализаторошибок)

Представим полученные результаты:

Входная последовательность:

1	0	1	1	0	1	1	1	0	0	0
1	1	0	1	1	0	0	0	0	1	1
0	1	0	1	0	0	0	1	1	1	0
1	0	0	0	1	0	0	0	0	1	0
1	1	0	0	1	0	1	1	0	0	1
0	1	0	0	0	1	1	0	1	1	1
0	1	1	0	1	1	1	0	1	1	0
0	0	1	1	1	1	1	0	0	0	0
0	0	0	1	1	1	1	1	0	1	0
0	1	0	0	1	0	1	0	1	0	0
1	1	1	1	1	0	0	1	0	0	0
1	0	1	0	1	0	1	1	1	1	1
1	0	0	1	1	0	1	1	1	1	0
0	0	0	0	0	0	0	1	1	1	0
1	1	1	1	0	0	1	1	0	1	1
1	1	1	0	0	0	0	1	0	0	1
0	1	0	1	1	0	1	1	0	1	1
1	0	0	0	1	0	1	1	1	0	1
0	1									

Комбинация на входе_1

Рис. 3.50. Входная последовательность

1	0	1	1	1	0	0	0	0	0	0
1	1	0	1	1	0	0	0	0	1	1
0	1	0	1	0	0	1	1	1	1	0
1	0	0	0	1	0	1	0	0	1	0
0	1	0	1	1	0	1	1	0	0	1
0	1	1	0	0	0	1	1	1	1	1
0	0	0	0	1	1	0	1	0	0	1
1	0	1	1	1	1	0	0	0	0	0
1	0	0	1	1	1	0	0	1	0	1
0	1	1	1	1	0	1	0	1	1	1
1	1	1	1	0	0	0	0	0	0	0
0	0	1	0	0	0	0	0	1	1	1
1	0	1	1	1	0	1	1	1	1	1
0	0	0	0	0	0	1	0	1	0	1
1	1	0	0	0	1	1	0	1	1	1
1	1	1	0	0	1	0	1	0	0	1
0	1	0	0	1	0	1	1	0	1	1
1	0	0	1	1	0	0	1	1	0	1
1	1									

Комбинация на выходе_1

Рис. 3.51. Выходная последовательность

0.2029	Частота ошибок
362	Число обнаруженных ошибок
1784	Общее количество символов по сравнению

Ошибки_1

Рис. 3.52. Счетчик ошибки

Анализируя рисунок выше, можно сделать вывод, что комбинация на входе совпадает с комбинацией на выходе, таким образом, передача осуществилась удачно. Что касается ошибок, то их частота равна 0,2029, число обнаруженных ошибок равно 362, общее

количество символов по сравнению равно 1784. В данной схеме осуществляется передача информации без преобразований. Если сравнивать этот результат с результатом с кодированием при вероятности ошибки 0,2, то они идентичны. При увеличении вероятности ошибки до 0,3 в декодированной последовательности были найдены ошибки.

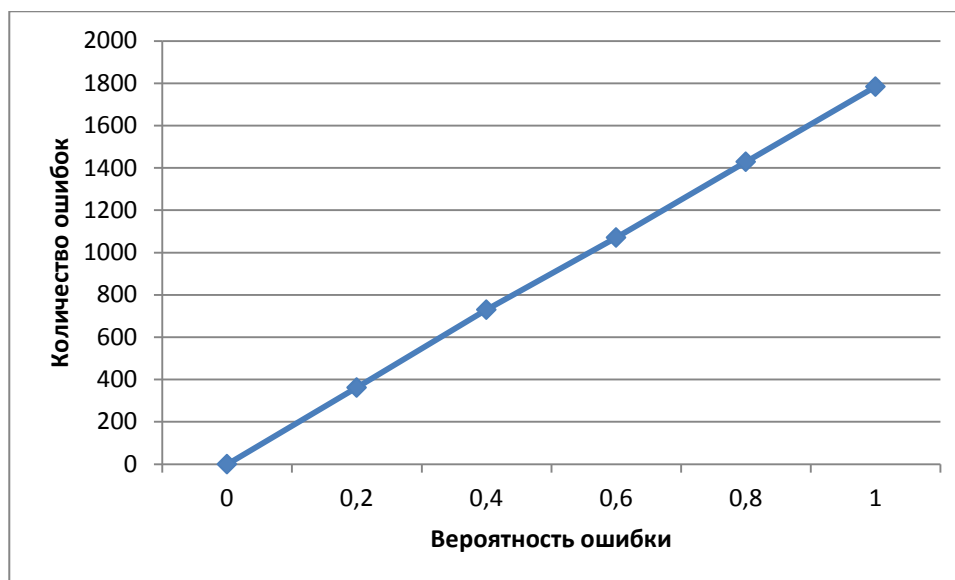


Рис. 3. 53. График зависимости числа обнаруженных ошибок от вероятности ошибки для кода (255,223)

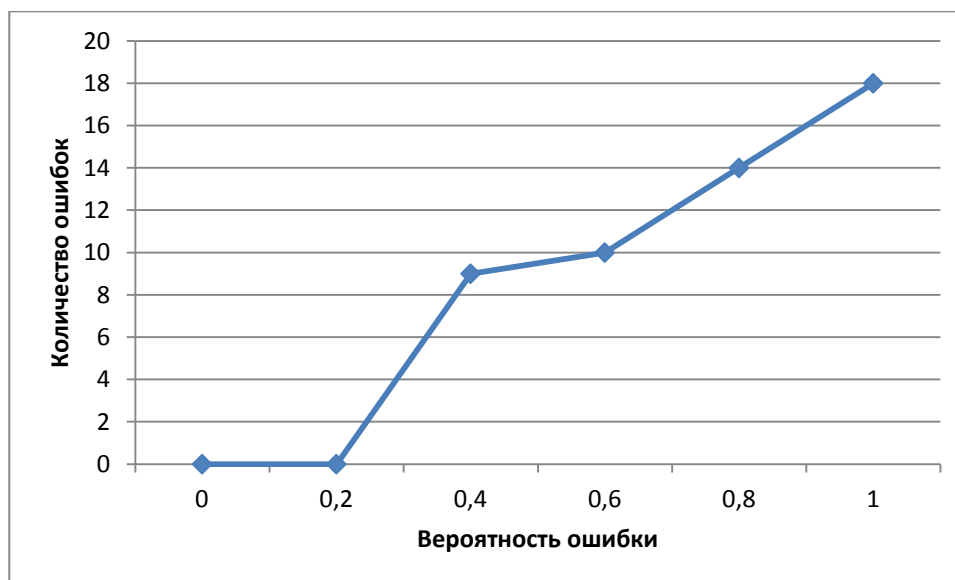


Рис. 3.54. График зависимости числа обнаруженных ошибок от вероятности ошибки для кода (7,3)

Были рассмотрены особенности кода Рида-Соломона. Построили графики зависимости числа обнаруженных ошибок от вероятности ошибки для кодов (255,223) и (7,3).

Изучена структура кодера и декодера.

1. Код Рида-Соломона — код с 8-битными символами и проверкой на чётность. В коде (255,223) 223 байта информационных символов, 32 байта проверки на четность.

2. Объем вычислительной мощности, необходимой для кодирования и декодирования кодов Рида-Соломона, зависит от числа символов четности. Большое значение t означает, что большее число ошибок может быть исправлено, но это потребует большей вычислительной мощности по сравнению с вариантом при меньшем t .

3. Код RS(255,223) может исправить до 16 ошибок в символах. В худшем случае, могут иметь место 16 битовых ошибок в разных символах (байтах). В лучшем случае, корректируются 16 полностью неверных байт, при этом исправляется $16 * 8 = 128$ битовых ошибок.

4. В результате анализа графиков рисунков 2.19-2.20 можно сделать вывод, что код с длинными последовательностями исправляет больше ошибок.

В разделе начало взяло аналитика выбора тех или иных параметров кодов, таких как, выбор размерности кодов, вероятность ошибки в канале и т.д. Сначала были выполнен предварительный анализ, позволившие выбрать и обосновать структурные схемы кодов по исходным данным технического задания.

Путем проведения компьютерной симуляции, была проверена достоверность расчетов. В программе Matlab 15 были собраны схемы:

- 1) Кода Хэмминга
- 2) Кода БЧХ
- 3) Кода Рида-Соломона

С помощью компьютерной симуляции была дана оценка ошибкам при передаче информации по каналу.

Также я познакомился с различным программным обеспечением, для построения различного вида схем.

Подводя итог своего индивидуального задания, можно сказать следующее, то, что я сделал, является основополагающим делом к дальнейшим, более трудным вещам. Индивидуальная работа была весьма увлекательна и полезна. С поставленными целями справился успешно.

Таблица 3.1. Результаты

	Вер-ть ошибки 0,2	Вер-ть ошибки 0,3
Код	Частота ошибок: 0,25	Частота ошибок: 0,4167
Хэмминга	Число обнаруженных ошибок:	Число обнаруженных ошибок: 5

(7,4)	3 Общее кол-во символов по сравнению: 12	Общее кол-во символов по сравнению: 12
Код BCH (15,7)	Частота ошибок: 0,2143 Число обнаруженных ошибок: 3 Общее кол-во символов по сравнению: 14	Частота ошибок: 0,4286 Число обнаруженных ошибок: 6 Общее кол-во символов по сравнению: 14
Код P-C (255,223)	Частота ошибок: 0,2029 Число обнаруженных ошибок: 362 Общее кол-во символов по сравнению: 1784	Частота ошибок: 0,31 Число обнаруженных ошибок: 553 Общее кол-во символов по сравнению: 1784

3.2. Циклические избыточные коды CRC (Cyclic redundancy check) [16]

Циклический избыточный код (CRC). Наиболее известными из методов обнаружения ошибок передачи данных являются [1]:

- *Посимвольный контроль чётности*, называемый также поперечным, подразумевает передачу с каждым байтом дополнительного бита, принимающего единичное значение по чётному или нечётному количеству единичных бит в контролируемом байте. Посимвольный контроль чётности прост как в программной, так и в аппаратной реализации, но его вряд ли можно назвать эффективным методом обнаружения ошибок, так как искажение более одного бита исходной последовательности резко снижает вероятность обнаружения ошибки передачи. Этот вид контроля обычно реализуется аппаратно в устройствах связи.

- *Поблочный контроль чётности*, называемый продольным. Схема данного контроля подразумевает, что для источника и приёмника информации заранее известно, какое число передаваемых символов будет рассматриваться ими как единый блок данных. В этой схеме контроля для каждой позиции разрядов в символах блока (поперёк блока) рассчитываются свои биты чётности, которые добавляются в виде обычного символа в конце блока. По сравнению с посимвольным контролем чётности, поблочный контроль чётности обладает большими возможностями по обнаружению и даже корректировке ошибок передачи, но всё равно ему не удаётся обнаруживать определённые типы ошибок.

- *Вычисление контрольных сумм.* В отличие от предыдущих методов, для метода контрольных сумм нет чёткого определения алгоритма. Каждый разработчик трактует понятие контрольной суммы по-своему. В простейшем виде контрольная сумма – это арифметическая сумма двоичных значений контролируемого блока символов. Но этот метод обладает практически теми же недостатками, что и предыдущие, самый главный из которых – нечувствительность контрольной суммы к чётному числу ошибок в одной колонке и самому порядку следования символов в блоке.

- *Контроль циклически избыточным кодом – CRC.* Это гораздо более мощный и широко используемый метод обнаружения ошибок передачи информации. Он обеспечивает обнаружение ошибок с высокой вероятностью. Кроме того, этот метод обладает рядом других полезных моментов, которые могут найти своё воплощение в практических задачах.

Циклический избыточный код (англ. Cyclic redundancy code, CRC) – алгоритм вычисления контрольной суммы, предназначенный для проверки целостности передаваемых данных. Алгоритм CRC обнаруживает все одиночные ошибки, двойные ошибки и ошибки в нечётном числе бит. Понятие циклических кодов достаточно широкое, однако на практике его обычно используют для обозначения только одной разновидности, использующей циклический контроль (проверку) избыточности.

Главная особенность значения CRC состоит в том, что оно однозначно идентифицирует исходную битовую последовательность и поэтому используется в различных протоколах связи, а также для проверки целостности блоков данных, передаваемых различными устройствами. Благодаря относительной простоте алгоритм вычисления CRC часто реализуется на аппаратном уровне.

При передаче пакетов по сетевому каналу могут возникнуть искажения исходной информации вследствие разных внешних воздействий: электрических наводок, плохих погодных условий и многих других. Сущность методики в том, что при хороших характеристиках контрольной суммы в подавляющем числе случаев ошибка в сообщении приведёт к изменению его контрольной суммы. Если исходная и вычисленная суммы не равны между собой, принимается решение о недостоверности принятых данных, и можно запросить повторную передачу пакета.

Основная идея алгоритма CRC состоит в представлении сообщения в виде огромного двоичного числа, делении его на другое фиксированное двоичное число и использовании остатка этого деления в качестве контрольной суммы [2]. Получив сообщение, приёмник может выполнить аналогичное действие и сравнить полученный остаток с «контрольной суммой».

На рисунках 2.1-2.2 изображено графическое представление кодирования и декодирования CRC-кода [3]:

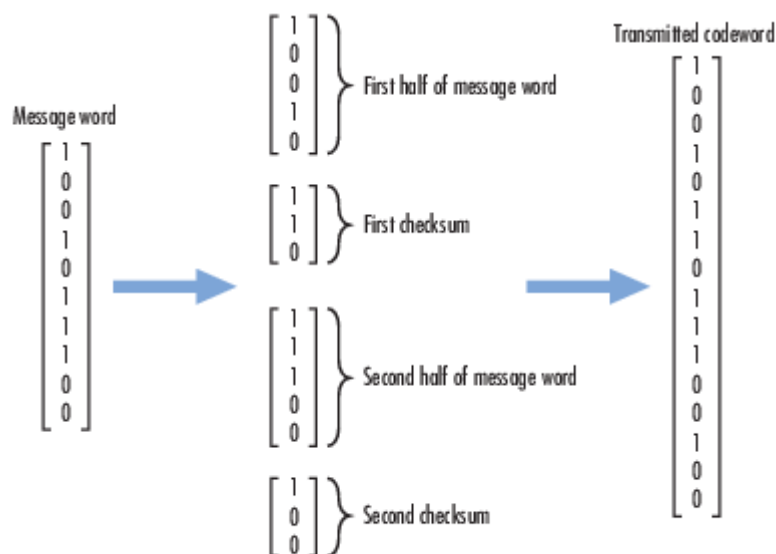


Рис. 3.55. Принцип работы кодера CRC

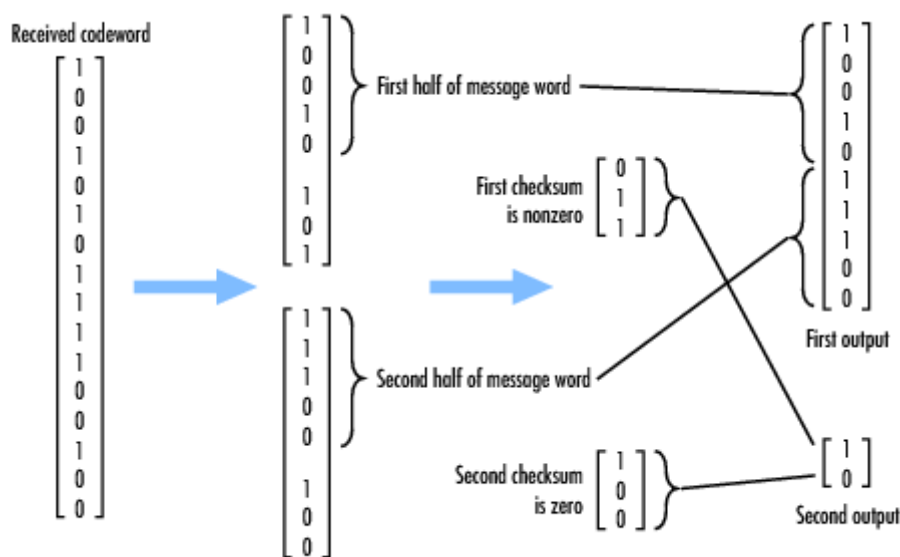


Рис. 3.56. Принцип работы декодера CRC

Степень CRC-полинома W называют позицию самого старшего единичного бита. Например, степень полинома 10011_2 равна 4.

Для вычисления CRC используют полиномиальную арифметику. Вместо представления делителя, делимого (сообщения), частного и остатка в виде положительных целых чисел, можно представить их в виде полиномов с двоичными коэффициентами или в виде строки бит, каждый из которых является коэффициентом полинома.

Например, десятичное число 23 в 16-ричной и 2-ичной системах будет иметь вид $23_{10}=17_{16}=10111_2$, что совпадает с полиномом: $1 \cdot x^4 + 0 \cdot x^3 + 1 \cdot x^2 + 1 \cdot x^1 + 1 \cdot x^0$ или упрощённо: $x^4 + x^2 + x^1 + 1$.

И сообщение, и делитель могут быть представлены в виде полиномов, с которыми можно выполнять любые арифметические действия без переносов.

Как правило, контрольная сумма добавляется к исходному сообщению и полученное расширенное сообщение передаётся через канал связи.

На другом конце канала приёмник может сделать одно из возможных действий (оба варианта совершенно равноправны):

1. Выделить текст полученного сообщения, вычислить для него контрольную сумму и сравнить её с переданной.
2. Вычислить контрольную сумму для всего переданного сообщения, и посмотреть, получится ли в результате нулевой остаток.

Поскольку исходное сообщение может быть очень большим (до нескольких Мбайтов) и так же из-за того, что для получения CRC используется CRC-арифметика, использовать обычную компьютерную операцию деления нельзя.

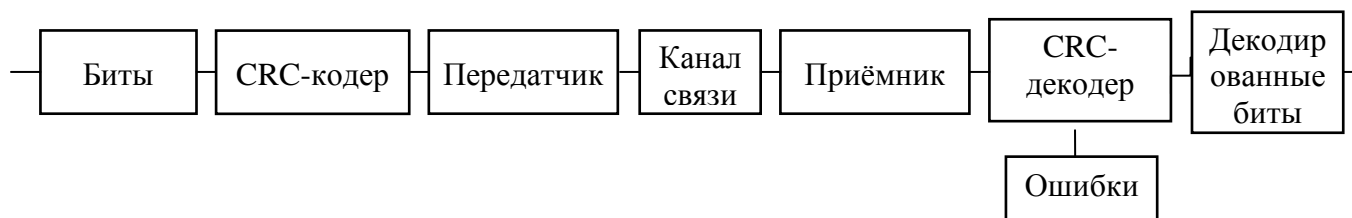


Рис. 3.57. Обобщенная структурная схема исследования CRC кодов

Самый популярный и рекомендуемый IEEE полином для CRC-32 используется в Ethernet, FDDI; также этот многочлен является генератором кода Хемминга. Использование другого полинома — CRC-32C — позволяет достичь такой же производительности при длине исходного сообщения от 58 бит до 131 кбит, а в некоторых диапазонах длины входного сообщения может быть даже выше — поэтому в наши дни он тоже пользуется популярностью. К примеру, стандарт ITU-T использует CRC-32C с целью обнаружения ошибок в полезной нагрузке.

Ниже в таблице перечислены наиболее распространённые многочлены — генераторы CRC [4]:

Таблица 3.2. Распространённые полиномы CRC кодов

Название	Полином
CRC-1	$x+1$ (используется в аппаратном контроле ошибок, также известен как бит чётности)
CRC-4-ITU	x^4+x+1
CRC-5-ITU	$x^5+x^4+x^2+1$
CRC-5-USB	x^5+x^2+1
CRC-6-ITU	x^6+x+1
CRC-7	x^7+x^3+1 (системы телекоммуникации, ITU-T G.707, ITU-T G.832, MMC, SD)
CRC-8	$x^8+x^7+x^6+x^4+x^2+1$
CRC-16-IBM	$x^{16}+x^{15}+x^2+1$ (Bisync, Modbus, USB, ANSI X3.28)
CRC-16-CCITT	$x^{16}+x^{12}+x^5+1$ (X.25, HDLC, XMODEM, Bluetooth, SD)
CRC-30	$x^{30}+x^{29}+x^{21}+x^{20}+x^{15}+x^{13}+x^{12}+x^{11}+x^8+x^7+x^6+x^2+x+1$ (CDMA)

Программная реализация виртуальных моделей кодирования

Описание реализации циклического избыточного кода (CRC)

Виртуальная модель передачи данных с обнаружением ошибок при помощи CRC-кода была реализована в среде Simulink Matlab. Модель демонстрирует работу CRC-кодера и декодера, позволяет исследовать обнаруживающую способность кода для разных генераторных полиномов.

На рисунке 3.58 приведена разработанная модель:

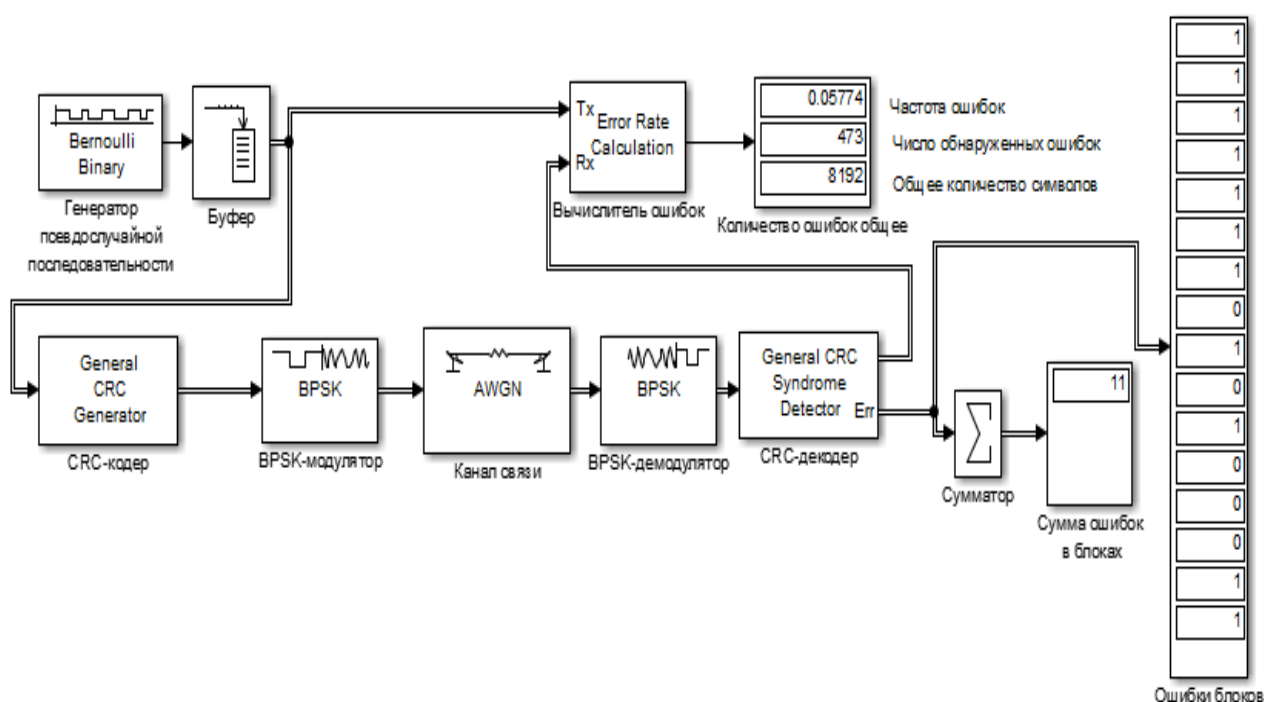


Рис. 3.58. Разработанная модель исследования CRC-кодов

В её основу положены следующие элементы, встроенные в библиотеку Simulink:

- Bernoulli Binary Generator;
- General CRC Generator;
- BPSK Modulator Baseband;
- AWGN Channel;
- BPSK Demodulator Baseband;
- General CRC Syndrome Detector;
- Error Rate Calculation;
- Buffer;
- Add;
- Display (Дисплей, отражающий ошибки).

Далее представлено описание основных блоков:

Bernoulli Binary Generator (генератор псевдослучайной последовательности) – генерирует случайную бинарную последовательность (рисунок 3.59).

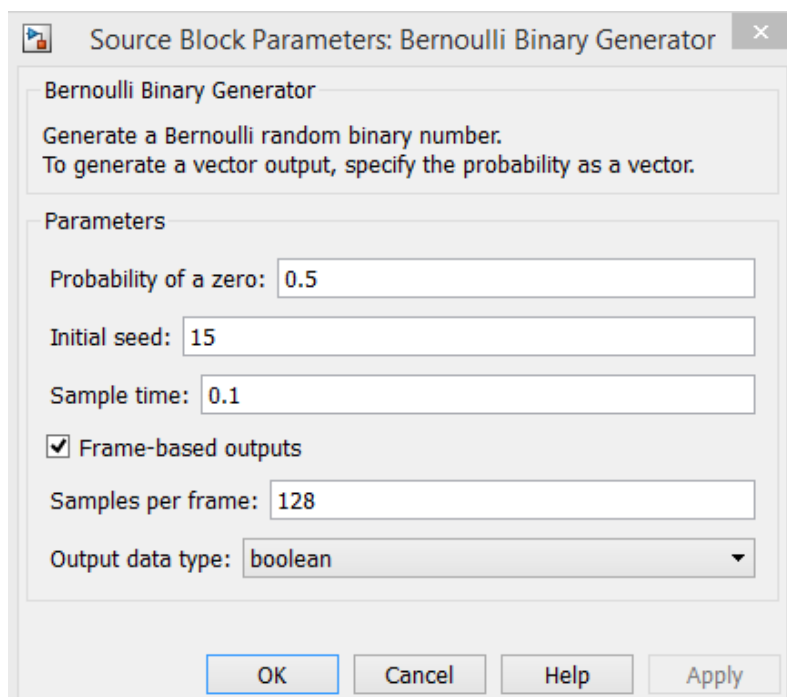


Рис. 3.59. Параметры блока «Bernoulli Binary Generator»

«Probability of a zero» - вероятность появления нуля;

«Initial seed» - начальное значение для генерации;

«Sample time» - длительность сэмпла;

«Samples per frame» - размер фрейма.

General CRC Generator (CRC-кодер) – циклический избыточный кодер (рисунок 3.3).

«Generator polynomial» - генераторный полином, может быть задан в 3 формах:

1) В обычной записи, например: $x^3 + x^2 + x + 1$.

2) в виде матрицы-строки с указанием степеней с ненулевыми коэффициентами, например: $[4 \ 1 \ 0] = x^4 + x + 1$.

3) в виде матрицы-строки с указанием нулевых и ненулевых коэффициентов, например: $[1 \ 1 \ 0 \ 1 \ 1] = x^4 + x^3 + x + 1$.

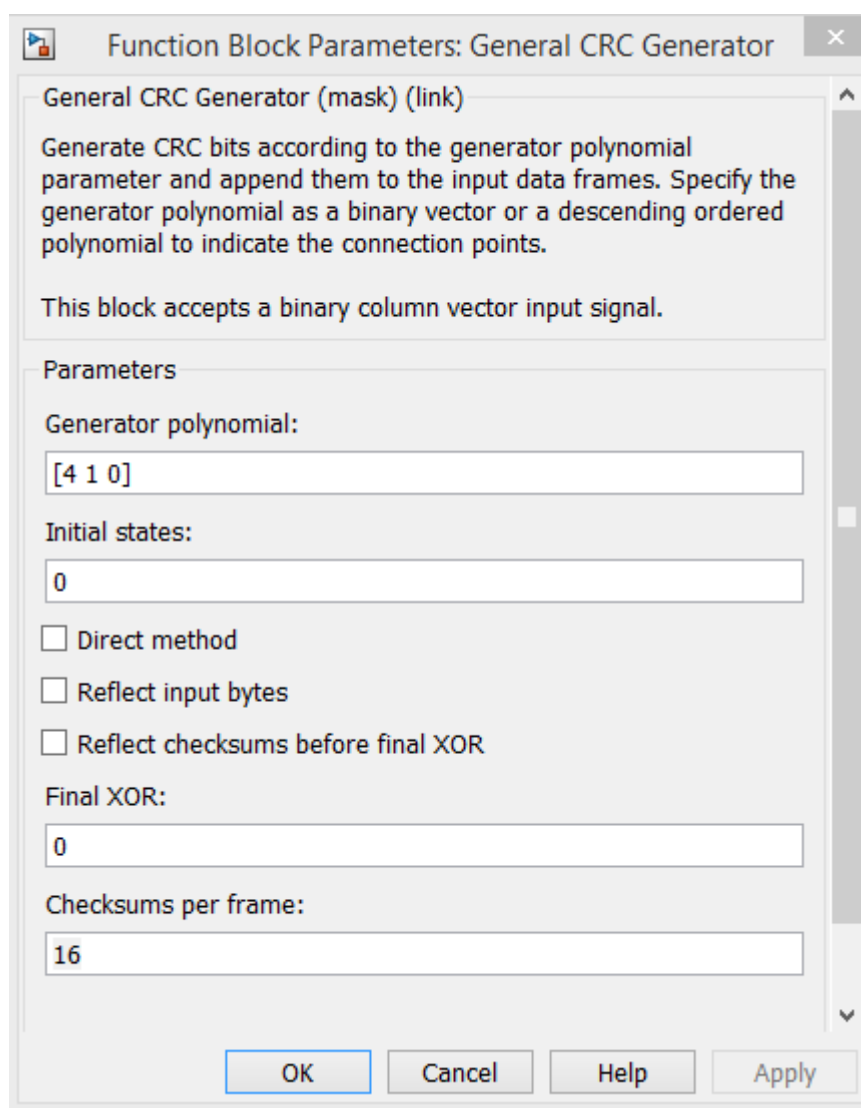


Рис. 3.60. Параметры блока «General CRC Generator»

«Initial states» - начальное состояние сдвиговых регистров.

«Direct method» - включение прямого метода вычисления CRC, иначе работает по табличному методу.

«Reflect input bytes» - инвертировать входной поток.

«Reflect checksums before final XOR» - инвертировать контрольные суммы перед конечной операцией XOR.

«Final XOR» - Выполнить операцию XOR в конце кодирования.

«Checksums per frame» - количество контрольных сумм во фрейме.

BPSK Modulator Baseband – BPSK модулятор.

BPSK Demodulator Baseband – BPSK демодулятор.

AWGN Channel (Канал связи) – добавляет «белый» гауссовский шум в канале (рисунок 3.61).

«SNR» - задаёт отношение сигнал/шум в канале.

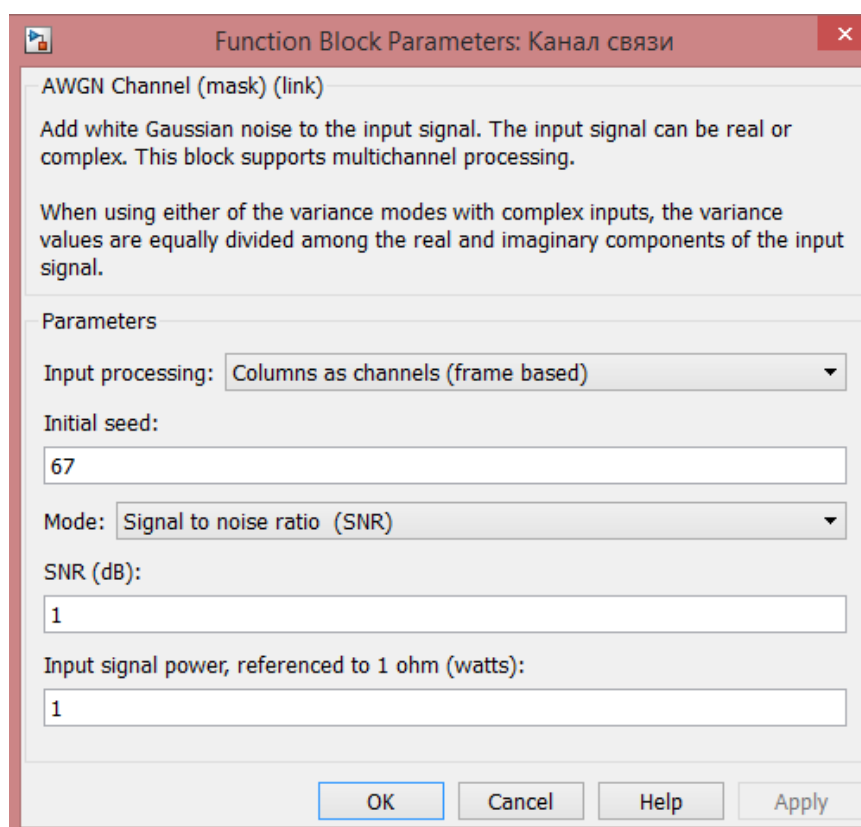


Рис. 3.61. Параметры блока «AWGN»

General CRC Syndrome Detector - циклический избыточный декодер. Все параметры декодера задаются аналогично параметрам блока «General CRC Generator» (рисунок 3.3).

Error Rate Calculation – вычислитель ошибок между переданной и принятой последовательностью.

Buffer – буфер. Переводит последовательность бит в один блок.

Add (сумматор) – суммирует ошибки от CRC-декодера.

Display - дисплей, отражающий ошибки.

Результаты моделирования

Исследование циклического избыточного кода

Модель циклического избыточного кода (срс), разработанная представленная на рисунке 3.58, позволяет исследовать обнаруживающую способность CRC кодов с различными полиномами.

Задаём одинаковый генераторный полином в блоки CRC-кодер и CRC-декодер (рисунок 3.62):

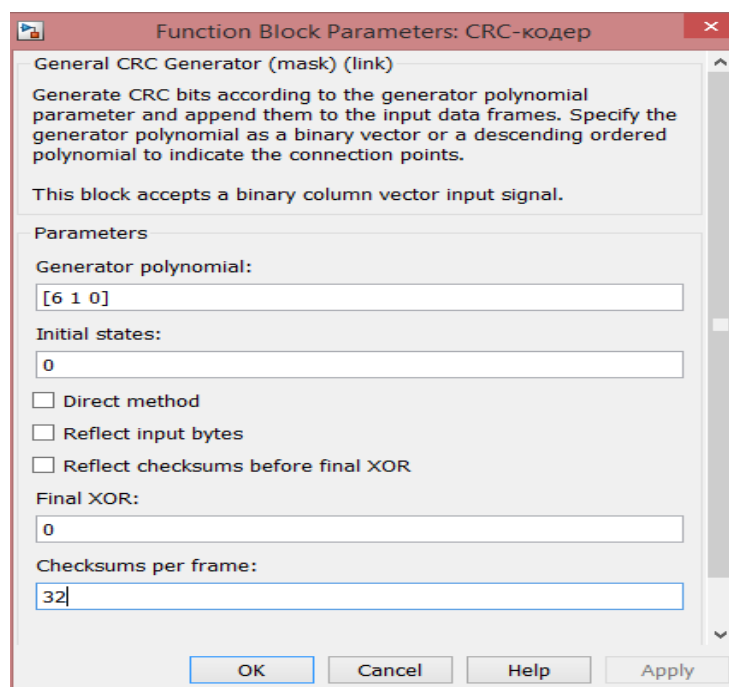


Рис. 3.62. Параметры блока CRC-кодер

Общее число передаваемых символов составляет 8192. Количество контрольных сумм изменяется от 2 до 8192, с увеличением каждого предыдущего значения в 2 раза (2, 4, 8, 16...8192).

Значение SNR в блоке «Канал связи» установлено в 1 дБ. Таким образом, битовая вероятность ошибки (BER) составит 0,05786.

На рисунке 4.2 представлен график зависимости числа обнаруженных ошибок от числа контрольных сумм для различных полиномов CRC-кода.

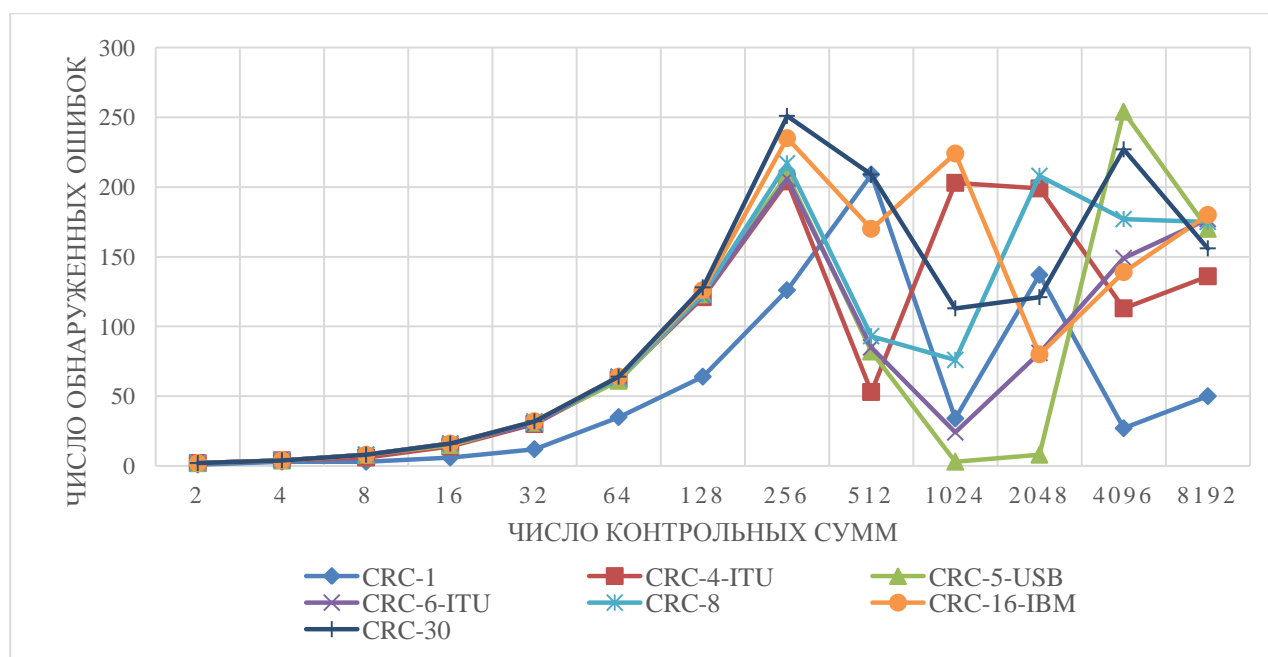


Рис. 3.63. График зависимости числа обнаруженных ошибок от числа контрольных сумм для различных полиномов CRC-кода

В данном разделе проведено исследование модели циклического избыточного кода (CRC).

Модель позволяет исследовать CRC-коды с возможностью задания любого генераторного полинома и изменении количества контрольных сумм во фрейме.

Получены следующие результаты и сделаны следующие выводы:

- 1) чем выше степень полинома, тем лучше его обнаруживающая способность;
- 2) для каждого полинома есть такое число контрольных сумм в блоке, при котором его обнаруживающая способность максимальна, причём у всех полиномов эти точки различны.

Однако, при выборе полинома CRC-кода также необходимо учитывать и другие факторы:

- 1) увеличение степени полинома приводит к усложнению реализации кодера и декодера;
- 2) чем выше частота вычисления контрольных сумм, т.е. чем больше контрольных сумм добавляется в блок данных, тем меньше пропускная способность канала;
- 3) CRC-коды используют для обнаружения ошибок, что означает наличие канала переспроса. При выборе между кодом CRC/каналом переспроса и помехоустойчивым кодированием, необходимо учитывать характеристики канала связи. При большом числе ошибок передача данных будет невозможна.
- 4) Выбор полинома зависит от размера передаваемого блока данных, чем больше блок – тем выше степень полинома необходимо подбирать. Таким образом, существует ограничение на размер блока данных, иначе в любом блоке на приёмном конце будет обнаруживаться ошибка.

3.3. Сверточные коды. Декодирование сверточных кодов

Современная теория кодов достаточно развита и содержит детальную классификацию. Все применяемые коды можно разбить на две большие группы: блочные, в которых кодирование и декодирование производится в пределах определенного участка кодовой последовательности- блока, и древовидные, в которых обработка символов производится непрерывно, без разделения на блоки. Часть кодов относится к разряду линейных, в которых кодовые последовательности представлены как элементы линейного векторного пространства. Можно применить также разбиение на коды, исправляющие независимые случайные ошибки, и коды, исправляющие пакетные ошибки.

В отличие от блочных, свёрточные коды обладают следующими преимуществами:

- сверточные коды позволяют производить кодирование и декодирование потоков данных непрерывно во времени;
- сверточные коды не нуждаются в блоковой синхронизации;

Используются два различных изображения регистров сдвига: с состыкованными впритык ячейками (рисунок 3.64, а) и с непосредственными последовательными связями между ячейками (рисунок 3.64, б), что дает возможность на схемах встраивать между соответствующими ячейками добавочные элементы (схемы суммирования по модулю 2).

Сумматор по модулю 2 осуществляет сложение поступающих на его входы символов 0 и 1. Правило сложения по модулю 2 следующее: сумма двоичных символов равна 0, если число единиц среди поступающих на входы символов четно, и равна 1, если это число нечетно.

Коммутатор осуществляет последовательное считывание поступающих на его входы (контакты) символов и устанавливает на выходе очередность посылки кодовых символов в канал связи.

По аналогии с блоковыми кодами, сверточные коды можно классифицировать на *систематические* и *несистематические*.

Систематическим сверточным кодом является такой код, для которого в выходной последовательности кодовых символов содержится без изменения породившая ее последовательность информационных символов. В противном случае сверточный код является несистематическим.

Сверточный код создается прохождением передаваемой информационной последовательности через линейный сдвиговый регистр с конечным числом состояний. В общем, регистр сдвига состоит из M (m -битовых) ячеек и линейного преобразователя, состоящего из n функциональных генераторов и выполняющего алгебраические функции, как показано на рисунке 3.65.

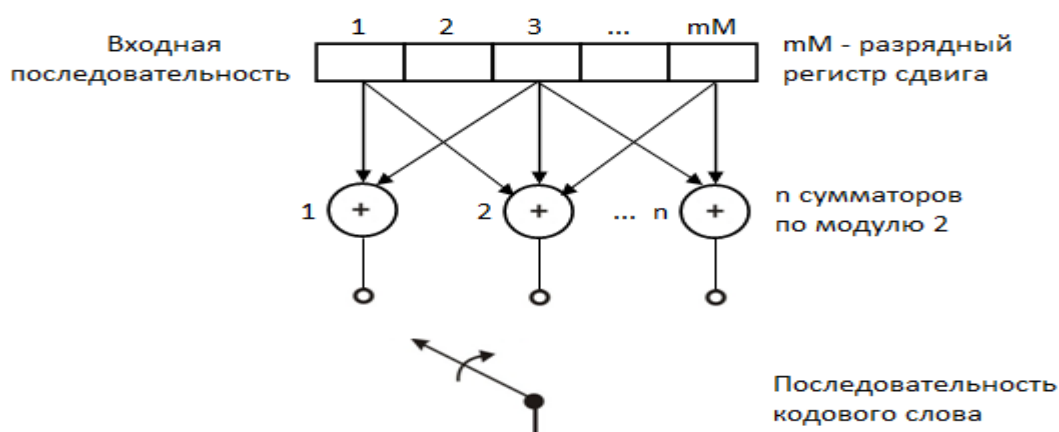


Рис. 3.65. Общий вид сверточного кодера

Входные данные к кодеру, которые считаются двоичными, продвигаются вдоль регистра сдвига по k бит за раз. Число выходных битов для каждой k -битовой входной последовательности равно n . Следовательно, кодовая скорость, определённая как $R = k/n$, согласуется с определением скорости блокового кода.

Представление сверточных кодов

Графическое представление

Рассмотрим свёрточный кодер со скоростью кода $1/2$, его графическое представление показано на рисунке 3.57. В этом кодере каждый раз, информационный бит поступает на вход регистров сдвига, а на выходе генерируется два бита.

В этом кодере каждый раз, информационный бит поступает на вход регистров сдвига, а на выходе генерируется два бита.

В качестве примера рассмотрена ситуация, когда на вход кодера подаётся некая последовательность информационных символов $V = 1\ 1\ 0\ 1\ 0\ 1$ на выходе имеем последовательность $U = 11\ 10\ 00\ 10\ 00\ 01$. Процесс образования выходных символов легко воспроизвести в уме глядя на рисунок.

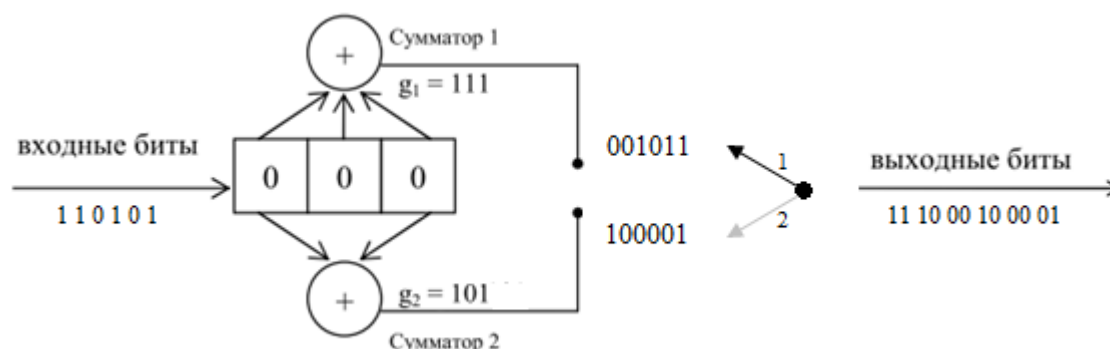


Рис. 3.66. Свёрточный кодер с $K=3$, $k=1$, $n=2$

Старшие разряды входных информационных бит для удобства располагаются справа и поступают в регистр сдвига первыми. При поступлении на вход регистра одного бита, на выходах сумматора 1 и сумматора 2 образуется по одному биту, которые в дальнейшем поочередно считываются коммутатором и образуют выходную последовательность.

Более подробно весь процесс кодирования информационного потока 00...101 приведен в таблице 3.3.

Таблица 3.3. Процесс кодирования информационного потока

№ такта	Входной информационный бит	Состояние регистра сдвига	Сумматор 1	Сумматор 2	Выходные кодовые комбинации
0	-	0000			-
1	1	1000	$1 \oplus 0 \oplus 0 = 1$	$1 \oplus 0 \oplus 0 = 1$	11

2	0	0100	$0 \oplus 0 \oplus 0 = 0$	$0 \oplus$ $1 \oplus 0 = 1$	01
3	1	1010	$1 \oplus 1 \oplus 0 = 0$	$1 \oplus$ $0 \oplus 0 = 1$	01
4	1	1101	$1 \oplus 0 \oplus 1 = 0$	$1 \oplus$ $1 \oplus 1 = 1$	01
5	0	0110	$0 \oplus 1 \oplus 0 = 1$	$0 \oplus$ $1 \oplus 0 = 1$	11
6	1	1011	$1 \oplus 1 \oplus 1 = 1$	$1 \oplus$ $0 \oplus 1 = 0$	10

В данной таблице видно, как меняется состояние регистра сдвига с приходом нового информационного бита, показан и процесс образования выходных кодовых комбинаций образуемых в результате считывания коммутатором символов с обоих сумматоров.

Полиномиальное представление

Иногда связи кодера описываются с помощью *полиномиального генератора*, аналогичного используемому для описания реализации обратной связи регистра сдвига циклических кодов. Сверточный кодер можно представить в виде набора из n полиномиальных генераторов, по одному для каждого из n сумматоров по модулю 2. Каждый полином имеет порядок $K - 1$ или меньше и описывает связь кодирующего регистра сдвига с соответствующим сумматором по модулю 2, почти так же как и вектор связи. Коэффициенты возле каждого слагаемого полинома порядка $(K - 1)$ равны либо 1, либо 0, в зависимости от того, имеется ли связь между регистром сдвига и сумматором по модулю 2. [7].

Для кодера на рисунке 3.57, можно записать полиномиальный генератор $g_1(X)$ для верхних связей и $g_2(X)$ - для нижних.

$$g_1(X) = 1 + X + X^2, \quad g_2(X) = 1 + X^2$$

Здесь слагаемое самого нижнего порядка в полиноме соответствует входному разряду регистра. Выходная последовательность находится следующим образом: $U(X) = m(X)g_1(X)$ чередуется с $m_1(X)g_2(X)$.

Прежде всего, необходимо выразить вектор некоего сообщения $m = 1 \ 0 \ 1$ в виде полинома, т.е. $m(X) = 1 + X^2$. Для очистки регистра предполагается использование нулей, следующих за битами сообщения. Тогда выходящий полином $U(X)$, или выходящая

последовательность U кодера (рисунок 3.57) для входного сообщения m может быть найдена следующим образом:

$$m(X)g_1(X) = (1+X^2)(1+X+X^2) = 1 + X + X^3 + X^4$$

$$\underline{m(X)g_2(X) = (1+X^2)(1+X^2) = 1+X^4}$$

$$m(X)g_1(X) = 1+X+0X^2+X^3+X^4$$

$$m(X)g_2(X) = 1+0X+0X^2+0X^3+X^4$$

$$U(X) = (1,1) + (1,0)X + (0,0)X^2 + (1,0)X^3 + (1,1)X^4$$

$$U = 11\ 10\ 00\ 10\ 11$$

В этом примере кодер был представлен в виде *полиномиальных генераторов*, с помощью которых описываются циклические коды [7]. *Полиномиальные генераторы кодера изображенного на рисунке 2.57 так же можно представить в виде векторов:*

$$g_1 = [111] \text{ и } g_2 = [101].$$

Представление в виде кодовых деревьев, решеток и диаграмм состояний

Помимо графического и полиномиального представления имеются три альтернативных метода, которые часто используются для описания свёрточного кода. Это древовидная диаграмма, решетчатая диаграмма и диаграмма состояний. Для примера, древовидная диаграмма для свёрточного кодера, показанного на рисунке 3.57 иллюстрируется на рисунке 3.58.

Рассмотрим свёрточный код со скоростью $1/2$, $K=3$, описанный ранее и показанный на рисунке 3.57. Первый входной бит может быть 0 или 1. Соответствующие выходные биты – 00 или 11. Когда следующий бит входит в кодер, первый бит передвигается в следующую ячейку. Соответствующие выходные биты зависят от бита, переместившегося во вторую ячейку и нового входного бита. Следовательно, древовидная диаграмма для этого кода, показанная на рисунке 3.58, имеет две ветви на узел, соответствующие двум возможным входным символам. Жирной линией на рисунке 3.58 обозначен путь, по которому бы шел кодер, если бы входной комбинацией была последовательность 101.

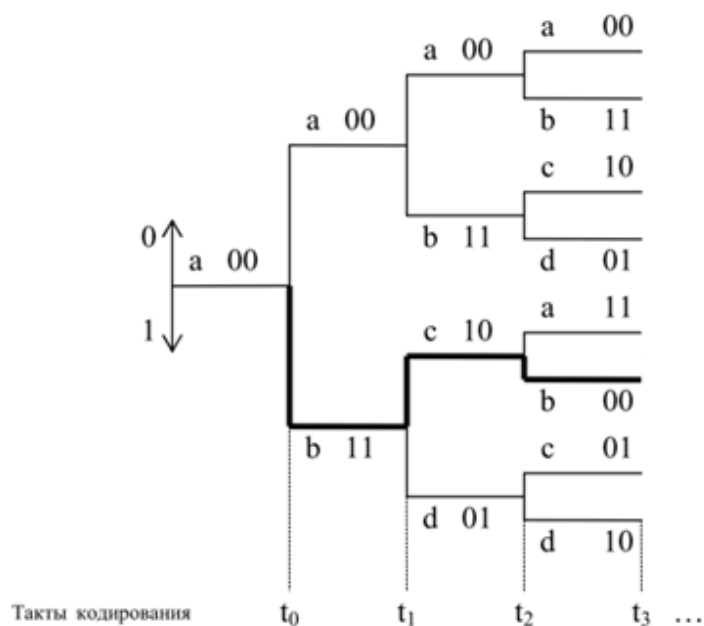


Рис. 3.67. Древоподобная диаграмма для сверточного кода, имеющего параметры $K=3$, $k=1$, $n=2$

Поскольку кодовое ограничение по входу кодера $K=3$, а значит дерево начнёт повторяться после третьего шага. Как показано на рисунке 3.58, все ветви, исходящие из узла, обозначенного a (состояния a) дают идентичные выходы. Путем слияния узлов, имеющих одинаковое название, мы получаем решётку, показанную на рисунке 3.68. На нём жирной линией обозначен путь по которому бы шел кодер, если бы входной комбинацией была последовательность 1100, в таком случае на выходе имела бы комбинация 11 01 01 11.

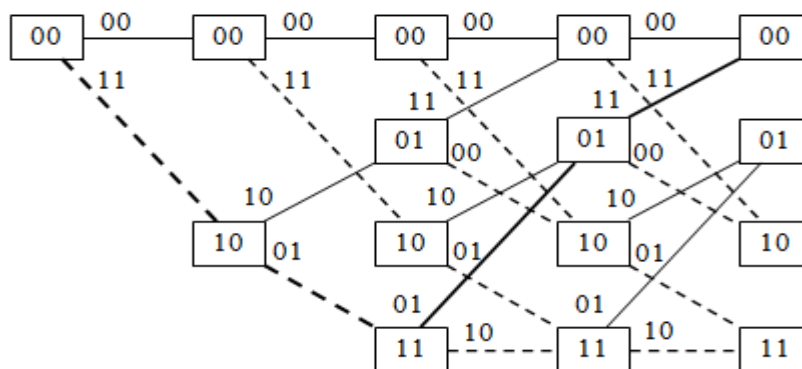


Рис. 3.68. Решётчатая диаграмма для сверточного кода, имеющего параметры $K=3$, $k=1$, $n=2$

Поскольку выход кодера определяется входом и состоянием кодера, ещё более компактной, чем решётка, является диаграмма состояний. Диаграмма состояний — это граф возможных состояний кодера и возможных переходов из одного состояния в другое. Итоговая диаграмма состояний для этого кодера показана на рисунке 3.60.

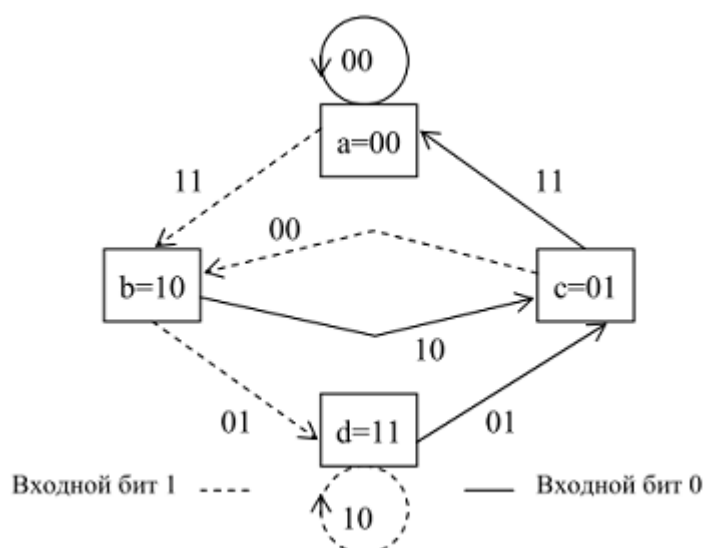


Рис. 3.69. Диаграмма состояний для свёрточного кода, имеющего параметры

$$K=3, k=1, n=2$$

Для обобщения стоит отметить, что свёрточный код со скоростью k/n и кодовым ограничением K характеризуется 2^k ветвями, уходящими от каждого узла на древовидной диаграмме. Решетка и диаграмма состояний имеют (каждая из них) $2^{k(K-1)}$ возможных состояний. Имеются 2^k ветвей, входящих в каждое состояние, и 2^k ветвей, покидающих каждое состояние (для решётки и дерева это верно после наступления установившегося режима) [1, 10].

Методы декодирования свёрточных кодов

Метод порогового декодирования

При пороговом декодировании свёрточных кодов вычисляются синдромы (признаки места ошибочных символов), затем эти синдромы или последовательности, полученные посредством линейного преобразования синдромов, подаются на входы порогового элемента, где путем “голосования” (мажоритарный метод) и сравнения его результатов с порогом выносится решение о значении декодируемого символа. Основное достоинство этого метода декодирования – простота реализации. Однако он не полностью реализует потенциальные корректирующие способности свёрточного кода. Кроме того, не все свёрточные коды могут быть декодированы этим методом. Пороговое декодирование, как правило, применяется для систематических кодов. [1]. Общая схема декодера для свёрточного кода ($R = 1/2$) представлена на рисунке 3.70.



Рис. 3.70. Общая схема декодера для сверточного кода

Декодер содержит аналог кодера, в котором по принимаемым информационным символам в сдвигающем регистре формируется копия проверочной последовательности. С этой целью синхронизатор декодера с помощью ключей 1 и 2 “расфасовывает” входную последовательность символов на 2 потока – информационный и проверочный, синхронизатор управляет работой всего декодера.

В формирователе синдрома (сумматоре по модулю 2) образуется последовательность синдромов S , которая поступает на вход синдромного регистра. В отсутствие в канале ошибок последовательности на входах формирователя синдрома всегда совпадают, и синдромная последовательность состоит из одних нулей. Различным наборам ошибок соответствуют определенные конфигурации синдромных последовательностей, в которых на определенных позициях появляются единичные символы. Закон формирования проверочных символов выбирается таким образом, чтобы по структуре синдромной последовательности можно было определить искаженные символы.

Логическая схема определяет по синдрому правильность записанного в информационном регистре блока информационных символов. Если имеется комбинация ошибок, которая может быть исправлена, то логическая схема исправляет ошибки в этом блоке путем подачи единичных символов на выходную схему суммирования по модулю 2 в моменты выхода из информационного регистра искаженных символов.

Ошибки, исправляемые в очередном информационном блоке, могут влиять на символы синдромов, соответствующих последующим блокам, поскольку сверточные коды непрерывны. Для того чтобы декодер смог полностью реализовать свои корректирующие возможности, следует исключить влияние этих ошибок. Это может быть достигнуто за счет обратной связи, которая на схеме (рисунок 3.62) представлена пунктирной линией.

Обратная связь преобразует синдромный регистр прямого действия в нелинейный регистр сдвига с обратной связью. Это может привести к явлению, называемому

размножением ошибок. Неисправимые ошибки в канале могут вызвать переход синдромного регистра в такое состояние, что и при отсутствии аддитивных ошибок в канале декодер будет продолжать всегда декодировать неправильно. Причина этого состоит в том, что выход нелинейного регистра сдвига с обратной связью, когда на его вход поступает нулевая последовательность, а начальное состояние – ненулевое, может быть периодическим [6].

Последовательный алгоритм декодирования

При рассмотрении алгоритмов последовательного декодирования удобно представлять сверточный код в виде кодового дерева. Как уже отмечалось ранее, исходному нулевому состоянию сдвигающего регистра кодера соответствует начальный узел дерева. Если входной информационный символ, поступающий в регистр равен 1, то ему приписывается линия (ребро дерева), идущая, как принято на этом рисунке, вниз, а если информационный символ равен 0, – то вверх. Тем самым получаем два новых узла, соответствующие следующему такту работы кодера, для каждого из которых дерево строится далее аналогичным образом, и так далее. Над каждым ребром дерева записываются кодовые символы, получаемые при этом на выходе кодера. Совокупность нескольких последовательных ребер, соединяющих какие-либо два узла, составляет ветвь дерева.

Каждая последовательность кодируемых информационных символов порождает определенный путь по кодовому дереву. Очевидно, задача декодера заключается в отыскании истинного (правильного) пути, то есть того пути, который в действительности был порожден кодером [1].

Таким образом, при алгоритмах последовательного декодирования декодер определяет наиболее правдоподобный путь по решетке, что позволяет исключить из анализа большую часть остальных путей, имеющих меньшее правдоподобие.

Для этого, сначала необходимо передвигаться поочередно вдоль каждого пути кодовой решетки, сравнивая принятую последовательность со значением, соответствующим пути, по которому происходит движение. Если при этом удастся обнаружить некоторый путь, значение которого, совпадает с принятой последовательностью, то естественно считать, что эта последовательность и передавалась.

Если быть точнее, происходит сравнение принятой комбинации с комбинациями возможными для данного шага декодирования. Путь считается наиболее правдоподобным, если метрика между принятой последовательностью и последовательностью соответствующей данному переходу минимальна.

Резюмируем: на каждом шаге декодирования мы имеем два исходящих из узла пути. Сначала мы вычисляем расстояние Хемминга между принятой комбинацией и комбинацией соответствующей одному исходящему пути, затем вычисляем расстояние Хемминга между

принятой комбинацией и комбинацией соответствующей второму пути. После чего выбираем тот путь, значение расстояния Хемминга которого меньше. Данный процесс показан на рисунке 3.71. В данном случае переход вниз по решетке будет считаться наиболее правдоподобным, так как расстояние Хемминга между значением соответствующему этому переходу и принятой комбинацией является наименьшим.

Принятая последовательность: 11

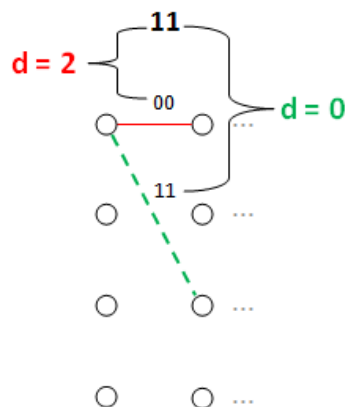


Рис. 3.71. Выбор пути с наименьшим расстоянием Хэмминга

Декодер выполняет операцию выбора пути с наименьшей метрикой до тех пор пока не столкнётся с неоднозначностью, при которой принять решение о том какой путь является кратчайшим не представляется возможным из-за того что расстояния Хемминга между принятой последовательностью и значениями двух возможных и путей одинаковы. В этой ситуации декодер вынужден проверить оба возможных пути, и по следующим метрикам принять решение касаясь неопределённости.

В нашем же случае, если расстояния Хемминга для обоих путей одинаковы, то прибегаем к жесткому методу принятия решений и пользуемся следующим правилом:

- Если мы находимся в узле 00 или 11, и пришла комбинация 00 или 01 – считаем что передавался ноль, 10 или 11 – единица.
- Если мы находимся в узле 01 или 10, и пришла комбинация 00 или 01 – считаем что передавалась единица, 10 или 11 – ноль.

Аналогичным образом декодер продолжает углубляться в решетку и принимать решения, касающиеся информационных битов, устраняя все пути, кроме одного.

На рисунке 3.72 изображена ситуация когда на вход декодера поступает последовательность, в которой биты выделенные цветом заведомо искажены так, как это бывает в реальных каналах связи.

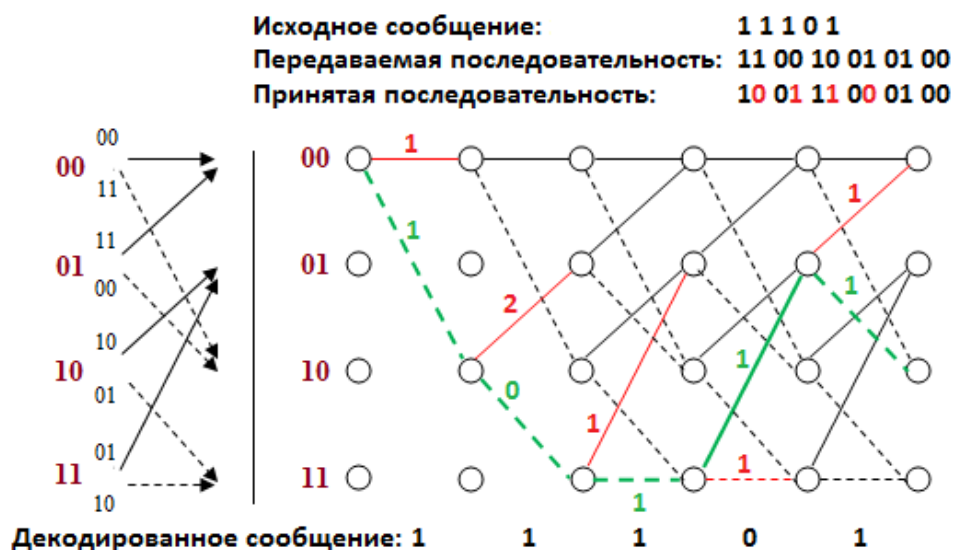


Рис. 3.72. Решетчатая диаграмма последовательного декодера

Сравнение алгоритма Витерби с последовательным алгоритмом

Принципиальное отличие алгоритма Витерби от последовательного алгоритма декодирования заключается в том, что алгоритм Витерби принимает решение по суммарным метрикам путей, причём в узлах, где эти пути сходятся. Этот шаг в процессе декодирования показан на рисунке 3.65. (метрики обозначены данным образом для наглядности).

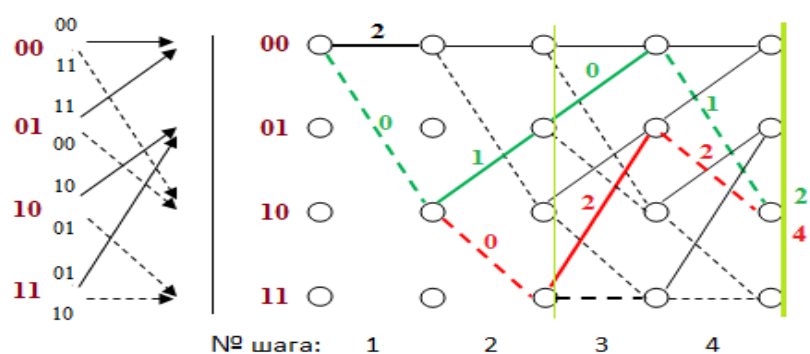


Рис. 3.73. Решетчатая диаграмма декодера Витерби

Таким образом, если бы в ситуации, представленной на рисунке 3.64, мы принимали решение на основе последовательного алгоритма, то на первом и втором шаге декодирования мы приняли бы решение идти вниз по решетке, выбирая кратчайшие метрики Хемминга. Однако в таком случае мы бы не учли метрики последующих ветвей нашего, на первый взгляд кратчайшего пути, которые на шаге три и четыре сильно возрастают, а значит, выбрали бы не самый короткий путь при декодировании. Поэтому, алгоритм Витерби является оптимальным. Однако он сложнее последовательного как в понимании, так и в реализации. В качестве метрики выступает расстояние Хемминга.

Программная реализация сверточного кодера

Подготовка данных к кодированию

Сначала, символы, введенные с клавиатуры ассоциируются со своими десятичными значениями в таблице ASCII. Затем, ASCII-код символа переводится в восьмеричную систему исчисления.

Ранее был объявлен массив «Bit_Array()», который имеет 8 ячеек, в которых содержится двоичное представление чисел от 0 до 7. Так как восьмеричная система имеет всего восемь цифр (0-7) то перебирается массив в цикле от 0 до 7, и находится равенство между индексом массива (счетчиком цикла) и восьмеричной цифрой ASCII-кода, после чего вместо нее подставляется двоичное представление из ячейки массива.

К примеру, цифра 0, будет иметь десятичный ASCII код - 48, в восьмеричный - 60. Что в двоичной системе исчисления будет 110000. Если ASCII-код - двузначное число, впереди дописывается 00, чтобы двоичные числа имели одинаковую длину. Формируется строка бит сообщения 00110000.

После того как символы исходного текста ассоциированы с кодами таблицы кодов ASCII и переведены в двоичный вид, осуществляется свёрточное кодирование.

Функция свёрточного кодирования

Разрабатываемый свёрточный кодер является несистематическим, без обратной связи и изображен на рисунке 3.74. Функциональную схему кодера (см. Приложении Б).

Кодер называется систематическим, если в каждом кодовом подблоке длины s имеется b символов, которые совпадают с символами соответствующего информационного подблока. В противном случае кодер - несистематический.

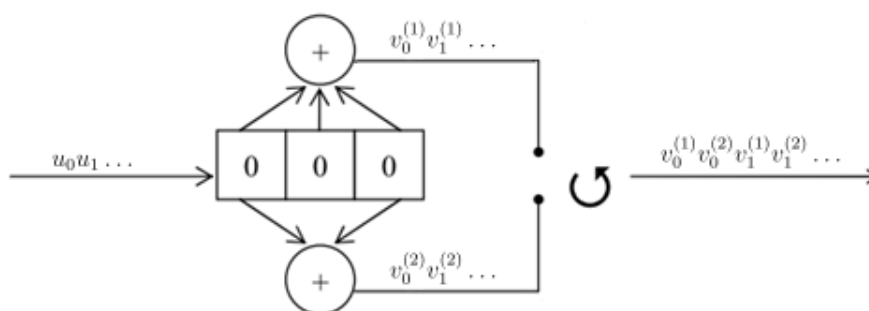


Рис. 3.74. Двоичный сверточный кодер скорости $R = 1/2$.

Бесконечная последовательность информационных символов $u = u_0 u_1 \dots$ поступает в регистр сдвига (в нашем примере памяти $m = 3$, то есть с двумя элементами задержки). Кодер генерирует две выходные последовательности $v^{(1)} = v_0^{(1)} v_1^{(1)} \dots$ и $v^{(2)} = v_0^{(2)} v_1^{(2)} \dots$, которые поступают в коммутатор, формирующий выходную (кодую) последовательность $v = v_0^{(1)} v_0^{(2)} v_1^{(1)} v_1^{(2)} \dots$, передаваемую по каналу. Поскольку на каждый

входящий в кодер информационный символ приходится два кодовых символа, кодовая скорость равна $R = 1/2$. Если одна из кодовых последовательностей $v^{(1)}$ или $v^{(2)}$ совпадает с информационной последовательностью u , то, как отмечалось ранее, кодер называется систематическим. В противном случае кодер называется несистематическим [9].

Генераторы свёрточного кодера изображенного на рисунке 3.75. можно описать при помощи полиномов:

$$G_1 = 1+X+X^2$$

$$G_2 = 1+X^2$$

Таким образом, порождающая матрица этого свертчного кодера, в терминах оператора задержки имеет вид:

$$G(X) = (1+X+X^2 \ 1+X^2).$$

Как уже отмечалось ранее, кодер – это автомат. В данном случае имеется классический алгоритм автомата Мили — конечный автомат, выходная последовательность которого (в отличие от автомата Мура) зависит от состояния автомата и входных сигналов. Это означает, что в графе состояний каждому ребру соответствует некоторое значение (выходной символ). В вершины графа автомата Мили записываются выходящие сигналы, а дугам графа приписывают условие перехода из одного состояния в другое, а также входящие сигналы. Диаграмма состояний свёрточного кодера представленного на рисунке 3.66 изображена на рисунке 3.75.

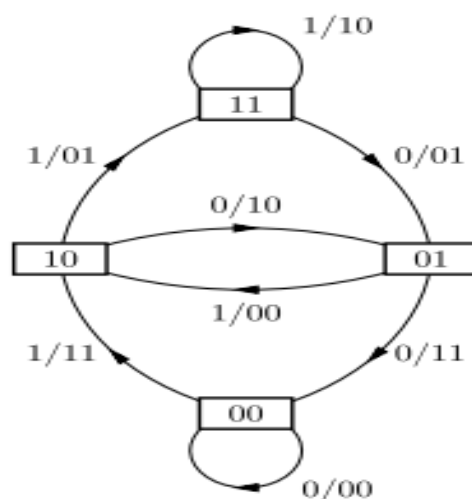


Рис. 3.75. Диаграмма состояний свертчного кодера скорости $R = 1/2$

Свёрточное кодирование кодером со скоростью $R = 1/2$ в программном комплексе выполняет функция Fun_BitCoder. После объявления массива и его целочисленных переменных обнуляются регистры сдвига, описывается цикл, осуществляющий побитное чтение информационных символов, их сдвиг в регистрах и суммирование по модулю два

согласно установленным правилам. Элемент исходного кода программы, отвечающий за реализацию свёрточного кодирования, представлен на рисунке 3.76.

```

    R1 = 0
    R2 = 0
    For i = 0 To length_array 'цикл обработки сообщения автоматом Мили
    X = S_mess_chr(i).ToString
    Y1 = X Xor R1 'сумма по модулю 2(исключающее или)
    Y2 = X Xor R2 '
    R2 = R1 'сдвиг битов в регистрах
    R1 = X '
    Y1 = Y1 Xor R2 'сумма по модулю 2(исключающее или)
    s_res = s_res + Y1.ToString + Y2.ToString 'результат кодирования
    Next
    Return s_res
End Function

```

Рис. 3.76. Элемент программного кода функции свёрточного кодирования

В роли сумматоров выступают переменные Y1 и Y2, а в роли ячеек регистра переменные R1 и R2. Коммутатором является переменная s_res. Для осуществления поочередной коммутации бит с выходов кодера символы переводятся в строковый тип.

Функция имитации канала связи

По мере прохождения по каналу связи сигнал подвергается воздействию помех, именно этот процесс имитирует данная функция. Помехи - случайные воздействия, искажающие передаваемый сигнал. Воздействие помехи на сигнал может быть двояким. Если помеха $\zeta(t)$ складывается с сигналом $s(t)$ и на вход приёмника действует их сумма $x(t) = s(t) + \zeta(t)$ такую помеху называют аддитивной. Если результирующий сигнал равен произведению помехи и передаваемого сигнала $x(t) = s(t) \cdot \zeta(t)$, то помеху называют мультипликативной.

В канале с мультипликативной помехой наблюдаются всплески искажений, причинами которых могут быть, например, коммутационные помехи, быстрые замирания радиосигнала. Выражается мультипликативная помеха в изменении характеристик линии связи (сопротивление линии связи (ЛС), частота среза ЛС, нелинейность характеристик ЛС).

Чтобы приблизить модель симметричного канала к действительности, вводится понятие пакета ошибок.

Пакетом ошибок длины L называется вектор ошибок E, все ненулевые компоненты которого расположены на отрезке из подряд следующих позиций, причем в начале и конце отрезка расположены ненулевые компоненты.

Аддитивная помеха не зависит от ЛС и определяется внешними воздействиями на среду передачи сигналов. В качестве модели аддитивной помехи обычно принимают белый шум, то есть стационарный гауссовский случайный процесс, имеющий нулевое математическое

ожидание и равномерный и бесконечно широкий спектр плотности мощности. При воздействии помех на передаваемый сигнал, имеющих аддитивный характер, $L=1$ имеем пакет длиной в один символ, то есть однократную ошибку.

В разрабатываемом программном комплексе имеется возможность пронаблюдать эффективность работы последовательного декодера свёрточных кодов как при наличии однократных ошибок, так и пачек длиной $L > 1$.

На рисунке 3.77 приведена блок-схема программно реализованного алгоритма имитации канала связи с белым шумом и канала с многолучевым распространением.

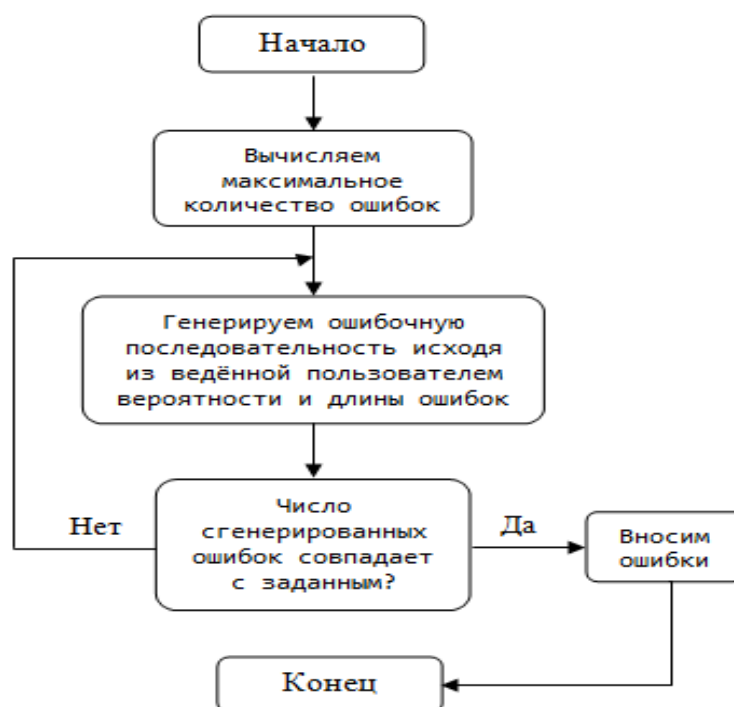


Рис. 3.77. Блок-схема для алгоритма имитации канала связи

За имитацию канала связи отвечает функция «Fun_SimulatorChannelCom». В ней, первым делом, определяется максимальное количество генерируемых ошибок:

$$\text{max_mistake_dou} = \text{length_array} * (\text{CInt}(\text{probability_mistake}) / 100),$$

где: max_mistake_dou – максимальное количество ошибок

length_array – длина сообщения

$\text{probability_mistake}$ – вероятность ошибки

Например, если длина сообщения равна 80 бит, а битовая вероятность ошибки – 0.1 то генератор сгенерирует от 0 до 8 ошибок.

Далее, активируется обычный генератор случайных чисел в диапазоне от 1 до 100:
`Dim value As Integer = CInt(Int((100 * Rnd()) + 1))`

Функция CInt приводит выражение в тип Integer и округляет его до ближайшего целого. После чего число сгенерированных ошибок сверяется с заданным числом ошибок,

если ошибок меньше чем задано, цикл продолжается. Проверяется также заданная вероятность с сгенерированным числом, если оно меньше или равно, то в сообщение вносится ошибка. Внесение ошибки осуществляется путем инвертирования 0 и 1. Элемент исходного кода программы, отвечающий за реализацию функции имитации канала связи представлен на рисунке 3.79.

```

If (max_mistake_int > 0) And (i <= (length_array_1 - 1)) And (remainder > 0) Then ' пока
кол-во введенных ошибок не достигло максимума
    If (value <= CInt(probability_mistake)) Then 'если случайное число или
равно, заданной вероятности ошибки
        j = 1 ' выставляем флаг ошибки

If (S_mess_chr(f).ToString = "0") Then ' заменяем биты в сообщении
    S_mess_chr(f) = "1"
Else
    S_mess_chr(f) = "0"
End If

```

Рис. 3.78. Элемент программного кода функции имитации канала связи

Если были заданы двойные ошибки, инвертируются по два соседних бита.

В конечном итоге, после того как подсчет ошибок осуществлен, и их количество и кратность соответствует заданным параметрам, сообщение с ошибками выводится на блок индикации, где искаженные биты выделяются цветом.

Функция вычисления метрик путей

Важным этапом разработки алгоритма декодирования является выбор алгоритма вычисления метрик путей. В качестве такой метрики будет принято расстояние Хэмминга. Расстояние Хэмминга — число позиций, в которых соответствующие символы двух слов одинаковой длины различны. Согласно выбранному алгоритму вычисления метрик путей, биты сообщения суммируются по модулю два со значениями возможных путей и полученному результату подставляется его десятичное значение, соответствующее величине расстояния Хемминга между слагаемыми. За это отвечает функция «Fun_Metric», программная реализация которой приведена на рисунке 3.70, где:

s_1 – биты сообщения;

s_2 – возможные значения путей решетки;

res – результат суммирования по модулю два бит сообщения со значениями возможных путей решетки;

res_1 - десятичное значение, соответствующее величине расстояния Хемминга между слагаемыми.

```

Public Function Fun_Metric(ByVal f_s_mess_1 As String, ByVal f_s_mess_2 As String)
    Dim s_1, s_2 As Integer
    Dim res, res_1 As Integer
    s_1 = CInt(f_s_mess_1)
    s_2 = CInt(f_s_mess_2)
    res = s_1 Xor s_2 ' сумма по модулю два
    Select Case res
        Case "00"
            res_1 = 0
        Case "01"
            res_1 = 1
        Case "10"
            res_1 = 1
        Case "11"
            res_1 = 2
    End Select
    Return res_1
End Function

```

Рис. 3.79. Программный код функции вычисления метрик путей

Из таблицы 3.4 видно, что значение 00 может стать результатом суммирования только двух слов, биты которых не отличаются ни в одной из позиций, а это значит, что расстояние Хемминга между такими словами равно 0. Значения $res = 01$ и $res = 10$ возможны лишь в том случае, если по модулю 2 суммировались слова, соответствующие биты которых отличаются друг от друга в одной позиции, то есть расстояние Хемминга между ними равно 1. Что же касается значения $res = 11$, то оно может получиться только в результате сложения двух слов, биты которых в обеих позициях отличаются, и расстояние Хемминга между этими словами равно 2.

Таблица 3.4. Вычисление метрик путей

№ такта	Входной информационный бит	Состояние регистра сдвига	Сумматор 1	Сумматор 2	Выходные кодовые комбинации
0	-	0000			-
1	1	1000	$1 \oplus 0 \oplus 0 = 1$	$1 \oplus$ $0 \oplus 0 = 1$	11
2	0	0100	$0 \oplus 0 \oplus 0 = 0$	$0 \oplus$ $1 \oplus 0 = 1$	01
3	1	1010	$1 \oplus 1 \oplus 0 = 0$	$1 \oplus$ $0 \oplus 0 = 1$	01
4	1	1101	$1 \oplus 0 \oplus 1 = 0$	$1 \oplus$	01

				$1 \oplus 1 = 1$	
5	0	0110	$0 \oplus 1 \oplus 0 = 1$	$0 \oplus$ $1 \oplus 0 = 1$	11
6	1	1011	$1 \oplus 1 \oplus 1 = 1$	$1 \oplus$ $0 \oplus 1 = 0$	10

Функция декодирования последовательным алгоритмом

Учитывая топологию решетки, первый шаг декодирования был реализован отдельно. В элементе исходного кода приведенного на рисунке 3.71 происходит проверка и выбор возможного пути, зависимости от битов сообщения (проверяются биты массива `S_str_Array()`).

Как мы уже знаем из теории, на первом шаге декодирования состояние декодера – 00, возможны только два пути, к узлам 00 и 10, которым соответствуют кодовые символы с выхода кодера 00 и 11.

Однако на вход декодера могут придти целых четыре комбинации бит сообщения, поэтому, на первом шаге декодирования, было решено связать приход кодовой комбинации 00 и 01 к переходу в узел 00, а кодовые комбинации 10 и 11 к переходу в узел 10. Так мы избавляемся от неопределённостей, заставляя декодер принять решение, даже если пришла не разрешенная комбинация.

```

If ((S_str_Array(i) = "00") Or (S_str_Array(i) = "01")) Then
    state_decoder = "00"
    state_metric = Fun_Metric(state_decoder, S_str_Array(i))
    s_mess = "0"
    f_x1 = Arr_XY(0, f)
    f_y1 = 284
    ...

ElseIf ((S_str_Array(i) = "10") Or (S_str_Array(i) = "11")) Then
    state_decoder = "10"
    state_metric = Fun_Metric(state_decoder, S_str_Array(i))
    s_mess = "1"
    f_x1 = Arr_XY(2, f)
    f_y1 = 384
    ...

End If
GoTo End_1_step

```

Рис. 3.80. Элемент программного кода функции декодирования последовательным алгоритмом

После того как путь выбран, в переменную `state_metric` записывается значение метрики пути, вычисленное описанной функцией `Fun_Metric`. Далее ставится метка 0 или 1,

указывающая на то какой символ передавался, после чего задаются конечные координаты по горизонтали и вертикали отрезка для выбранного пути.

На этом заканчивается первый шаг декодирования и осуществляется безусловный переход на метку End_1_step. Сюда же переходит программа и после отработки всех остальных шагов.

Далее происходит аналогичные операции, только количество возможных путей увеличивается до 4-х (00, 01, 10, 11), осуществляется переход ко второму шагу декодирования. Граф автомата, каковым является наш декодер, принимает вид показанный на рисунке 3.82

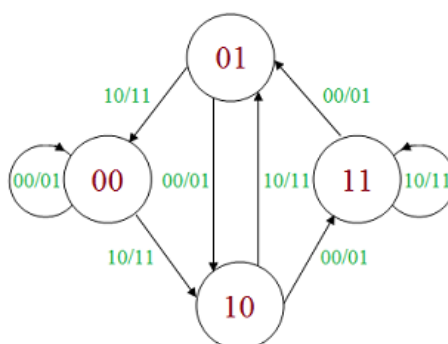


Рис. 3.81. Граф реализуемого последовательного декодера

Стратегия представленная на рисунке 3.82 реализуется при помощи алгоритма подобному тому, что был использован на первом шаге декодирования, теперь декодер так же выбирает путь с кратчайшей метрикой, однако в случае если расстояния Хемминга для обоих путей одинаковы, то прибегает жесткому методу принятия решений:

- Если мы находимся в узле 00 или 11, и пришла комбинация 00 или 01 – считаем что передавался ноль (двигаемся вверх), 10 или 11 – передавалась единица (двигаемся вниз).
- Если мы находимся в узле 01 или 10, и пришла комбинация 00 или 01 – считаем что передавалась единица (двигаемся вниз), 10 или 11 – передавался ноль (двигаемся вверх).

Процесс происходит до тех пор пока не закончится сообщение. Полный исходный код функции декодирования приведен в приложении (см. Приложение И).

Подготовка к выводу результатов декодирования

Для того чтобы в конечном итоге получить символьный результат организуем цикл реверсирования массива символов сообщения, иначе мы получим строку в зеркальном отражении. Организуем цикл перебора массива символов сообщения, разделяем их на пачки по 8 бит и каждую пачку преобразуем в десятичное число по принципу: $2^7 + 2^6 + 2^5 + 2^4 +$

$2^3 + 2^2 + 2^1 + 2^0$ а затем получаем символ из его ASCII-кода. Повторяем операции пока не закончится сообщение, после чего результаты выводятся в соответствующих окнах.

Сверточное декодирование. Практическая часть

ЗАДАНИЕ №1

Вариант 1

Сверточный код: $g_1(X) = 1 + X^2 + X^3$, $g_2(X) = 1 + X + X^3$.

Входная последовательность: $V = 10110$.

На рисунке 3.73 представлен сверточный кодер.

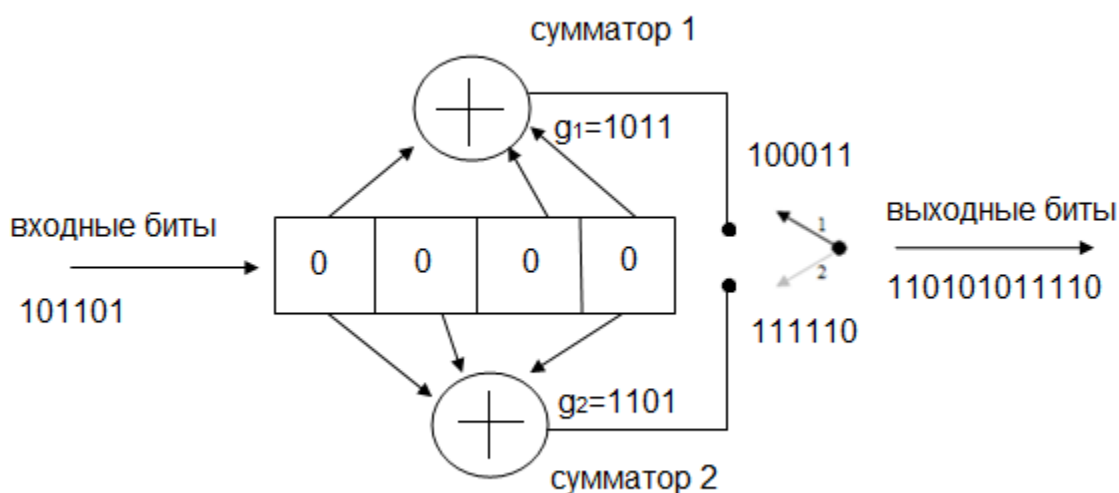


Рис. 3.82. Сверточный кодер

Произвели кодирование входной последовательности V данным сверточным кодером и нарисовали таблицу кодирования информационного потока. Данные представлены в таблице 3.5.

Таблица 3.5. Процесс кодирования информационного потока

№ такта	Входной информационный бит	Состояние регистра сдвига	Сумматор 1	Сумматор 2	Выходные кодовые комбинации
0	-	0000			-
1	1	1000	$1 \oplus 0 \oplus 0 = 1$	$1 \oplus 0 \oplus 0 = 1$	11
2	0	0100	$0 \oplus 1 \oplus 0 = 1$	$0 \oplus 1 \oplus 0 = 1$	01

				$1 \oplus 0 = 1$	
3	1	1010	$1 \oplus 1 \oplus 0 = 0$	$1 \oplus$ $0 \oplus 0 = 1$	01
4	1	1101	$1 \oplus 0 \oplus 1 = 0$	$1 \oplus$ $1 \oplus 1 = 1$	01
5	0	0110	$0 \oplus 1 \oplus 0 = 1$	$0 \oplus$ $1 \oplus 0 = 1$	11
6	1	1011	$1 \oplus 1 \oplus 1 = 1$	$1 \oplus$ $0 \oplus 1 = 0$	10

Построили решетку кодера:

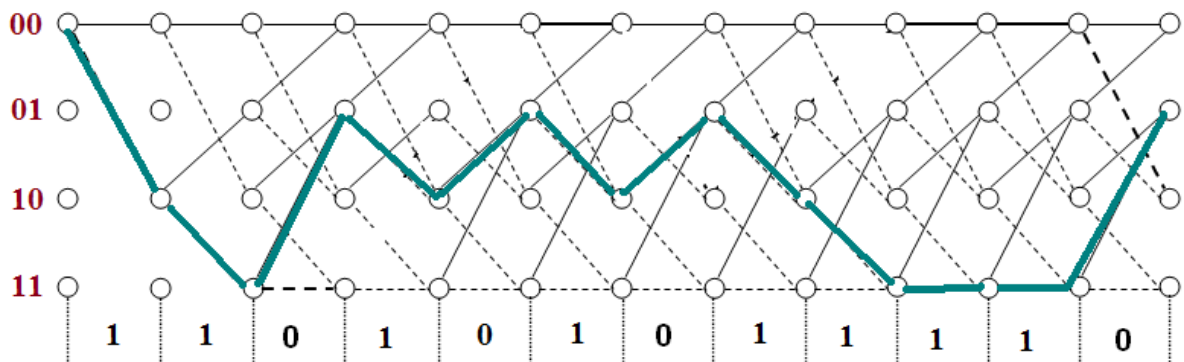


Рис. 3.83. Решетка кодера

Изобразили решетку декодера и произвели декодирование:

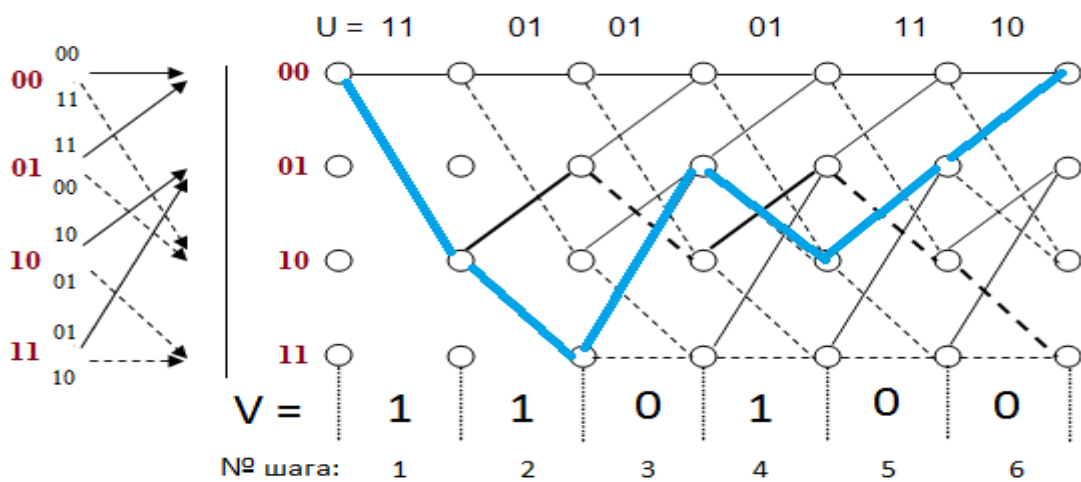


Рис. 3.84. Решетка декодера

В результате были выявлены 6 ошибок: После декодера был получен результат: 11 10 11 01 10 01. Подчеркиванием выделены места, где были совершены ошибки.

Вариант 2

Сверточный код: $g_1(X) = X + X^2 + X^3$, $g_2(X) = 1 + X + X^2$.

Входная последовательность: $V = 010100$.

На рисунке 3.85 представлен сверточный кодер.

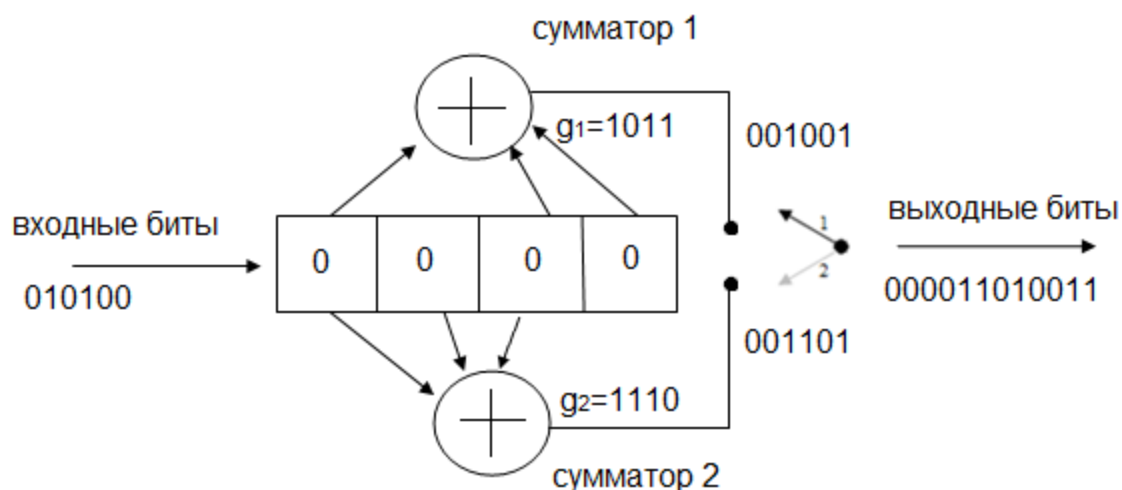


Рис. 3.85. Сверточный кодер

Произвели кодирование входной последовательности V данным сверточным кодером и нарисовали таблицу кодирования информационного потока. Данные представлены в таблице 3.6.

Таблица 3.6. Процесс кодирования информационного потока

№ такта	Входной информационный бит	Состояние регистра сдвига	Сумматор 1	Сумматор 2	Выходные кодовые комбинации
0	-	0000			-
1	0	0000	$0 \oplus 0 \oplus 0 = 0$	$0 \oplus 0 \oplus 0 = 0$	00
2	0	0000	$0 \oplus 0 \oplus 0 = 0$	$0 \oplus 0 \oplus 0 = 0$	00
3	1	1000	$1 \oplus 0 \oplus 0 = 1$	$1 \oplus 0 \oplus 0 = 1$	11

				$0 \oplus 0 = 1$	
4	0	0100	$0 \oplus 0 \oplus 0 = 0$	$0 \oplus$ $1 \oplus 0 = 1$	01
5	1	1010	$1 \oplus 1 \oplus 0 = 0$	$1 \oplus$ $0 \oplus 1 = 0$	00
6	0	0101	$0 \oplus 0 \oplus 1 = 1$	$0 \oplus$ $1 \oplus 0 = 1$	11

Построили решетку кодера:

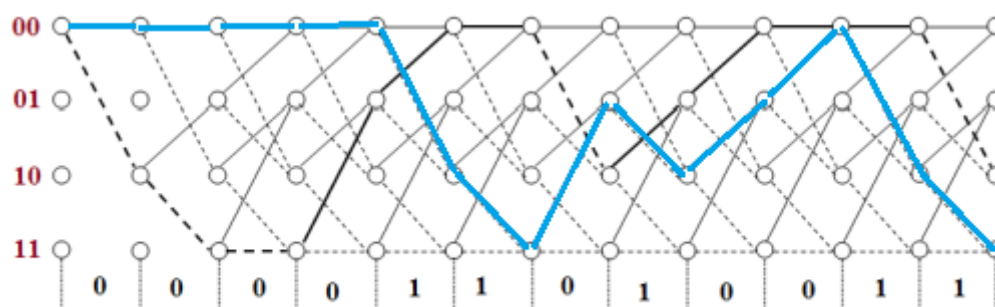


Рис. 3.86. Решетка кодера

Изобразили решетку декодера и произвели декодирование:

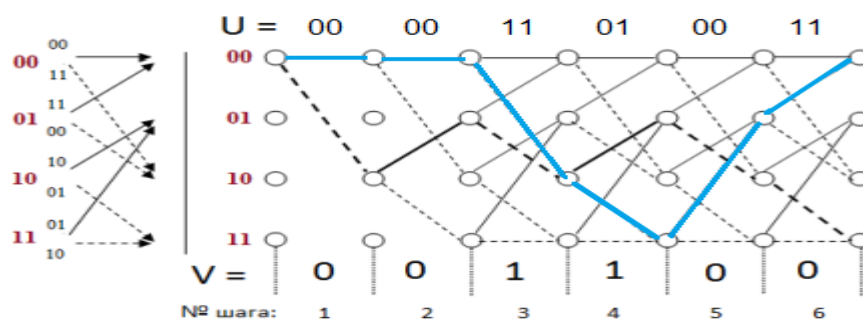


Рис. 3.87. Решетка декодера

В результате были выявлены 5 ошибок: После декодера был получен результат: 00 00 10 11 01 00. Подчеркиванием выделены места, где были совершены ошибки.

ЗАДАНИЕ №2

Построить графики зависимости числа ошибок от вероятности ошибок для одиночной и двойной длины ошибок.



Рис. 3.88. Процесс кодирования и декодирования

График зависимости числа ошибок от вероятности ошибок для одиночной длины ошибок представлен на рисунке 3.89, а для двойной длины ошибок на рисунке 3.90.

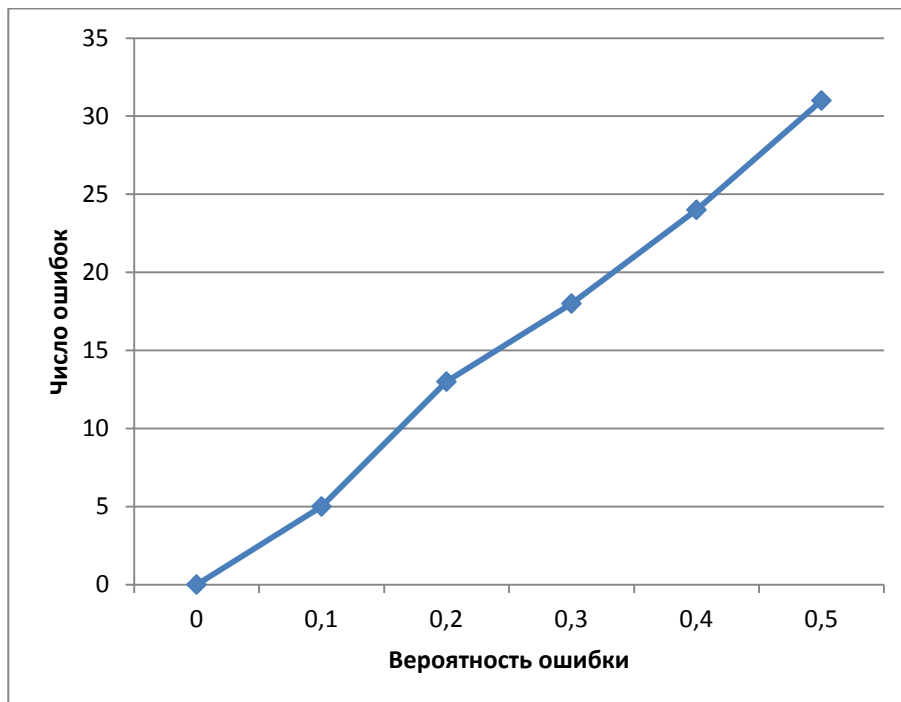


Рис. 3.89. График зависимости числа ошибок от вероятности ошибок для одиночной длины ошибок

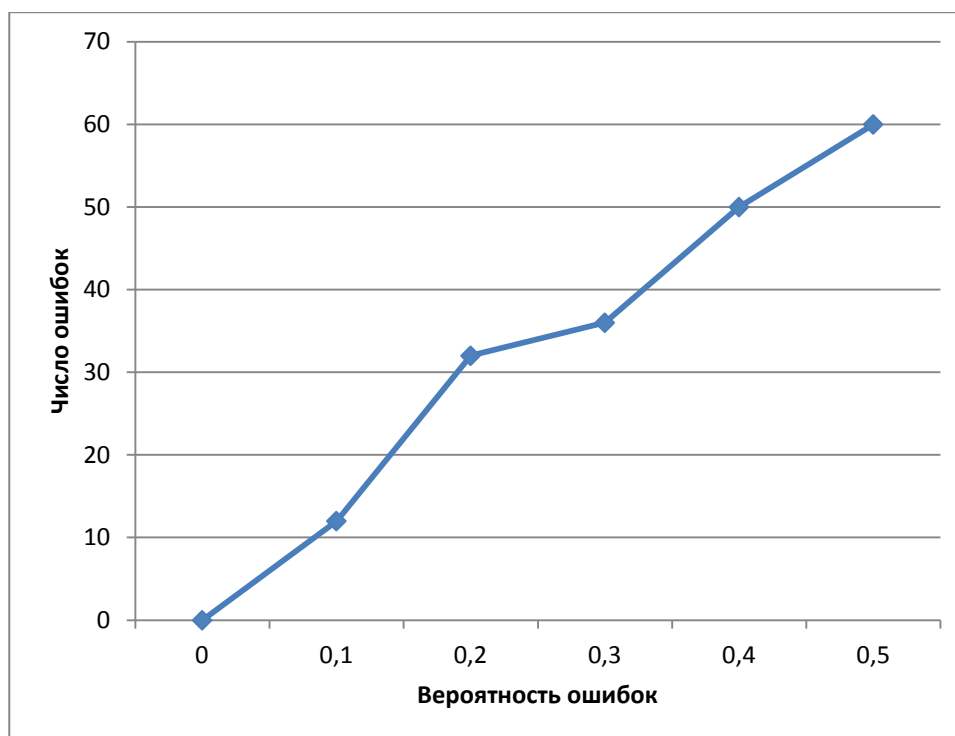


Рис. 3.90. График зависимости числа ошибок от вероятности ошибок для двойной длины ошибок

В процессе выполнения индивидуальной работы была осуществлена работа по созданию программного комплекса для визуализации и исследования метода свёрточного декодирования на основе последовательного алгоритма. Программный комплекс выполнен в полном соответствии с техническим заданием, и отвечает всем заявленным в нём требованиям.

Были произведены лабораторные испытания программного комплекса, в ходе которых была отмечена высокая эффективность его использования в качестве макета наглядно демонстрирующего процессы свёрточного кодирования и последовательного алгоритма декодирования.

3.4. Декодирование сверточных кодов по методу Витерби с использованием ПО MATLAB

Алгоритм свёрточного декодирования Витерби [10]

В 1967 году Витерби разработал и проанализировал алгоритм, в котором, по сути, реализуется декодирование, основанное на принципе максимального правдоподобия; однако в нем уменьшается вычислительная нагрузка за счет использования особенностей структуры конкретной решетки кода. Преимущество декодирования Витерби, по сравнению с декодированием по методу "грубой силы", заключается в том, что сложность декодера Витерби не является функцией количества символов в последовательности кодовых слов.

Алгоритм включает в себя вычисление *меры подобия* (или *расстояния*), между сигналом, полученным в момент времени t , и всеми путями решетки, входящими в каждое состояние в момент времени t_i . В алгоритме Витерби не рассматриваются те пути решетки, которые, согласно принципу максимального правдоподобия, заведомо не могут быть оптимальными. Если в одно и то же состояние входят два пути, выбирается тот, который имеет лучшую метрику; такой путь называется *выживающим*. Отбор выживающих путей выполняется для каждого состояния. Таким образом, декодер углубляется в решетку, принимая решения путем исключения менее вероятных путей. Предварительный отказ от маловероятных путей упрощает процесс декодирования. В 1969 году Омуре (Omura) показал, что основу алгоритма Витерби составляет оценка максимума правдоподобия. Отметим, что задачу отбора оптимальных путей можно выразить как выбор кодового слова с *максимальной метрикой правдоподобия* или *минимальной метрикой расстояния*.

Алгоритм Витерби является достаточно мощным алгоритмом декодирования, при этом сложность его аппаратной реализации невысока. В настоящее время алгоритм Витерби применяется во многих стандартах беспроводной связи, таких как IEEE 802.11a/g, WiMAX, DAB/DVB, WCDMA, GSM.

Наиболее подходящей элементной базой для аппаратной реализации декодера Витерби являются программируемые логические интегральные схемы (ПЛИС).

Работа декодера Витерби

Алгоритм Витерби находит широкое применение и реализует *поиск максимально правдоподобного пути* на кодовой решетке с отбрасыванием части наименее правдоподобных вариантов путей на каждом шаге декодирования.

Алгоритм Витерби характеризуется постоянством вычислительной работы, однако *сложность* декодера Витерби растет, как при всех переборных алгоритмах, *по экспоненциальному закону* от длины кодового ограничения сверточного кода. Поэтому алгоритм Витерби используется для декодирования коротких сверточных кодов.

Алгоритм заключается в повторении одного основного шага. На каждой из последующих диаграмм рис. 3.91 этот шаг изображен подробно.

Метрика пути (МП) есть сумма метрик ветвей, образующих некоторый путь на решетчатой диаграмме. Путь конечной длины оканчивается в определенном состоянии. Метрика состояния (МС) равна метрике пути, который заканчивается в данном состоянии.

Шаг декодирования состоит в обработке декодером принимаемых из канала данных в интервале между двумя соседними отсчетами.

На рисунке показано развитие процесса декодирования символов СК со скоростью $R_{\text{код}} = 1/2$ и длиной кодирующего регистра $K = 3$. На вход декодера поступают пары символов из канала: (...11,10,00,11,01...) (декодирование с жестким решением). Цифрами около ветвей обозначены метрики ветвей, цифры в кружках обозначают метрики состояний.

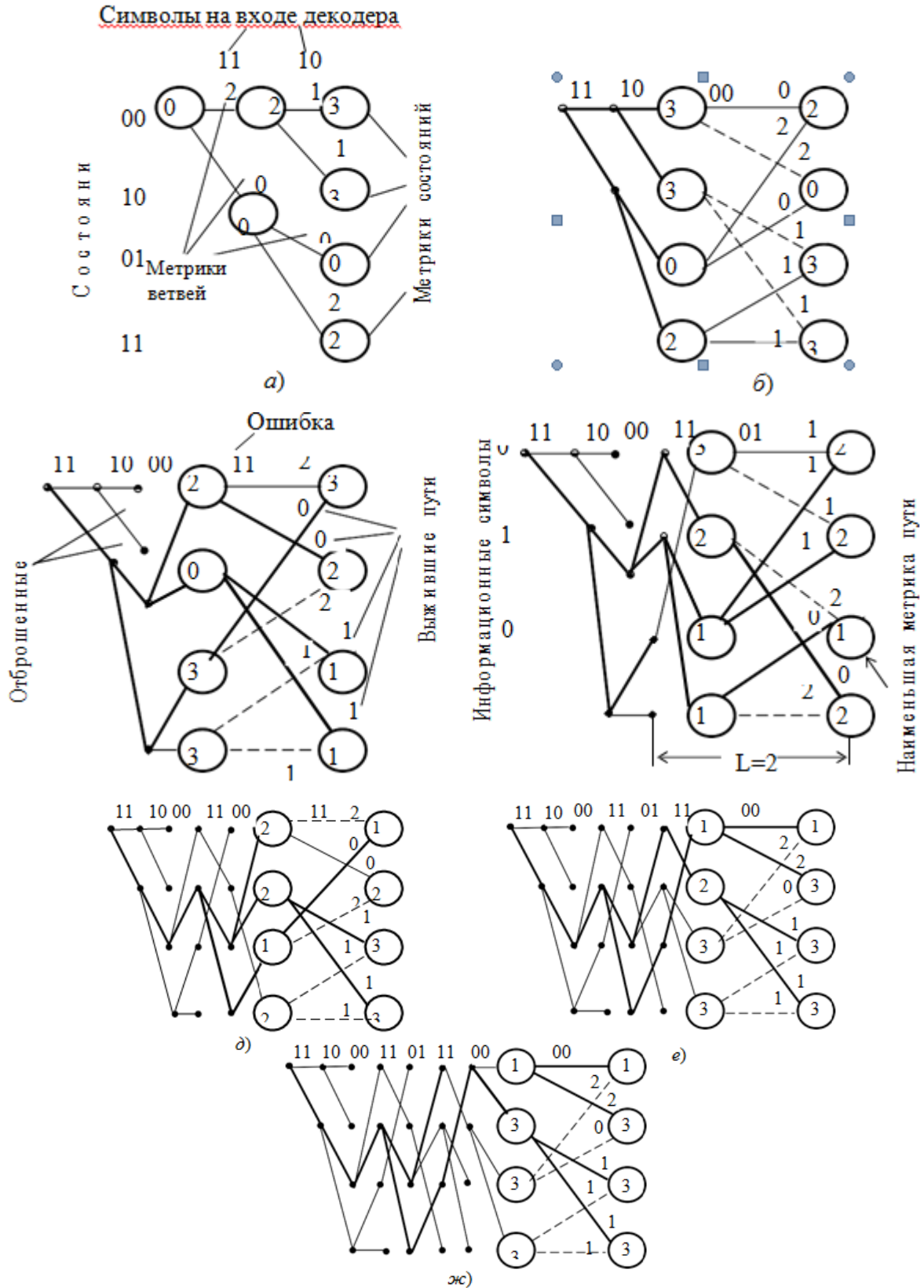


Рис. 3.91. Декодирование кода(7,5)

К каждому новому состоянию ведут два пути. К примеру, к состоянию 00 ведут пути из предыдущих состояний 00 и 01. На i -м шаге декодирования декодер вычисляет метрики путей как суммы метрик предыдущих состояний и метрик входящих ветвей.

Далее производится попарное сравнение метрик путей, входящих в каждое из состояний (пары показаны фигурными скобками). В результате сравнения *выбирается меньшая метрика*, и она считается метрикой данного состояния для последующего шага декодирования. *Путь*, входящий в данное состояние с меньшей метрикой, считается *выжившим*. На рис. 3.91 отрезки выживших путей показаны сплошной линией. Пути, входящие в состояния с большими метриками, считаются отмершими (оборванными). Они показаны на решетчатой диаграмме пунктиром.

Таким образом, на каждом шаге декодирования в соответствии с алгоритмом Витерби, *в каждом из состояний* решетчатой диаграммы производятся *однотипные операции*:

1) *Сложение метрик* предыдущих состояний с метриками соответствующих ветвей.

2) *Сравнение метрик* входящих путей.

3) *Выбор путей* с наименьшими метриками, величины которых используются как метрики состояний на последующем шаге декодирования. Если метрики сравниваемых путей одинаковы, то выбор одного из двух путей производится случайным образом.

На каждом шаге декодирования половина возможных продолжений путей отбрасывается. Другая половина образует продолжения путей для следующего шага декодирования, на котором вновь появляются два варианта продолжения каждого пути. Это обеспечивает *постоянство количества вычислений*, производимых на каждом шаге. *Декодер прослеживает* по кодовой решетке *путь*, имеющий *минимальное расстояние* от пути, который порождает кодер.

Таким образом, декодер, выбирающий на решетчатой диаграмме путь с наименьшей метрикой, *минимизирует вероятность ошибки*. Поскольку при декодировании анализу подвергаются последовательности конечной длины L , алгоритм не является строго оптимальным. Результаты расчетов и моделирования показывают, что при соответствующем выборе величины $L > (6...7)v$ можно получить результаты декодирования, достаточно близкие к оптимальным. *Сложность реализации алгоритма Витерби* для декодирования СК можно оценить по количеству ветвей кодовой решетки,

обрабатываемых декодером на длине декодирования L , с учетом сложности каждого шага решетки.

Декодер состоит из АЦП в каналах X и Y , вычислителя метрик ветвей, процессора, в котором производятся операции сложения, сравнения и выбора, устройства памяти путей, которые выжили, и мажоритарного элемента МЭ, в котором выбирается путь с наибольшей метрикой. Оптимальное значение шага квантования зависит от отношения сигнал/шум на входе АЦП. При восьми уровнях квантования минимум потерь обеспечивается при отношении размаха сигнала к шагу квантования, равном (4,5...5,5).

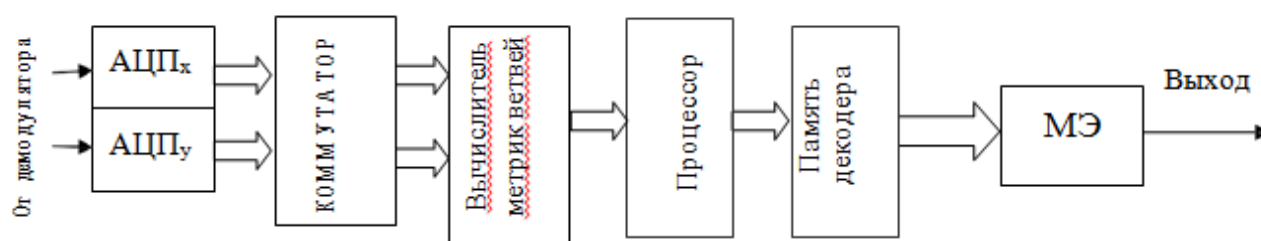


Рис. 3.92. Структурная схема декодера Витерби

Разработка программного интерфейса для исследования характеристик и визуализации основных преобразований.

На данном этапе необходимо реализовать и проанализировать описанный выше алгоритмы декодирования Витерби на практике в программной среде Matlab 2015. Необходимо построить соответствующие схемы, которые будут в себя включать источник сообщения, кодер, декодер, канал передачи, анализатор ошибок и соответственно устройства визуализации полученных результатов.

Индивидуальности сверточного кодировки. Библиотека Communications Blockset поддерживает разомкнутые либо бинарные замкнутые коды, которые могут быть описаны структурой сетки либо рядом порождающих полиномов. Это употребляет метод Viterbi, чтоб выполнить твердое и мягкое декодирование. Библиотека также включает декодер, максимизирующий апостериорную возможность, который может употребляться для мягкого декодирования сверточных кодов.

Характеристики сверточного кодировки. Процессы сверточного кодировки употребляют блоки Convolutional Encoder, Viterbi Decoder, и/либо APP Decoder из библиотеки Convolutional. Если параметр маски требуется и в кодирующем устройстве и в декодере, используйте одно и то же значение в обоих блоках. Блоки в подбиблиотеке Convolutional подразумевают, что используется один из 2-ух разных представлений сверточного кодера: если проектируется кодирующее устройство, используя диаграмму со

сдвиговыми регистрами и сумматорами по модулю 2, можно вычислить порождающую полиномиальную матрицу кода и потом применять функцию `poly2trellis` (из Communications Toolbox), чтоб произвести подобающую решетчатую (trellis) диаграмму маски параметра автоматом. Если проектируется кодирующее устройство, используя решетчатую диаграмму, можно выстроить решетчатую (trellis) диаграмму в MATLAB и применять его в качестве параметра маски.

Внедрение полиномиального описания в блоках. Для использования полиномиального описания в блоках Convolutional Encoder, Viterbi Decoder, либо APP Decoder используется сервисная функция `poly2trellis` из Communications Toolbox. Эта функция воспринимает полиномиальное описание и преобразовывает это в решетчатое описание. К примеру, последующая команда вычисляет решетчатое описание кодирующего устройства, длина ограничения которого равна 5 и чьи порождающие полиномы 35 и 31: `trellis = poly2trellis (5, [35 31])`

Определение характеристик кодировки. Блоки Convolutional Encoder и Viterbi Decoder могут выполнить этот код, если у их характеристик установлены надлежащие значения. Длина ограничения кодирующего устройства - вектор длины 2, потому кодирующее устройство имеет два входа. Элементы этого вектора указывают на число сохраненных битов в каждом сдвиговом регистре, включая текущие входные биты. Подсчет ячеек памяти в каждом сдвиговом регистре на диаграмме и добавлении того для текущих входов приводят к длине ограничения, равного [5 4]. Определяя порождающую матрицу кода как матрицу размера 2 X 3 восьмеричных чисел, используйте элемент в i - ой строке и j - ом столбце, чтоб указать, как i - ый вход повлияет на j - ый выход. К примеру, чтоб вычислить элемент во 2-ой строке и 3-ем столбце, заметьте, что последнее левое и два самых правых элемента во 2-м сдвиговом регистре диаграммы влияют на сумму, что сформировывает 3-ий выход. Эта информация в двоичном коде равна 1011, что эквивалентно восьмеричному числу 13. Полная порождающая матрица кода равна [27 33 0; 0 5 13]. Для использования длины кодового ограничения и характеристик порождающей матрицы кода в блоках Convolutional Encoder and Viterbi Decoder, примените функцию `poly2trellis`, чтоб преобразовать эти характеристики в решетчатую диаграмму.

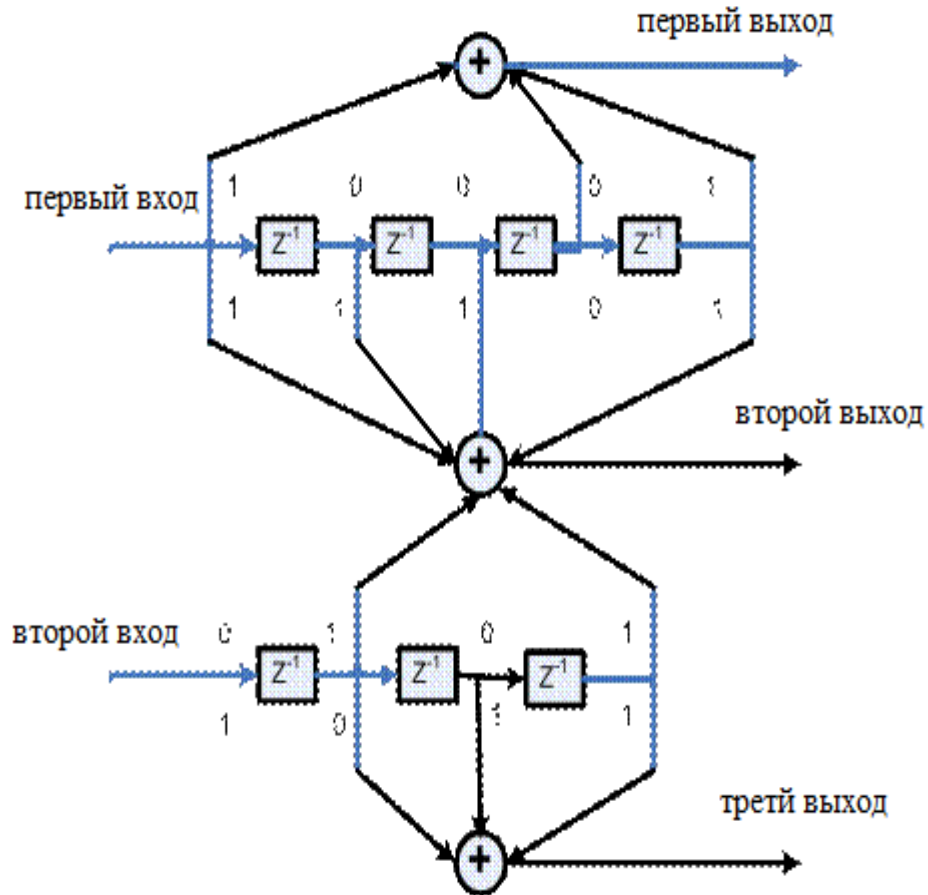


Рис. 3.93. Модель кодирующего устройства

Моделирование кодера Витерби

На рисунке представлена модель кодирующего устройства.

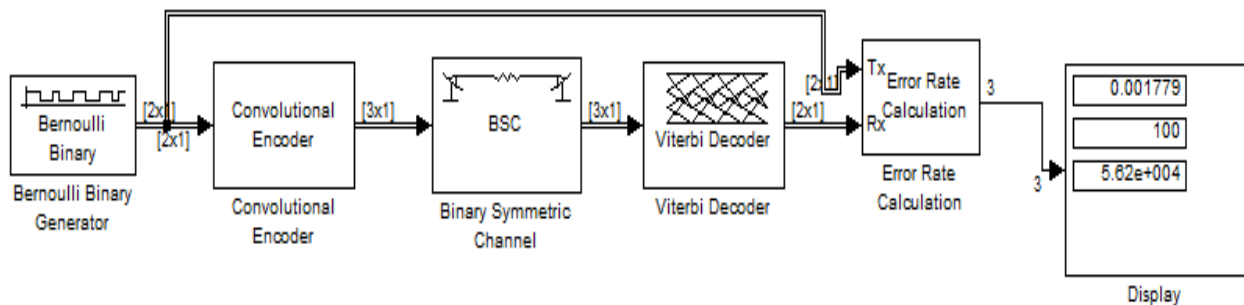
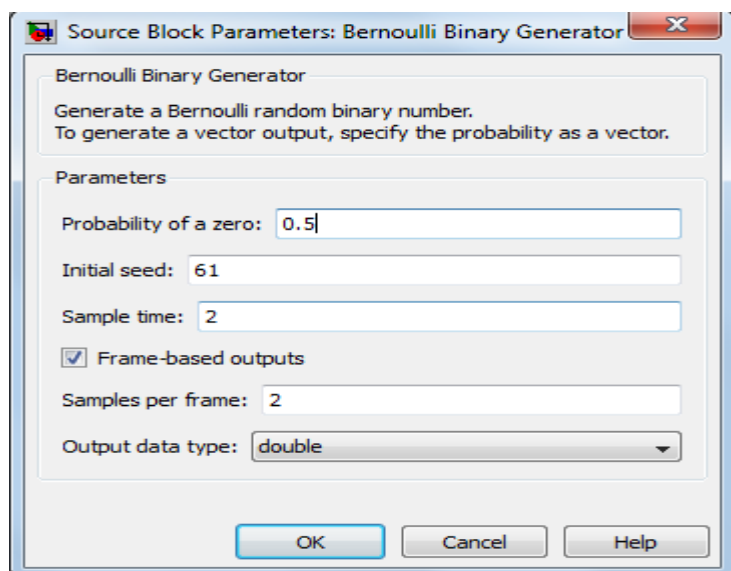


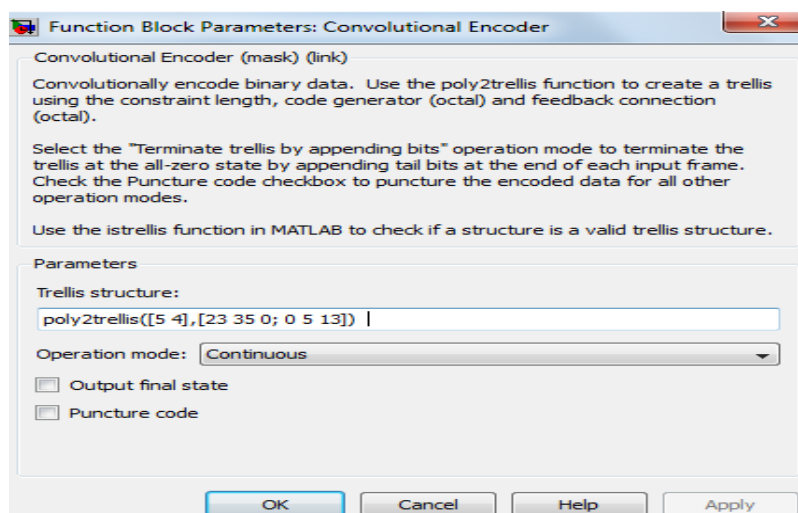
Рис. 3.94. Линия передачи с применением декодера Витерби

Чтоб выстроить модель, соберите и конфигурируйте блоки.

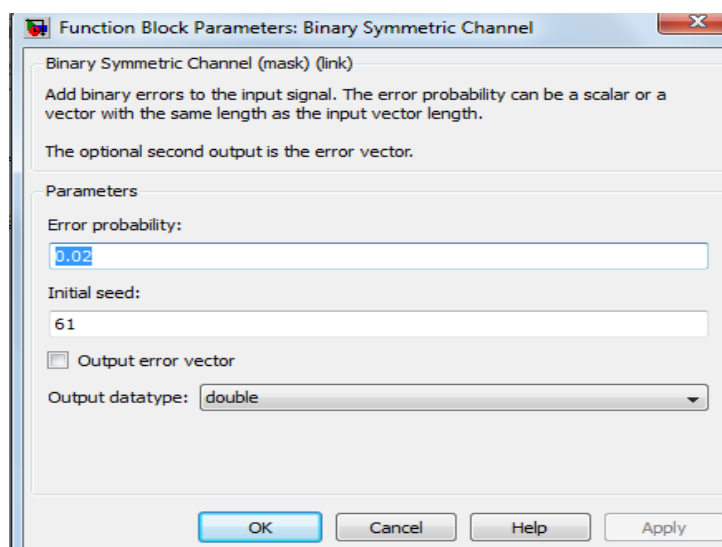
1. Блок Bernoulli Binary Generator из библиотеки Comm Sources.



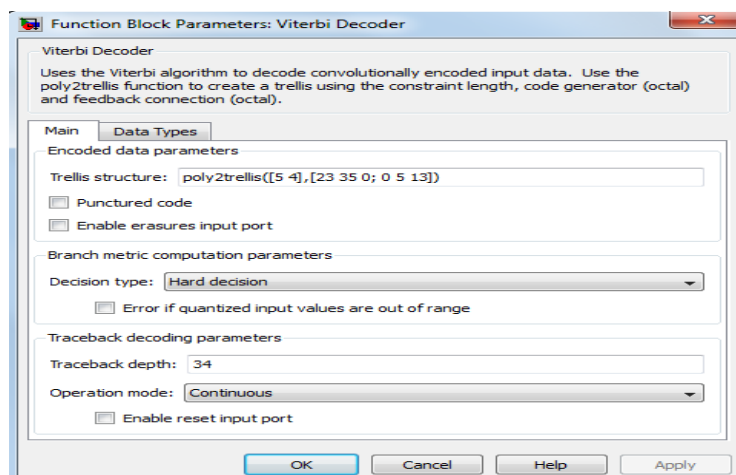
2. Блок Convolutional Encoder



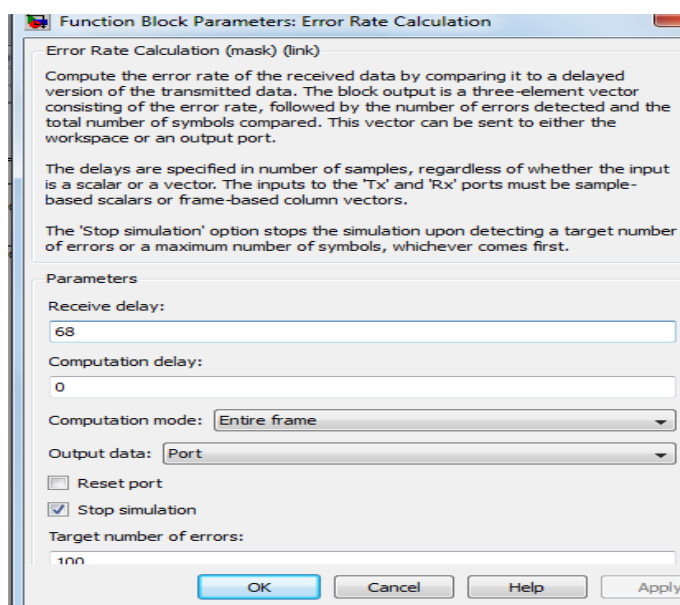
3. Блок Binary Symmetric Channel из библиотеки Channels



4. Блок Viterbi Decoder



5. Блок Error Rate Calculation из библиотеки Comm Sinks.



6. Блок Display из библиотеки Simulink Sinks. Потяните базисный край знака, чтоб сделать показ, довольно огромным для 3-х записей. Соедините блоки как на рисунке. Из меню Simulation, выберите Configuration parameters. В диалоговом окне установите параметр Stop time, равным inf.

Инструкции по размерам матрицы возникают на соединительных линиях, лишь если выбрать параметр Signal Dimensions подменю Port/signal displays из меню Format модели.

Кодирующее устройство воспринимает вектор кадра размера 2×1 и производит вектор кадра размера 3×1 , в то время как декодер делает обратное. Параметр Samples per frame parameter в блоке Bernoulli Binary Generator принят равным 2, так как блок должен произвести слово сообщения длины 2. Параметр Receive delay в блоке Error Rate Calculation принят равным 68, который является длиной вектора (2) восстановленного сообщения значения Traceback depth (34) в блоке Viterbi Decoder. Если исследуется переданные и

приобретенные сигналы, как матрицы в рабочем пространстве MATLAB, Вы видите, что 1-ые 34 строки восстановленного сообщения состоят из нулей, в то время как следующие ряды - расшифрованные сообщения. Таким образом, задержка приобретенного сигнала - 34 вектора длины 2, либо 68 отсчетов. Пуск модели производит выход монитора, состоящий из 3-х чисел: скорость роста ошибок, общее количество ошибок и общее количество ошибок сравнений, которые делает блок Error Rate Calculation во время моделирования. (1-ые два числа меняются в зависимости от Ваших значений Initial seed в блоках Bernoulli Binary Generator и Binary Symmetric Channel). Остановка моделирования после 100 ошибок происходит, так как параметр Target number of errors установлен равным 100, в блоке Error Rate Calculation. Скорость роста ошибки обязана быть меньше 0,02, как Error probability в блоке Binary Symmetric Channel.

Мягкое декодирование. Моделирование делает случайный сигнал двоичного сообщения, кодирует сообщение в сверточный код, модулирует код, используя двоичную манипуляцию со смещением фазы, (BPSK) и добавляет белый Гауссовский шум к модулированному сигналу, чтоб моделировать шум канала. Тогда, моделирование готовит полученную информацию для блока декодирования. В конце концов, моделирование ассоциирует декодированную информацию с сигналом начального сообщения, чтоб вычислить скорость роста ошибок. Конец моделирования наступает после обработки 100 битов ошибок либо 107 битов сообщения, после прибытия первого.

Длина ограничения кодирующего устройства - скаляр, потому что у кодирующего устройства есть один вход. Значение длины ограничения - число битов, сохраненных в сдвиговом регистре, включая текущий вход. Есть 6 регистров памяти, и текущий вход составляет один бит. Таким образом, длина ограничения кода - 7. Кодовый генератор является матрицей размера 1 x 2 матрица восьмеричных чисел, так как кодирующее устройство имеет вход и два выхода, 1-ый элемент в матрице показывает на значение входа, влияющего на 1-ый выход, и 2-ой элемент в матрице показывает, на значение входа, влияющего на 2-ой выход.

К примеру, 1-ый выход в диаграмме кодирующего устройства – сумма по модулю 2 самого правого и 4 последних левых частей в массиве входных значений диаграммы. Двоичное число с семью цифрами 1111001 отражает эту информацию и эквивалентно восьмеричному числу 171. Восьмеричное число 171, таким образом, становится первым входом порождающей кодовой матрицы. Тут, любая тройка употребляет последний левый бит как самый значимый бит. 2-ой выход соответствует двоичному числу 1011011, который эквивалентен восьмеричному номеру 133. Порождающий генератор кода потому [171 133].

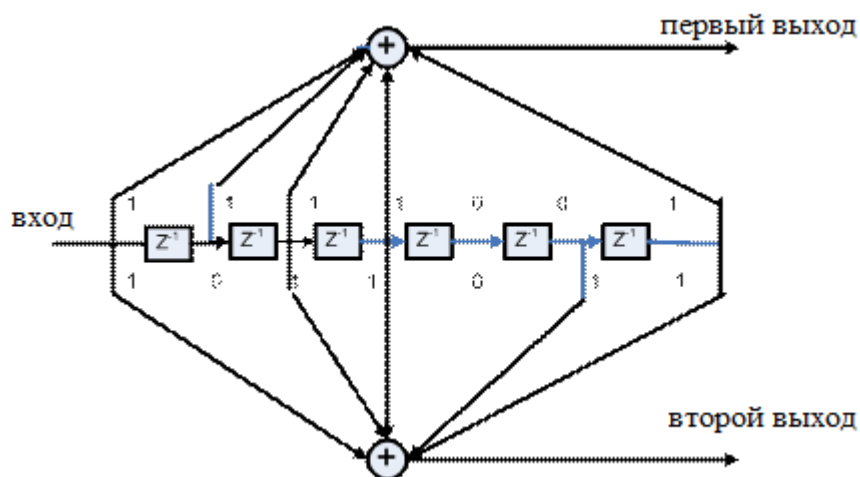


Рис.3.95. Модель кодирующего устройства

Параметр Trellis structure в блоке Convolutional Encoder говорит о том, какой блок, кодирует, когда поступают данные. В этом варианте, функция poly2trellis, в Communications Toolbox, преобразовывает длину кодового ограничения и пару восьмеричных чисел в решетчатую диаграмму. В то время как данные о сообщении, входящие в блок Convolutional Encoder, являются скалярным потоком битов, закодированные данные являются потоком бинарных векторов длиной 2.

Карта приобретенных данных. Приобретенные данные, другими словами, выход блока Convolutional Encoder, состоят из комплексных чисел, которые находятся в спектре меж -1 и 1. Чтоб вернуть оригинальное двоичное сообщение, приемная часть модели обязана декодировать сверточный код. Блок Viterbi Decoder в данной модели ждет, что ее входные данные будут целые числа меж 0 и 7. Демодулятор, рядовая подсистема в данной модели, преобразовывает приобретенные данные в формат, который блок Viterbi Decoder может интерпретировать подабающим образом. Наиболее точно, подсистема демодулятора: Конвертирует приобретенные данные в настоящий сигнал, удаляя его мнимую часть. Уместно представить, что мнимая часть приобретенных данных не содержит существенную информацию, так как мнимая часть переданных данных – есть ноль (игнорирование малых ошибок округления) и так как шум канала не чрезвычайно мощнейший. Восстанавливает приобретенные данные, деля на его текущее обычное отклонение и потом умножая на -1. Квантует нормализованные данные, используя три бита. Композиция данной карты и решения карты блока Viterbi Decoder на сто процентов изменяет модуляцию BPSK, чтоб блок BPSK Modulator Baseband выступил на передающей стороне данной модели. Декодирование сверточного кода. После получения данных подабающим образом нанесенных на карту векторов длиной 23-битовых значений решения, блок Viterbi Decoder

декодирует его. Блок употребляет метод мягенького декодирования с 23 разными входами поэтому, что параметр Decision type - есть Soft Decision и параметр Number of soft decision bits равен 3. Интерпретация данных мягенького решения. Когда характеристики Decision type установлен в состояние Soft Decision, блок Viterbi Decoder просит входных значений меж 0 и $2b-1$, где b – есть параметр Number of soft decision bits. Блок интерпретирует 0 как самое достоверное решение того, что бит кодового слова, является 0 и интерпретирует $2b-1$ как самое достоверное решение того, что бит кодового слова, является 1. Значения меж этими крайностями представляют наименее достоверные решения.

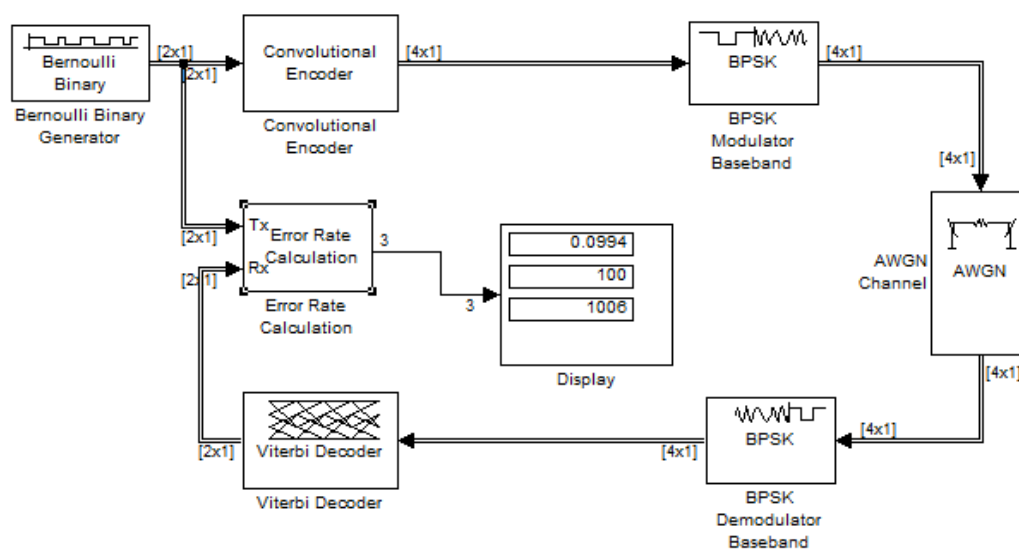


Рис. 3.96. Схема декодера Витерби в MATLAB

	Вер-ть ошибки 0,02	Вер-ть ошибки 0,03
Код (5, [35 31])	скорость роста ошибок: 0.5435 общее количество ошибок: 100 общее количество ошибок сравнений: 184	скорость роста ошибок: 0.5435 общее количество ошибок: 100 общее количество ошибок сравнений: 184
Код ([5 4],[23 35 0; 0 5 13])	скорость роста ошибок: 0,001779 общее количество ошибок: 100 общее количество ошибок сравнений: 5.62e+0.04	скорость роста ошибок: 0,01489 общее количество ошибок: 100 общее количество ошибок сравнений: 6714

Блоки программы декодера Витерби, реализованного в Matlab [21]

В блоке формирования узлов сетки выполняются следующие функции:

1. Формируется матрица фрагмента решетчатой диаграммы размера $n \times m$, в которой количество столбцов соответствует числу состояний кодера (узлов), а количество строк равно числу путей, подходящих к каждому узлу. Каждому элементу матрицы соответствуют ответвляемые слова при переходах между состояниями. Переходы между состояниями повторяются на протяжении всей решетки.

2. Формируется матрица переходов между состояниями.

3. Каждому переходу к конкретному узлу соответствует определенный входной бит (единица, либо ноль). Формируем матрицу размера $(n \times m)$ состоящую только из нулей и единиц .

4. Для каждого входного дибита (слова) декодера вычисляются все возможные значения метрик по фрагменту решетчатой диаграммы кодера, сформированной в предыдущем блоке. В каждый момент времени происходит операция сложения по модулю два входного дибита с соответствующим данному моменту времени ответвляемым словом. Для исключения нижней части решетчатой диаграммы в начальные моменты времени мы искусственно увеличиваем значение метрик переходов нижней части решетки.

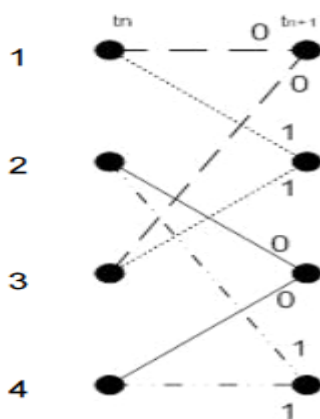


Рис. 3.97. Фрагмент решетчатой диаграммы декодера

Блок удаления путей:

1. В каждом узле, сравнивая метрики входных путей, обнуляем путь с наибольшей метрикой, тем самым убирая его из рассмотрения, но при этом могут возникнуть «тупиковые» пути.

2. Для удаления «тупиковых» путей, поочередно перебирая возможные m состояний в данный момент времени ищем отсутствие какого-либо элемента матрицы переходов состояний, если элемент отсутствует, то пути в предыдущих фрагментах решетки (матрицы) удаляются.

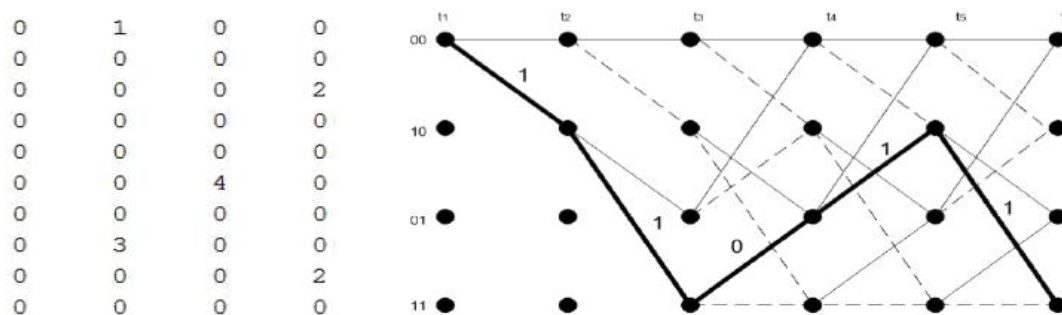


Рис. 3.98. Декодирование по выделенному пути

Блок декодирования:

Последовательно проверяем количество ненулевых переходов в n строках. Если такой переход единственный, то находим позицию этого элемента, и определяем какой бит был подан на вход кодера

Вопрос 1. Исследование методов построения кодеров непрерывных кодов.

В рабочем поле необходимо собрать схему для работы сверточного кода (5, [35 31]). Схема представлена на рисунке 3.99

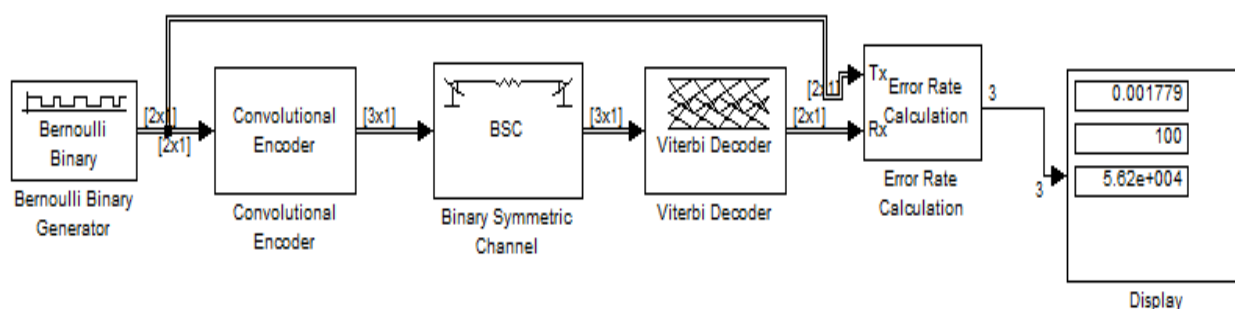


Рис. 3.99. Линия передачи с применением декодера Витерби.

В состав линии с кодированием входят:

1. BernoulliBinaryGenerator
2. Convolutional Encoder
3. Binary Symmetric Channel (каналпередачи)
4. Viterbi Decoder

5. Error Rate Calculation (анализаторошибок)

6. Display

Устанавливаем характеристики блоков для кода (5, [35 31])

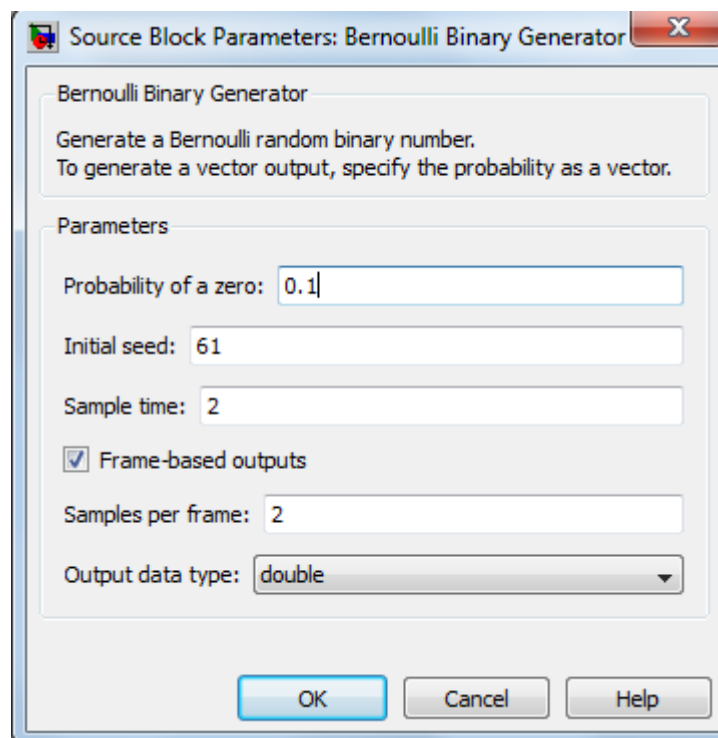


Рис. 3.102. Параметры Bernoulli Binary Generator

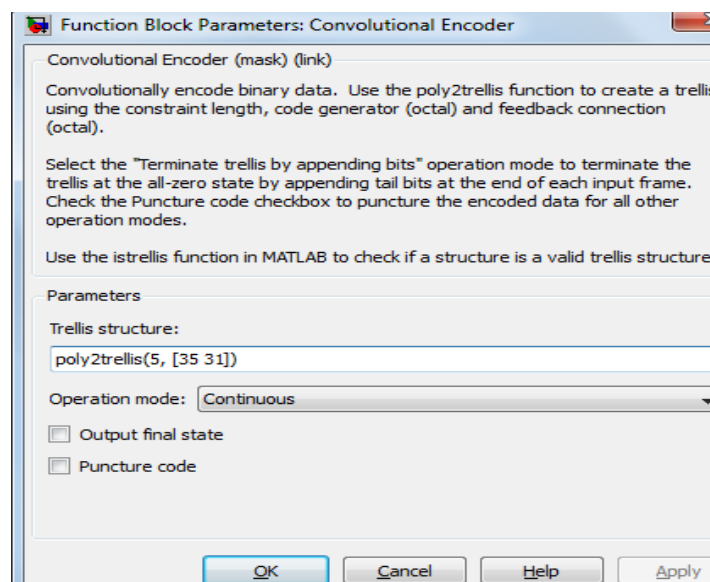


Рис. 3.103. Convolutional Encoder

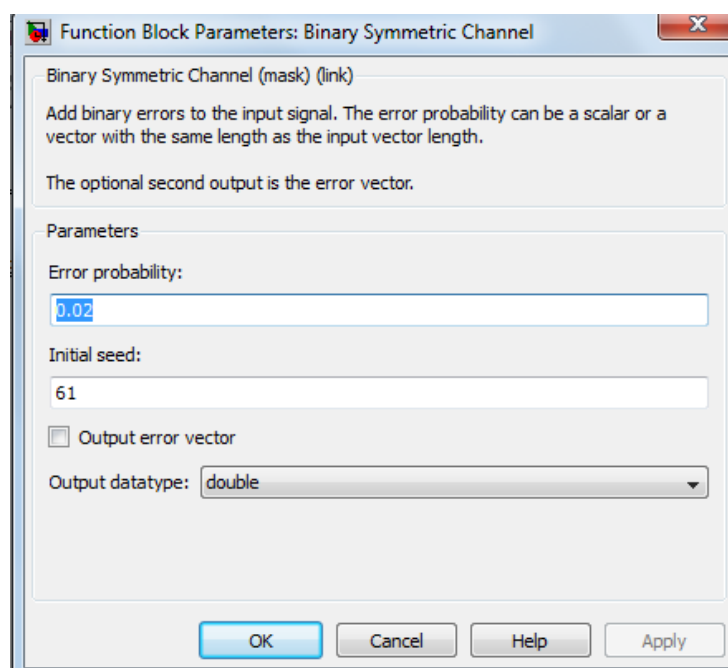


Рис. 3.104. Binary Symmetric Channel

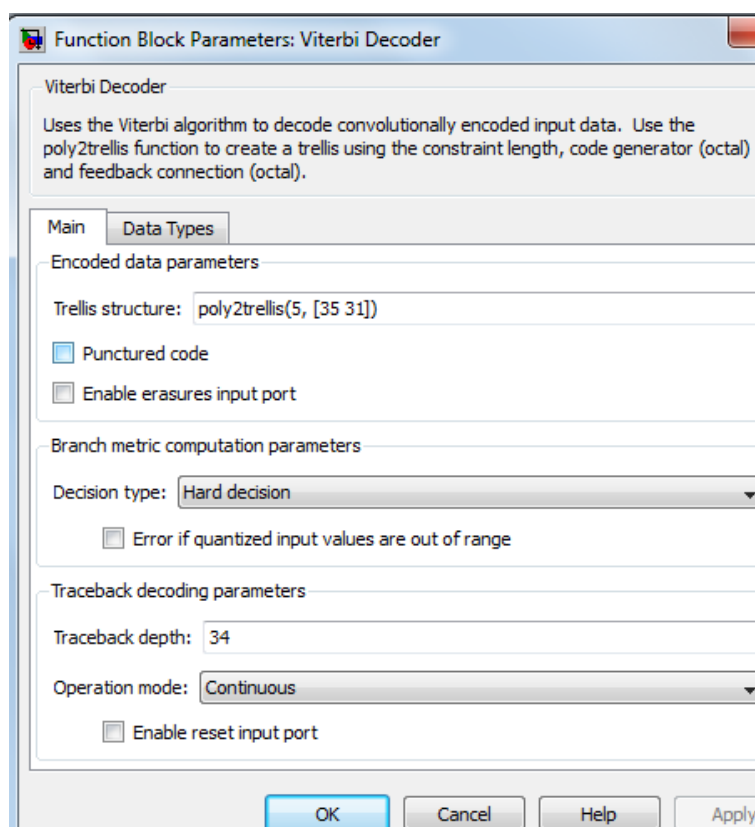


Рис. 3.100 Viterbi Decoder

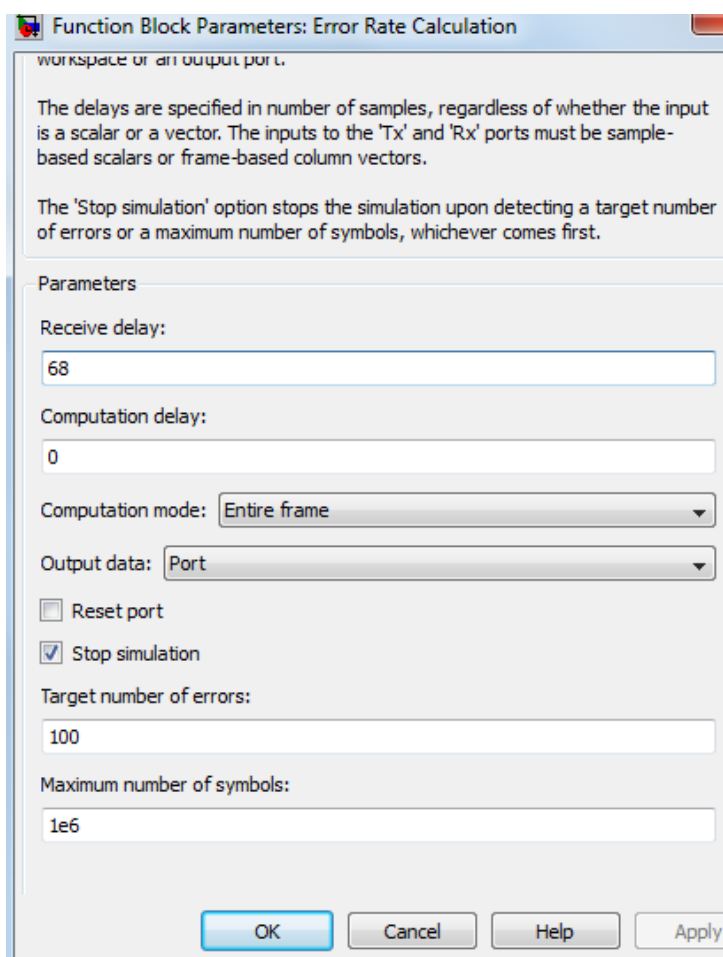


Рис. 3.101. Параметры Error Rate Calculation

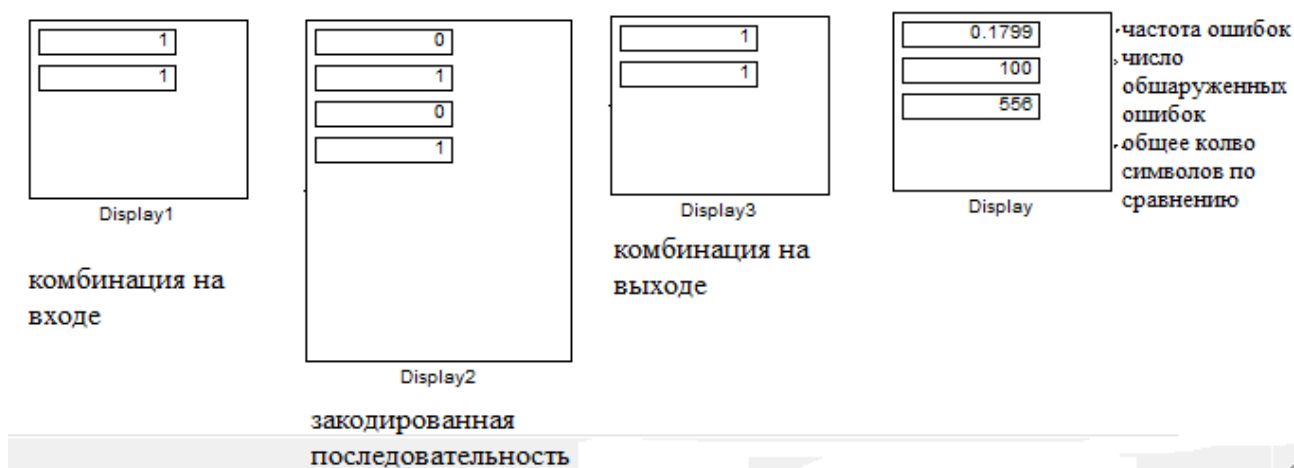


Рис. 3.102. Комбинация на входе, Закодированная последовательность, Комбинация на выходе, Ошибки (вероятность ошибок равна 0,02)

Анализируя рисунок выше, можно сделать вывод, что комбинация на входе совпадает с комбинацией на выходе, таким образом, передача осуществилась удачно. Что касается

ошибок, то их частота равна 0,1799, число ошибок равно 100, общее количество символов по сравнению равно 556.

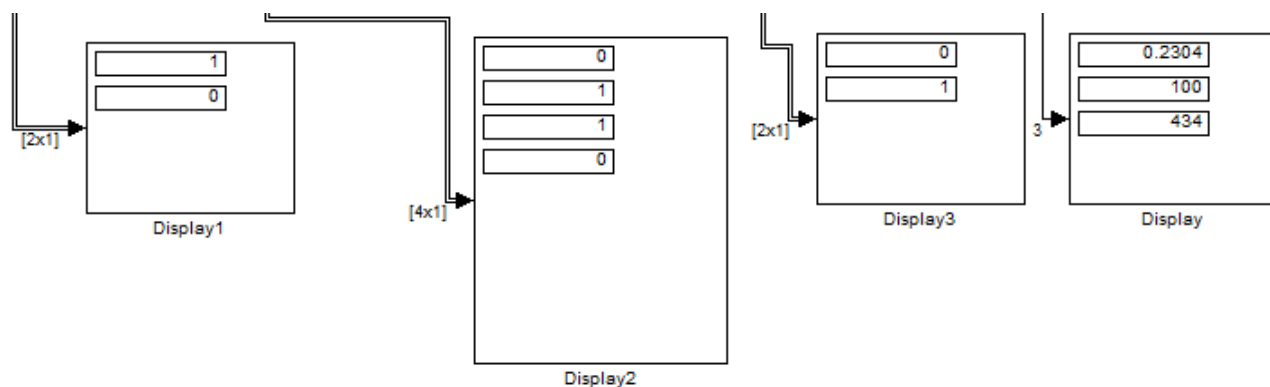


Рис. 3.103. Комбинация на входе, Закодированная последовательность, Комбинация на выходе, Ошибки (вероятность ошибок равна 0,08)

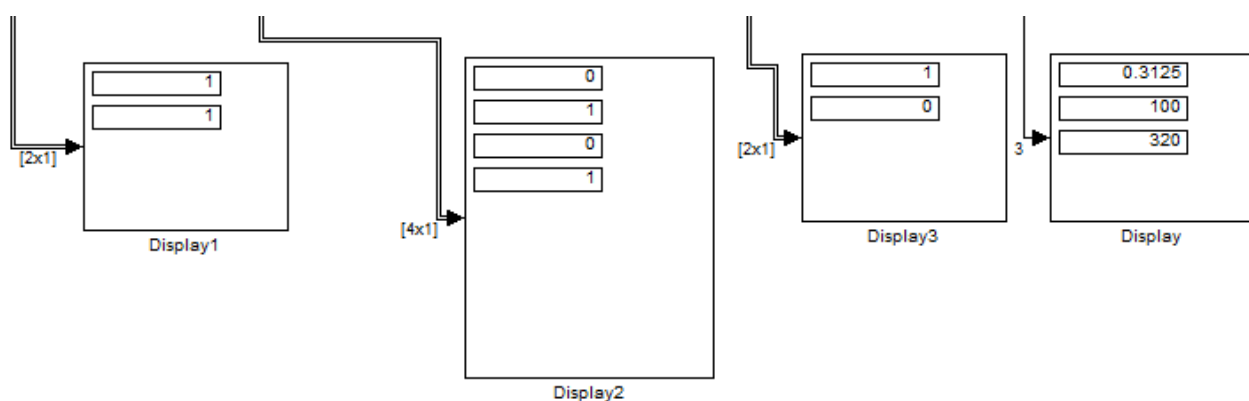


Рис. 3.104. Комбинация на входе, Закодированная последовательность, Комбинация на выходе, Ошибки (вероятность ошибок равна 0,12)

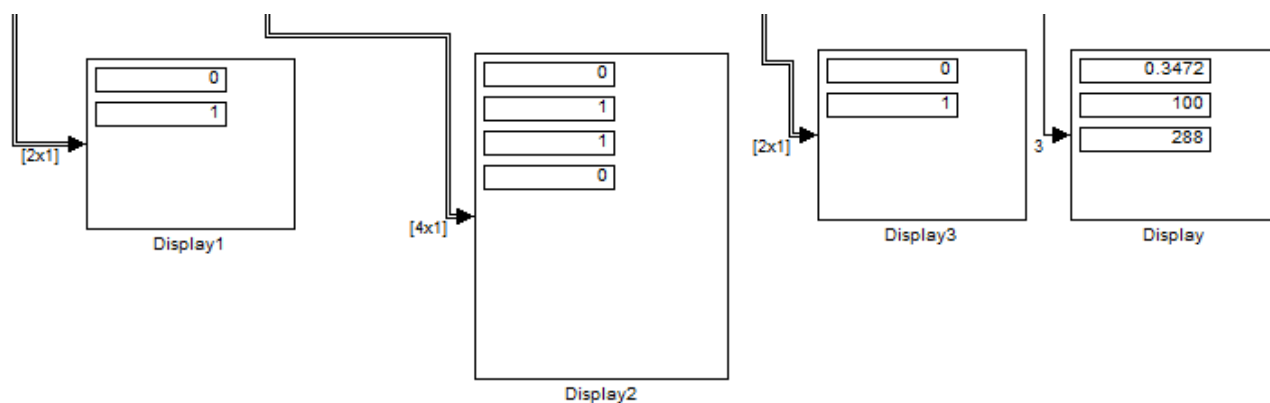


Рис. 3.105. Комбинация на входе, Закодированная последовательность, Комбинация на выходе, Ошибки (вероятность ошибок равна 0,14)

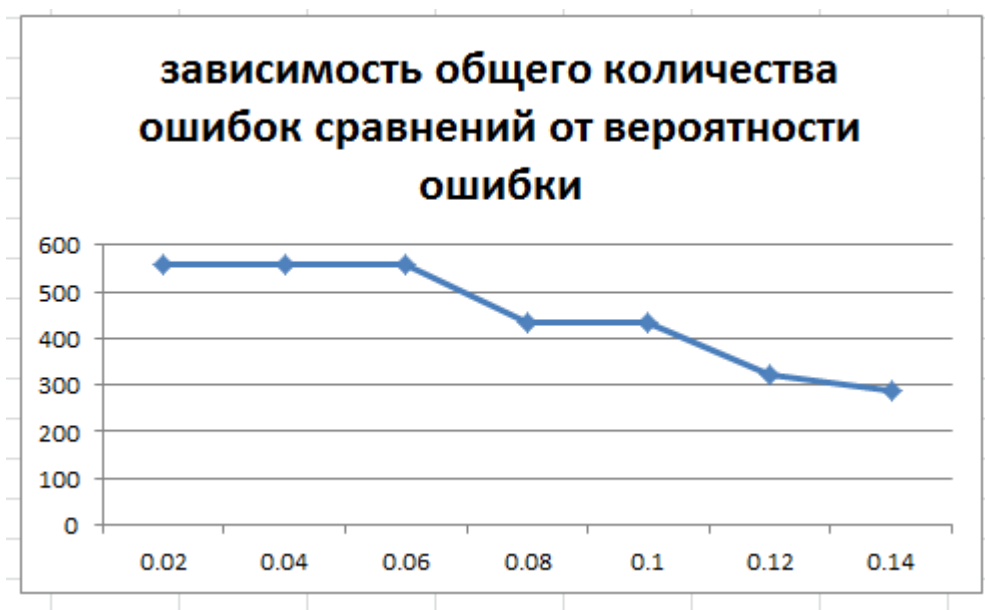


Рис. 3.106. График зависимости общего количества ошибок сравнений от вероятности ошибки для кода (5, [35 31])

Устанавливаем характеристики блоков для кода ([5 4],[27 33 0; 0 5 13])

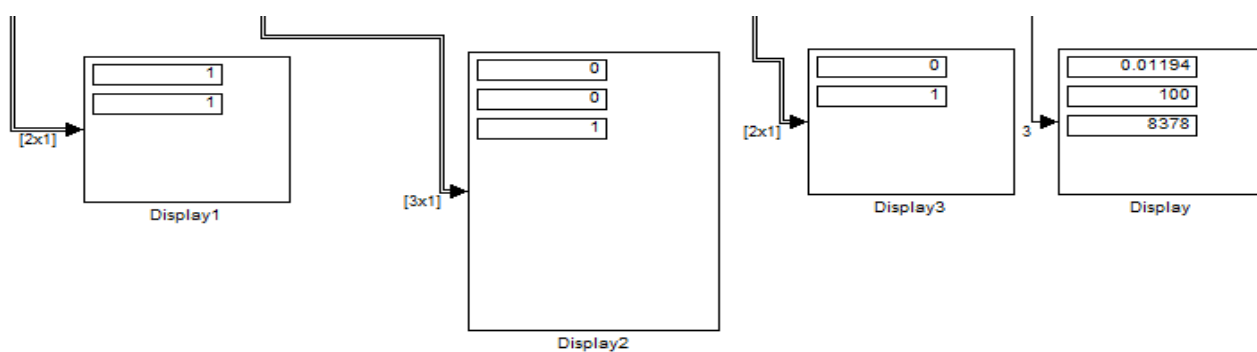


Рис. 3.107. Комбинация на входе, Закодированная последовательность, Комбинация на выходе, Ошибки (вероятность ошибок равна 0,02)

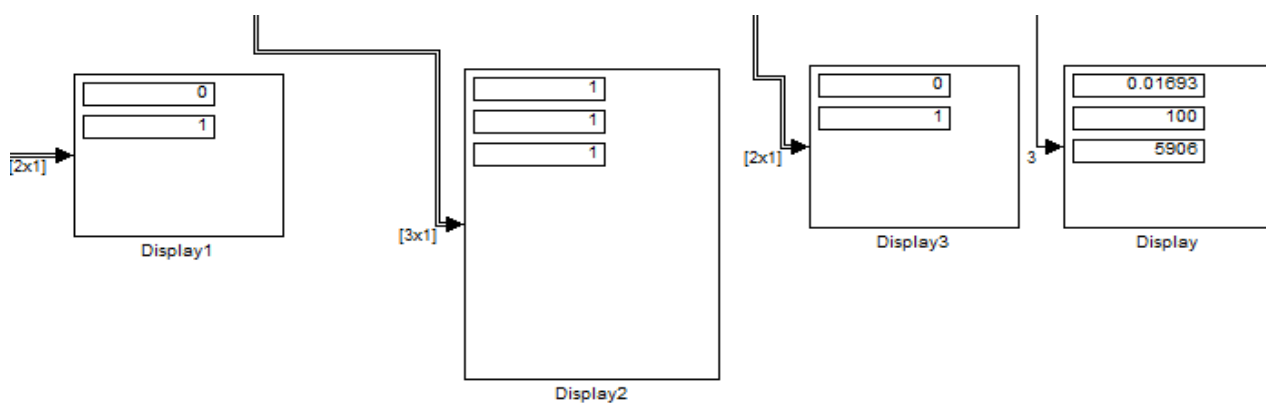


Рис. 3.108. Комбинация на входе, Закодированная последовательность, Комбинация на выходе, Ошибки (вероятность ошибок равна 0,025)

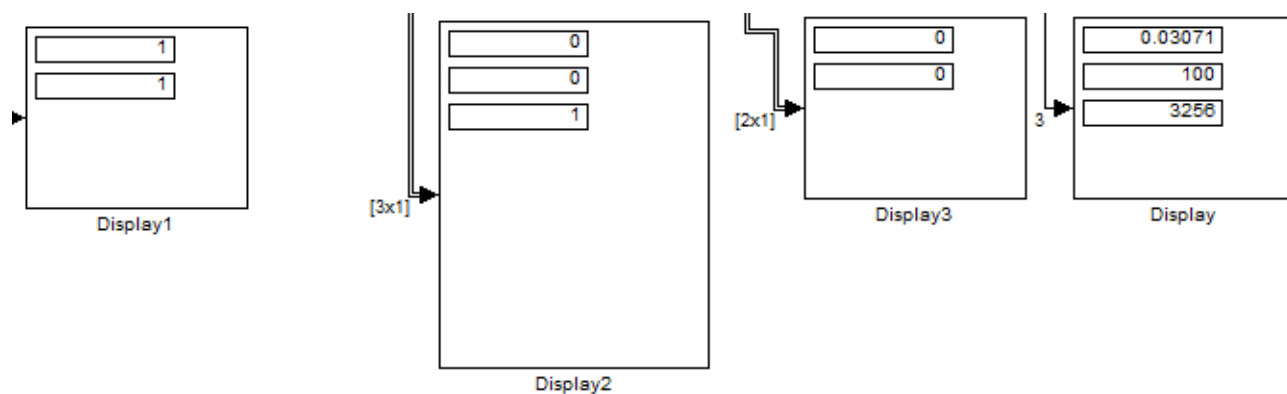


Рис. 3.109. Комбинация на входе, Закодированная последовательность, Комбинация на выходе, Ошибки (вероятность ошибок равна 0,03)

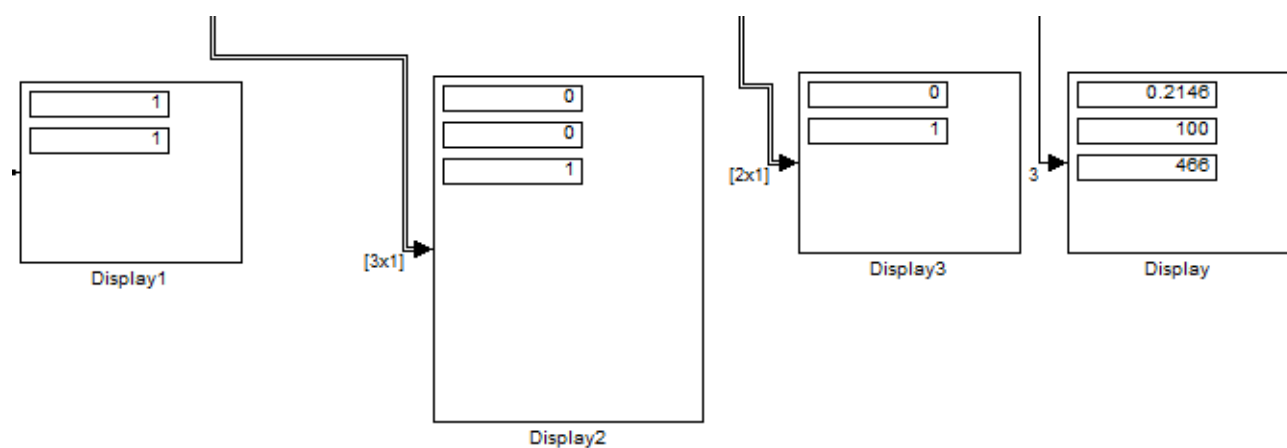


Рис. 3.110. Комбинация на входе, Закодированная последовательность, Комбинация на выходе, Ошибки (вероятность ошибок равна 0,035)

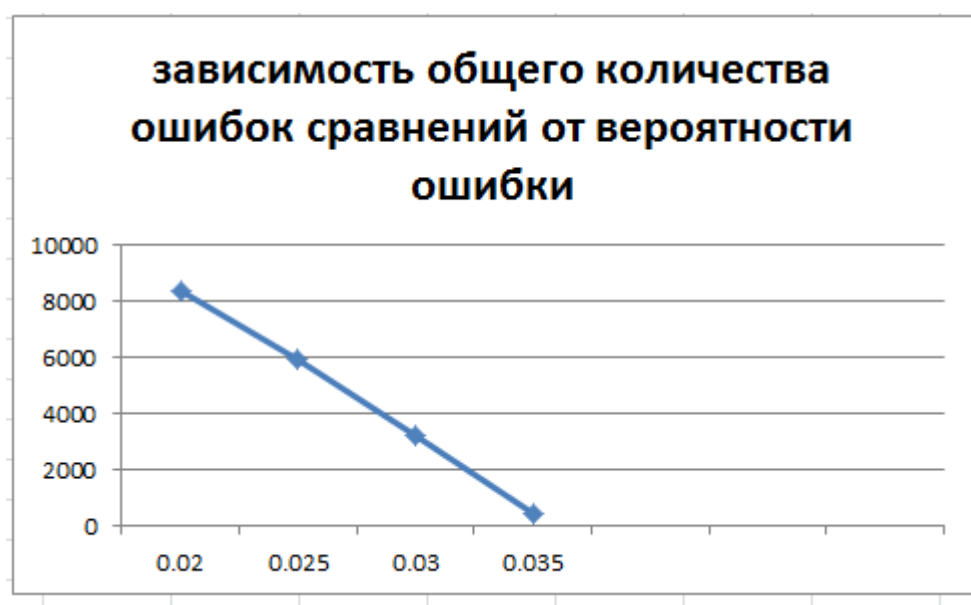


Рис. 3.111. График зависимости числа ошибок от вероятности ошибки для ([5 4], [27 33 0; 0 5 13])

Вопрос 2. Построение схемы декодирования непрерывных кодов в канале с Гауссовскими помехами.

В рабочем поле необходимо собрать схему для работы сверточного кода (7, [171 133]).
Схема представлена на рисунке 3.112

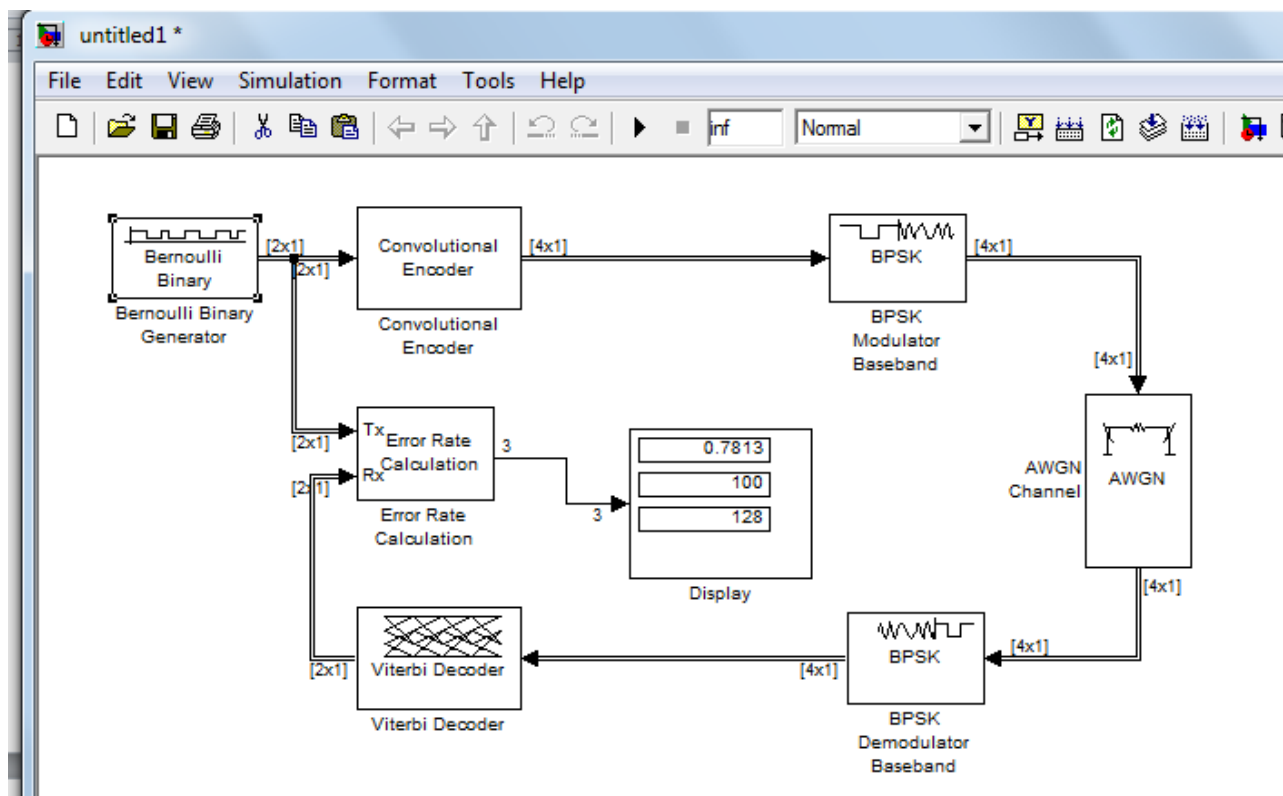


Рис. 3.112. Линия передачи с применением декодера Витерби

В результате практической работы построена схема линии передачи с декодированием Витерби в среде Simulink. Построены графики зависимостей общего количества ошибок сравнений на выходе декодера от вероятности ошибки в канале связи для кодов(5, [35 31]) и ([5 4],[27 33 0; 0 5 13]). Из графиков видно, что число ошибок сравнений уменьшается с ростом вероятности ошибок.

3.5. Турбокодирование. Обобщенная схема турбокодера ТСС с параллельным каскадированием. Сверточные турбокоды. Декодирование турбокодов. Характеристики помехоустойчивости сверточных турбокодов ТСС. Исследование турбокодов с использованием ПО MATLAB.

Принципы построения турбо-кодов

В данном разделе приведены основные теоретические данные одного из важнейших классов итеративно декодируемых кодов – турбо-кодов (turbo codes) и с их обобщением –

многократными турбо-кодов (multiple turbo codes). Турбо-кодирование основано на двух фундаментальных идеях: построение кодов с кодовыми словами, обладающими квазислучайными свойствами и построение декодеров, основанных на легко реализуемых алгоритмах итеративного декодирования [1]. Турбо-коды обладают хорошими характеристиками, особенно при больших длинах кодов и умеренных требованиях к вероятности ошибки на бит.

Параллельное каскадирование двух сверточных кодов

Турбо-код можно определить как параллельное каскадирование (parallel concatenation) двух сверточных кодов. Кодер состоит из двух параллельных систематических кодеров скорости $R=1/2$ и памяти $m=2$ (компонентные кодеры (constituent encoders)) с обратной связью (recursive encoders) имеющих рациональные порождающие подматрицы:

$$G = \left(1 + \frac{1 + D^2}{1 + D + D^2}\right) \quad (3.1)$$

и простого блочного перемежителя (ПБП). Входом турбо-кодера служит двоичная информационная последовательность длины K

$$u = u_0 u_1 \dots u_{K-1} \quad (3.2)$$

Входом первого составного кодера и перемежителя, ассоциированного с вторым составным кодером, служит последовательность u длины $K = K + m$,

$$u = u_0 u_1 \dots u_{K-1} \quad (3.3)$$

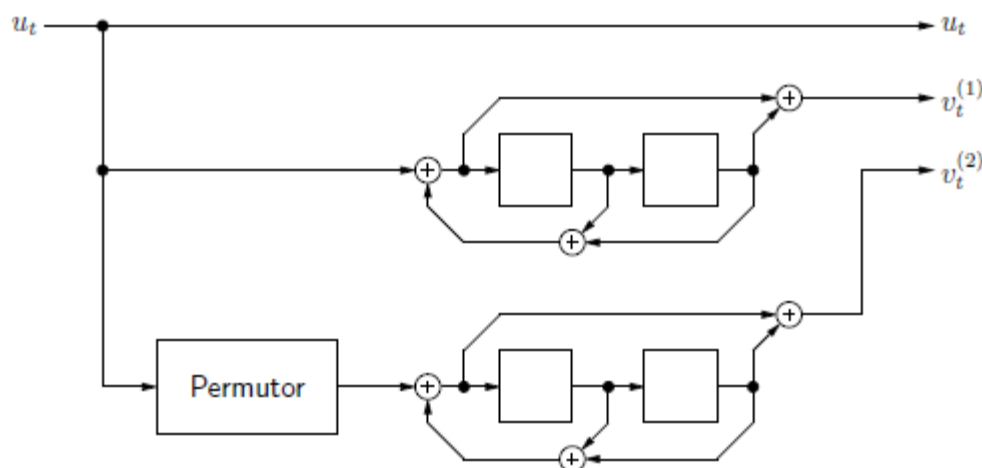


Рис. 3.113. Систематический сверточный кодер скорости $R = 1/3$, основанный на параллельном каскадировании двух сверточных кодеров

где m память составного кодера. Первые K символов этой последовательности – это символы информационной последовательности u , последние m символов образуют хвост,

приводящий первый компонентный кодер в нулевое состояние. Поскольку кодирование систематическое, входная последовательность (включая хвост) имеет вид:

$$v(0) = v_0^{(0)} v_1^{(0)} \dots v_{K-1}^{(0)} = u. \quad (3.4)$$

Первый компонентный кодер генерирует проверочную последовательность:

$$v(1) = v_0^{(1)} v_1^{(1)} \dots v_{K-1}^{(1)}. \quad (3.5)$$

Поскольку хвост заставляет первый компонентный кодер вернуться в нулевое состояние, мы можем использовать БКВР алгоритм апостериорно вероятностного декодирования кодовой последовательности первого компонентного кодера.

Параллельно последовательность u поступает в ПБП, определенный перестановочной матрицей $P = (p_{ik})$, $i = 0, 1, \dots, K-1$, $k = 0, 1, \dots, K-1$ размеров $(K \times K)$. Выходная последовательность $u(2) = uP$ ПБП поступает во второй компонентный кодер, генерирующий проверочную последовательность

$$v(2) = v_0^{(2)} v_1^{(2)} \dots v_{K-1}^{(2)} \quad (3.6)$$

Вообще говоря, второй кодер не переходит в конце кодирования в нулевое состояние. В этом случае нельзя использовать для декодирования БКДР алгоритм, но можно использовать однопутевой алгоритм апостериорно вероятностного декодирования. При больших длинах K это не сильно отражается на характеристиках декодирования и вызывает лишь их небольшую деградацию.

ПБП, использованный во втором кодере, перестановочной матрицей P . С помощью введения вектора перестановки индексов (indexpermutation vector) длины K можно дать альтернативное определение ПБП:

$$\pi = (\pi_0 \pi_1 \dots \pi_{K-1}) \quad (3.7)$$

где $0 \leq \pi_i' \leq K-1$, $i = 0, 1, \dots, K-1$, $\pi' \neq \pi''$ если $i \neq i'$ и π' таково, что $p_{\pi' i} = 1$, то есть $p_{\pi' i}$ – это единственный ненулевой элемент в i -ом столбце перестановочной матрицы P .

Комбинация трех последовательностей, $v^{(0)}$, $v^{(1)}$, и $v^{(2)}$ дает полную передаваемую последовательность (кодовое слово)

$$v = v_0 v_1 \dots v_{K-1} \quad (3.8)$$

где $v_n = v_n^{(0)} v_n^{(1)} v_n^{(2)}$, $n = 0, 1, \dots, K-1$, такая что общая длина блока кода $N = 3K$. Поскольку, как правило $K \gg m$, скорость кода $R \approx 1/3$.

Как было отмечено выше, второй кодер, вообще говоря, не возвращается в нулевое состояние. Важным исключением является турбо-кодер, у которого оба компонентных кодера – это рекурсивные сверточные кодеры памяти $m = 1$ с порождающей матрицей.

$$G = \left(1 + \frac{1}{1+D}\right) \quad (3.9)$$

В этом случае хвост состоит из одного символа u_{k-1} , равному 0, если вес информационной последовательности четный, и равен 1, если вес нечетный. Такая входная последовательность приводит оба компонентных сверточных кодера памяти $m = 1$ в нулевое состояние. Поэтому решетки компонентных сверточных кодеров начинаются и заканчиваются в нулевом состоянии и мы можем использовать БКДР алгоритм для декодирования обоих составных кодов.

Важной особенностью обычных турбо-кодов является слабый рост минимального расстояния d_{\min} с ростом длины кода.

Параллельное каскадирование трех и более сверточных кодов

Обычные турбо-коды имеют относительно плохое минимальное расстояние. При относительно высоких отношениях сигнал-шум это вызывает эффект “зависания” вероятности ошибки декодирования, при котором вероятность ошибки очень медленно убывает с ростом отношения сигнал-шум. Этот участок графика называется “полом” (floor).

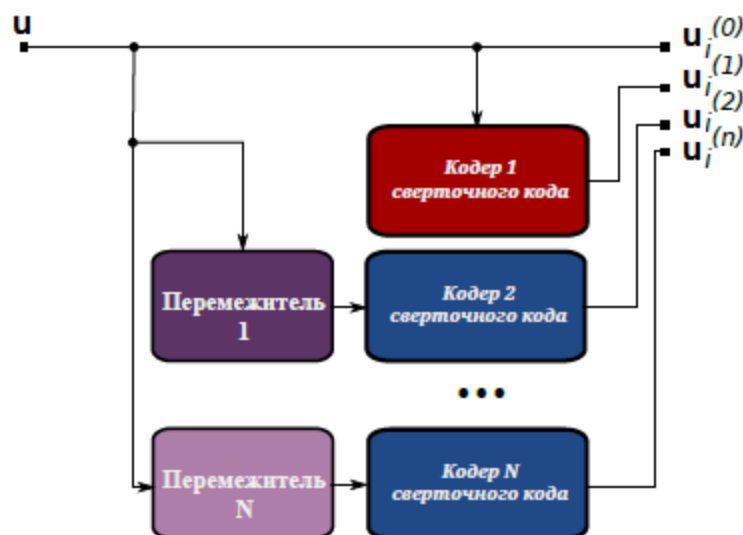


Рис. 3.114. Структура J -размерного сверточного турбокодера

Хотя минимальное расстояние турбо-кода может быть улучшено выбором хорошего перемежителя, оно не может расти быстрее, чем логарифмически с ростом длины блока. Поведение минимальное расстояния МТК существенно лучше.

Кодер J -мерного МТК состоит из J , $J \geq 3$, параллельных компонентных сверточных кодеров и J единичных блоковых перемежителей $P(j)$, $j = 1, 2, \dots, J$, размеров $K \times K$, как показано на рис. 7.2. (Определим формально нулевой перемежитель $P(0)$ как единичную матрицу размеров $K \times K$.) Предполагается, что все компонентные кодеры – это

систематические рекурсивные кодеры скорости $R = 1/2$ и памяти m . Двоичная последовательность длины K

$$u = u_0 u_1 \dots u_{K-1} \quad (3.10)$$

это входная последовательность МТК. Это может быть информационная последовательность; в этом случае все компонентные кодеры – это коды с нейтрализацией хвоста. Аналогично обычным турбо-кодам следует рассмотреть случай, когда u состоит из K информационных символов плюс хвост из m двоичных символов, так что $K = K + m$. В частности, если все компонентные кодеры – это систематические рекурсивные сверточные кодеры памяти 1, хвост состоит из 1 символа.

Последовательность u поступает на входы J ПБП, описываемым перестановочными матрицами $P(j) = (p(j)_{ik})$, где $j = 1, 2, \dots, J$, $i, k = 0, 1, \dots, K - 1$, размеров $K \times K$. За перемежителями следуют компонентные кодеры, такие что j -ая переставленная версия u , то есть $u(j) = uP(j)$, $j = 1, 2, \dots, J$, поступает на вход j -ого компонентного кодера.

Выходная последовательность j -ого компонентного кодера (проверочная последовательность) – это последовательность:

$$v^{(j)} = v_0^{(j)} v_1^{(j)} \dots v_{K-1}^{(j)}, j = 1, 2, \dots, J. \quad (3.11)$$

Входная последовательность МТК,

$$u = v^{(0)} = v_0^{(0)} v_1^{(0)} \dots v_{K-1}^{(0)},$$

перемешивается с выходными последовательностями J компонентных кодеров и передается по каналу. Кодовое слово многократного турбо-кода,

$$v = v_0 v_1 \dots v_{K-1} \quad (3.12)$$

где $v_n = v_n^{(0)} v_n^{(1)} \dots v_n^{(J)}$, $n = 0, 1, \dots, K - 1$, так что длина блока кода равна $N = (J+1)K$ и результирующая скорость равна $R = 1/(J+1)$. Кодовая скорость может быть увеличена выкалыванием символов.

Комбинация тождественного перемежителя и остальных J перемежителей называется $(J + 1)$ -мерным перемежителем. Он может быть представлен в виде $(J + 1)$ -мерной таблицы, сконструированной так, что проекция ненулевого элемента на плоскость, образованную 0-ой и j -ой, $j = 1, 2, \dots, J$, осями давало бы матрицу перестановок $P(j)$, $j = 1, 2, \dots, J$.

Итеративное декодирование турбо-кодов

В данном разделе описано декодирование турбо-кодов. Предположим, что кодовая последовательность $v = v_0 v_1 \dots v_{K-1}$, сконструированная из трех последовательностей, $v^{(0)} = u$, $v^{(1)}$, and $v^{(2)}$, передается по дискретному каналу без памяти. Пусть

$$r = r_0 r_1 \dots r_{K-1} \quad (3.13)$$

где $r_n = r_n^{(0)} r_n^{(1)} r_n^{(2)}$, $n = 0, 1, \dots, K - 1$, означает принятую последовательность и пусть

$$\pi_n^{(0)}(0) = P(u_n = 0) \quad (3.14)$$

означает априорную вероятность, что информационный символ $u_n = 0$, $n = 0, 1, \dots, K-1$.

На практике обычно $P(u_n = 0) = 1/2$.

Итеративный алгоритм декодирования турбо-кодов похож на итеративный алгоритм распространения доверия кодирования МППЧ кодов. Декодер состоит из двух компонентных апостериорно-вероятностных декодеров. Процесс декодирования состоит из I итераций; каждая итерация, в свою очередь, состоит из двух фаз. В результате e -ой фазы, $e = 1, 2$, i -ой итерации $i = 1, 2, \dots, I$, вычисляется последовательность:

$$\pi^{(e)}(i) = \pi_0^{(e)}(i) \pi_1^{(e)}(i) \dots \pi_{K-1}^{(e)}(i) \quad (3.15)$$

где $\pi^{(e)}(i)$ апостериорная вероятность 2, что информационный символ $u_n = 0$, $n = 0, 1, \dots, K-1$. Удобно вместо $\pi^{(e)}(i)$ оперировать с соответствующими логарифмическими отношениями правдоподобия, то есть,

$$z_n^{(e)} = \log \frac{\pi_n^{(e)}}{1 - \pi_n^{(e)}} \quad (3.16)$$

и вместо $\pi^{(e)}(i)$ оперировать с последовательностью:

$$z^{(e)}(i) = z_0^{(e)}(i) z_1^{(e)}(i) \dots z_{K-1}^{(e)}(i) \quad (3.17)$$

Можно использовать обозначение:

$$z_n^{(0)}(0) = \log \frac{P(u_n = 0 | r_n^{(0)})}{P(u_n = 1 | r_n^{(0)})} = \log \frac{\pi_n^{(0)}}{1 - \pi_n^{(0)}} + \log \frac{P(u_n = 0 | r_n^{(0)})}{P(u_n = 1 | r_n^{(0)})} \quad (3.18)$$

для логарифмического отношения апостериорных вероятностей значений символа u_n , при условии, что принят символ r_n . Оно называется внутренней информацией (intrinsic information) об информационном символе $u_n = v^{(0)}(n)$. Предполагается, что оба компонентных декодера хранят в памяти последовательность $z^{(0)}(0) = z_0^{(0)}(0) z_1^{(0)}(0) \dots z_{K-1}^{(0)}(0)$ и используют ее на каждом шаге итеративного процесса декодирования.

Пусть $r^{(0)}$, $r^{(1)}$, and $r^{(2)}$ означают последовательности принятых символов, соответствующих переданной информационной последовательности $v^{(0)} = u$, первой проверочной последовательности $v^{(1)}$, и второй проверочной последовательности $v^{(2)}$, соответственно. Пусть, далее, $r_n^{(0)}$ означает последовательность принятых символов, соответствующих последовательности u за исключением принятого символа $r_n^{(0)}$, то есть,

$$r_n^{(0)} = r_0^{(0)} r_1^{(0)} \dots r_{n-1}^{(0)} r_{n+1}^{(0)} \dots \quad (3.19)$$

Аналогично, пусть $\pi^{(e)}(i)$ означает последовательность апостериорных вероятностей для символов последовательности u за исключением $\pi^{(e)}(i)$ после e -ой фазы i -ой итерации, то есть

$$\pi_n^{(e)}(i) = \pi_0^{(e)}(i)\pi_1^{(e)}(i) \dots \pi_{n-1}^{(e)}(i)\pi_{n+1}^{(e)}(i) \dots \quad (3.20)$$

На первой фазе первой итерации первый компонентный декодер, используя $r(0)n$, $r(1)$ и

$$\pi_n^{(0)}(0) = \pi_0^{(0)}(0)\pi_1^{(0)}(0) \dots \pi_{n-1}^{(0)}(0)\pi_{n+1}^{(0)}(0) \dots \quad (3.21)$$

вычисляет логарифмическое отношение правдоподобия

$$y_n^{(1)}(1) = \log \frac{P(\mathbf{r}_n^{(0)}, \mathbf{r}^{(1)} | r_n^{(0)} \mathbf{u}_n = 0)}{P(\mathbf{r}_n^{(0)}, \mathbf{r}^{(1)} | \mathbf{u}_n = 1)} \quad (3.22)$$

для информационных символов $u_n = v_n^{(0)}$, $n = 0, 1, \dots, K-1$.

Для вычисления $y(1)n(1)$ декодер использует БКВР алгоритм апостериорновероятностного декодирования описанный в параграфе 4.5. Величина $y(1)n(1)$ называется внешней информацией (extrinsic information) первого компонентного декодера об информационном символе u_n при первой итерации. Первый компонентный декодер вычисляет последовательность:

$$z^{(1)}(1) = z_0^{(1)}(1)z_1^{(1)}(1) \dots z_{K-1}^{(1)}(1) \quad (3.23)$$

используя формулу

$$z_n^{(1)}(1) = z_n^{(0)}(0) + y_n^{(1)}(1) \quad (3.24)$$

и посылает ее второму компонентному декодеру.

На второй фазе первой итерации второй компонентный декодер, зная последовательность $z_n^{(1)}(1)$, вычисляет апостериорные вероятности $\pi_n^{(1)}(1)$, $n = 0, 1, \dots, K-1$, и, следовательно, $\pi_n^{(1)}(1)$. Используя $r_n^{(0)}$, $r^{(2)}$, и $\pi_n^{(1)}(1)$, декодер, вычисляет внешнюю информацию $y_n^{(2)}(1)$,

$$y_n^{(2)}(1) = \log \frac{P(\mathbf{r}_n^{(0)}, \mathbf{r}^{(2)} | r_n^{(0)} \mathbf{u}_n = 0)}{P(\mathbf{r}_n^{(0)}, \mathbf{r}^{(2)} | \mathbf{u}_n = 1)} \quad (3.25)$$

Затем декодер вычисляет новое логарифмическое отношение правдоподобия и посылает:

$$z^{(2)}(1) = z_0^{(2)}(1)z_1^{(2)}(1) \dots z_{K-1}^{(2)}(1) \quad (3.26)$$

первому компонентному декодеру. Это завершает первую итерацию.

Рассмотрим теперь i -ую итерацию, $i = 2, 3, \dots, I$. На первой фазе i -ой итерации первый компонентный декодер, зная последовательность логарифмических отношений правдоподобия $z^{(2)}(i-1)$, вычисляет последовательность апостериорных вероятностей $\pi^{(2)}(i-1)$ и, следовательно, $\pi_n^{(2)}(i-1)$. Используя $r_n^{(0)}$, $r^{(1)}$, и $\pi_n^{(2)}(i-1)$, первый компонентный декодер вычисляет $y_n^{(1)}(i)$, $n = 0, 1, \dots, K-1$, и $z_n^{(1)}(i)$, $n = 0, 1, \dots, K-1$. Затем он посылает $z^{(1)}(i)$ второму компонентному декодеру и т.д.. Идея состоит в том, что на каждой фазе итеративного процесса декодирования активный компонентный декодер использует (аппроксимированные) апостериорные вероятности, вычисленные другим

компонентным декодером на предыдущем шаге декодирования в качестве априорных вероятностей и вычисляет новую аппроксимацию апостериорных вероятностей.

Декодер чередует эти фазы декодирования в течение I итераций. Решение u_n , $n = 0, 1, \dots, K - 1$, получается сравнением финального логарифмического отношения $z_n^{(2)}(I)$ с порогом 0:

$$u_n = \begin{cases} 0, & \text{если } z_n^{(2)}(I) \succ 0; \\ 1, & \text{если } z_n^{(2)}(I) \prec 0. \end{cases} \quad (3.27)$$

(Если $z_n^{(i)} = 0$, тогда декодер использует правило бросания монеты и принимает решение $u_n = 0$ или $u_n = 1$ с вероятностью $1/2$.) Итеративный декодер иллюстрируется на рис. 2.3.1, где $u = u_0 u_1 \dots u_{K-1}$. Итеративное декодирование многократных турбо-кодов аналогично итеративному декодированию обычных турбо-кодов. Декодирование многократных турбокодов с J компонентными кодами состоит из J итераций, каждая итерация аналогична итерациям декодирования обычных турбо-кодов.

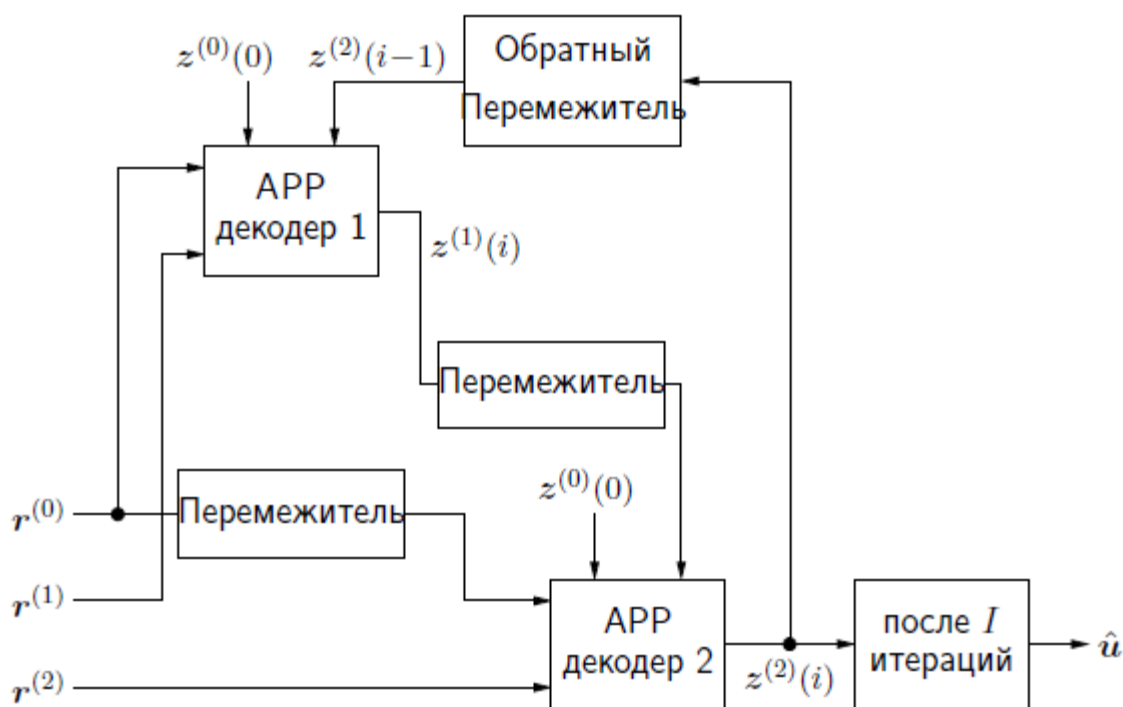


Рис. 3.115. Итеративный декодер для турбо-кодов

Разработка модели системы связи с использованием турбо-кодов

В диалоговом окне приложения Simulink необходимо собрать схему, представленную на рисунке 3.116.

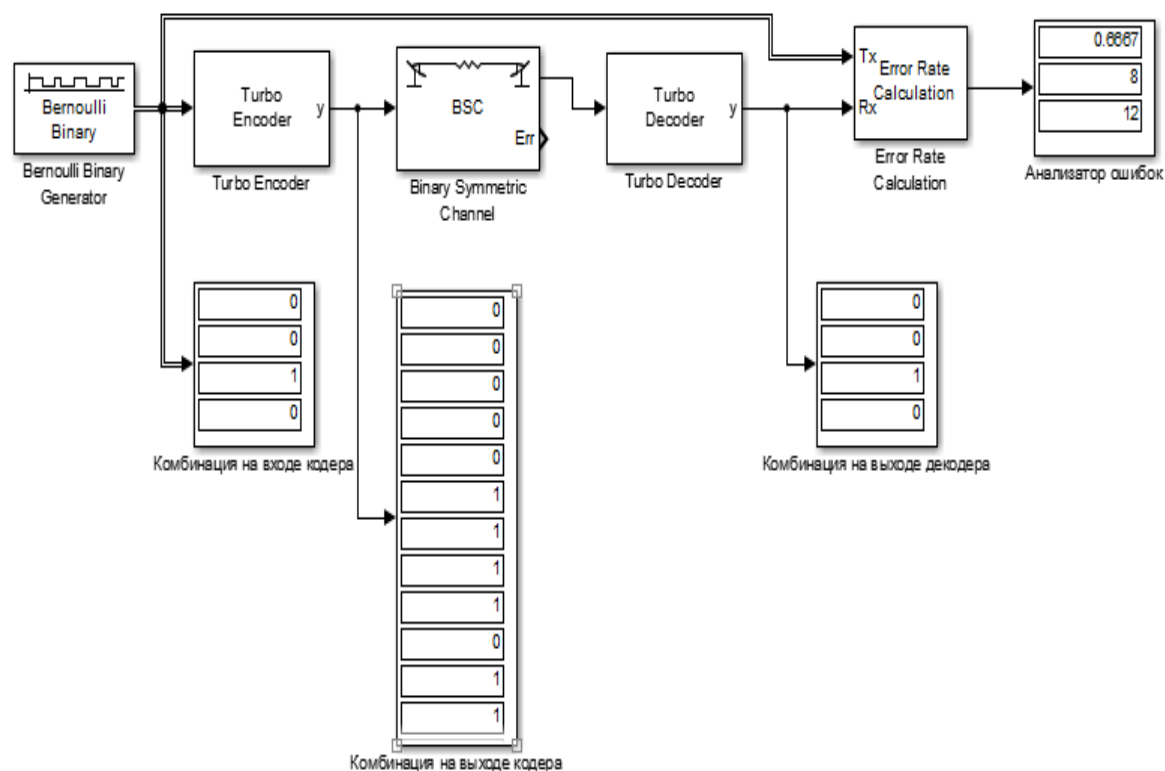


Рис. 3.116. Модель системы связи с использованием турбо-кодов

В состав проектируемой системы связи входят следующие элементы:

1. Bernoulli Binary Generator;
2. Turbo Encoder;
3. Binary Symmetric Channel (канал передачи);
4. Turbo Decoder;
5. Error Rate Calculation (анализатор ошибок);
6. Дисплей, отображающий комбинацию на входе кодера;
7. Дисплей, отображающий комбинацию на выходе кодера;
8. Дисплей, отображающий комбинацию на выходе декодера;
9. Дисплей анализатора ошибок, отображающий частоту ошибок, число обнаруженных ошибок, количество символов по сравнению.

На следующем этапе моделирования необходимо произвести настройку функциональных элементов исследуемой системы связи.

Сначала необходимо выполнить настройку генератора бинарных последовательностей Бернулли (Bernoulli Binary Generator) (рис. 3.110):

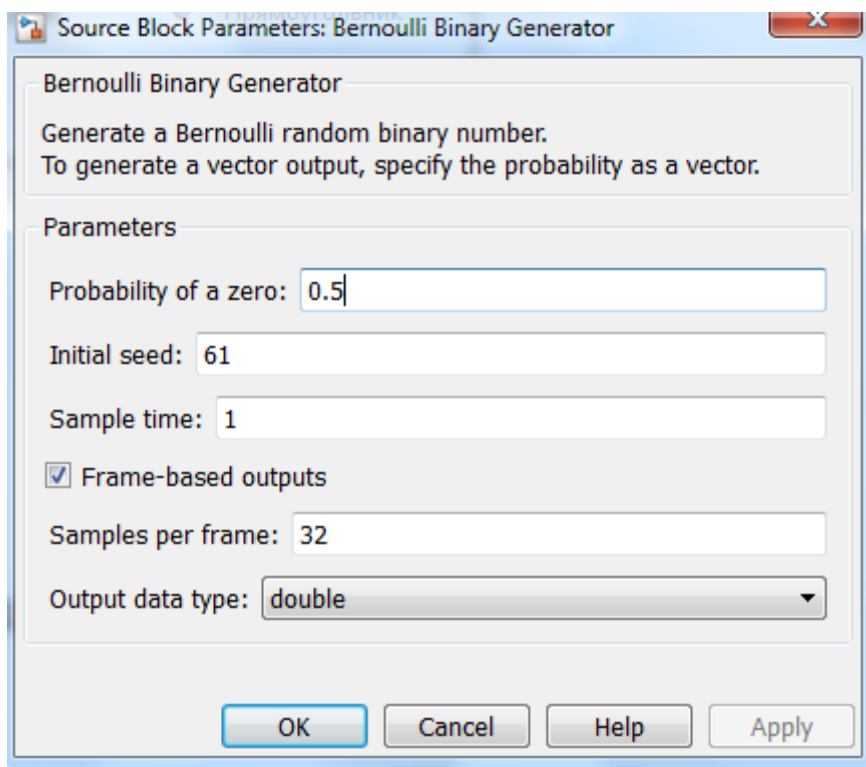


Рисунок 3.110. Настройка Bernoulli Binary Generator для 32 бит

В представленном случае будет использоваться входная последовательность, состоящая из 32 бит

Далее необходимо выполнить настройку турбо-кодера и декодера (рис. 3.111 и рис. 3.112):

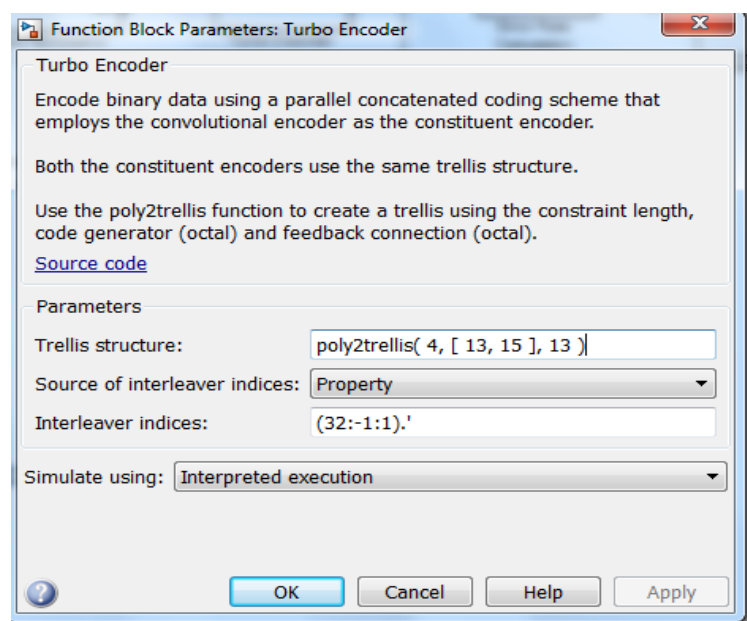


Рис. 3.111. Настройка Turbo Encoder

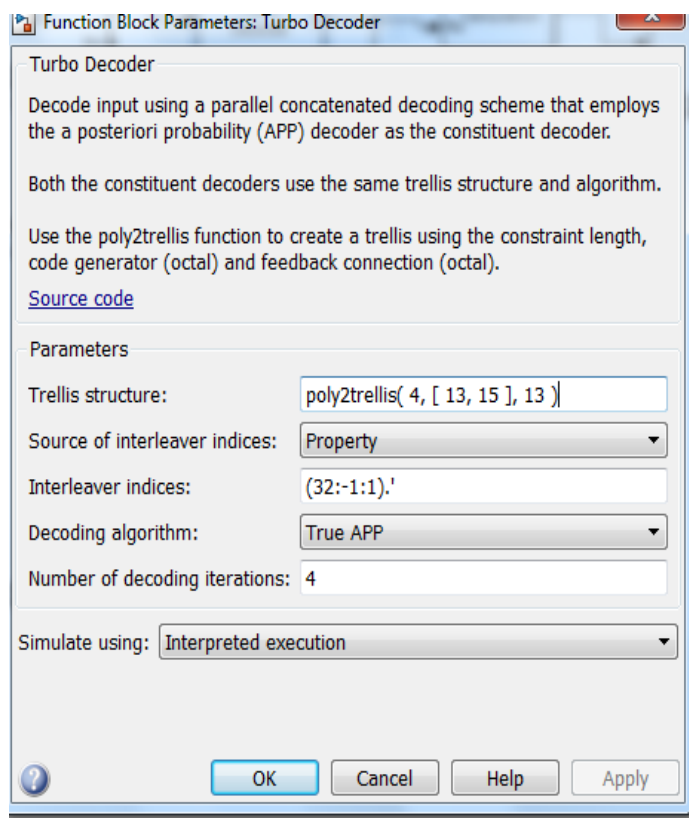


Рис. 3.112. Настройка Turbo Decoder

На следующем шаге необходимо выполнить настройку двоичного канала передачи информации (Binary Symmetric Channel) (рис. 3.113):

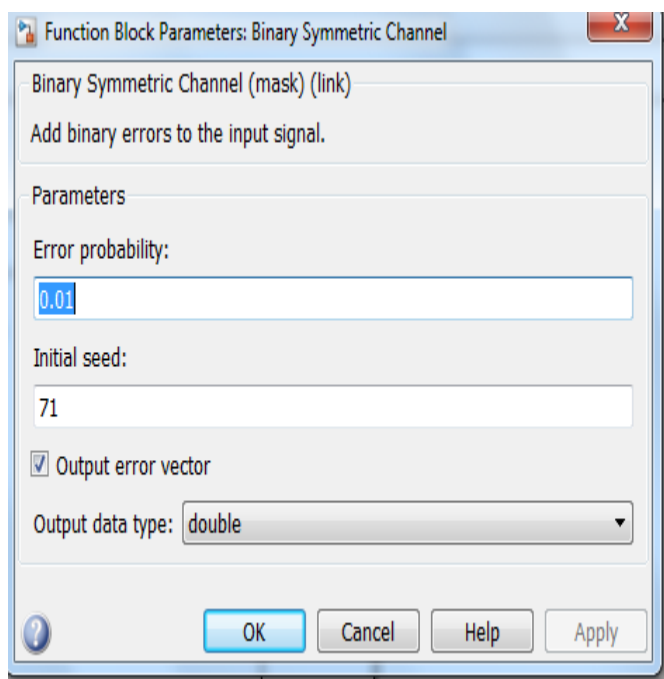


Рис. 3.113. Настройка Binary Symmetric Channel

Стоит отметить, что на данном этапе параметр «вероятность ошибки» необходимо задать малым, для того чтобы проверить работоспособность системы в целом (вероятность ошибки выступает аналогом отношения сигнал/шум в реальном аналоговом канале связи).

Далее настройка анализатора ошибок (Error Rate Calculation) (рис.3.6).

Parameters

Receive delay:
0

Computation delay:
0

Computation mode: Entire frame

Output data: Port

Reset port

Stop simulation

Рис. 3.114. Настройка Error Rate Calculation

Заключительный шаг – запуск работы всей системы. После запуска спроектированной системы будут получены следующие результаты:

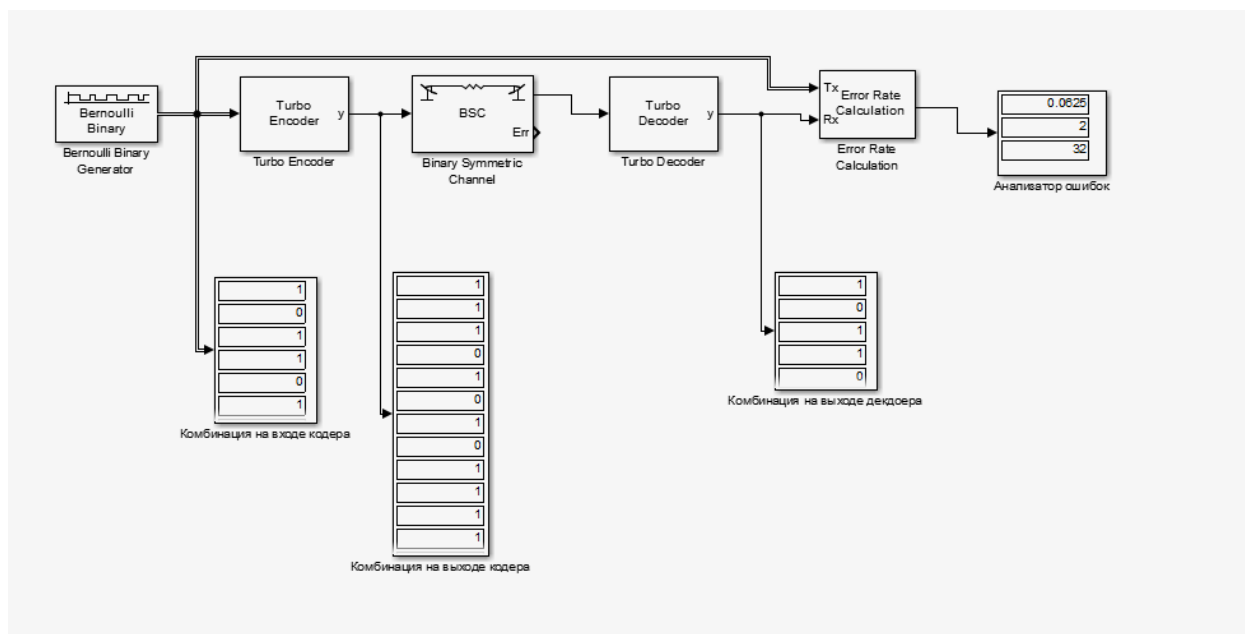


Рис. 3.115. Результаты работы системы связи на основе турбо-кодера

На основе полученных результатов можно сделать вывод, что операции кодирования и декодирования выполнены правильно, следовательно, спроектированная система исправна и работоспособна.

Одной из основных характеристик, характеризующих помехоустойчивость кода, является зависимость количества обнаруженных ошибочных бит от вероятности ошибки. Вероятность ошибки – является аналогом шумов в реальном аналоговом канале передачи. Таким образом для исследования помехоустойчивости турбо-кода представляется целесообразным построить вышеупомянутую зависимость.

Для турбо кода, с входной последовательностью, состоящей из четырех бит и со скоростью кодирования $R=1/3$ зависимость количества ошибок от вероятности ошибок представлена в таблице 3.7:

Таблица 3.7. Зависимость вероятности ошибок от количества ошибочных бит для турбо кода

P	0	0	0	0	0	0,	0,	0,7	0,8	0,9	1
ош	,01	,1	,2	,3	,4	5	6				
N	0	0	0	0	0	0,	0,	0,5	0,6	0,5	0,56
bits	,062	,281	,343	,375	,437	437	406	31	25	62	2

Для удобства обработки полученной информации, представляется необходимым и целесообразным представить таблицу 3.7 в виде графика:

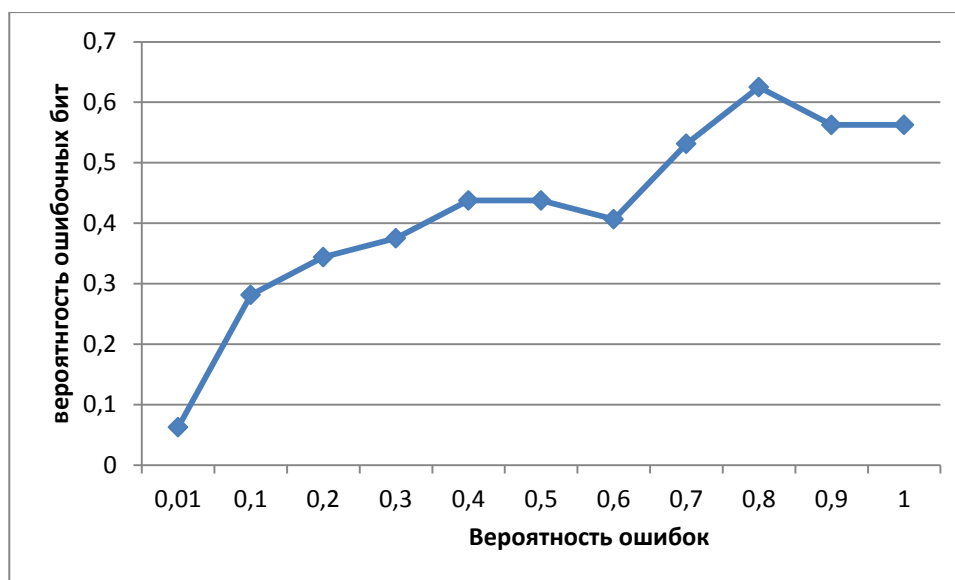


Рис. 3.116. Зависимость вероятности ошибок от количества ошибочных бит

Для турбо кода, с входной последовательностью, состоящей из 64 бит и со скоростью кодирования $R=1/3$ зависимость количества ошибок от вероятности ошибок представлена в таблице 3.8:

Таблица 3.8. Зависимость вероятности ошибок от количества ошибочных бит для турбо кода

P	0	0	0	0	0	0	0	0	0	0	0	1
ош	,01	,1	,15	,2	,3	,4	,5	,6	,7	,8	,9	
N	0	0	0	0	0	0	0	0	0	0	0	0
bits	,250	,375	,390	,468	,5	,5	,5	,51	,54	,57	,56	,54

Для удобства обработки полученной информации, представляется необходимым и целесообразным представить таблицу 3.8 в виде графика:

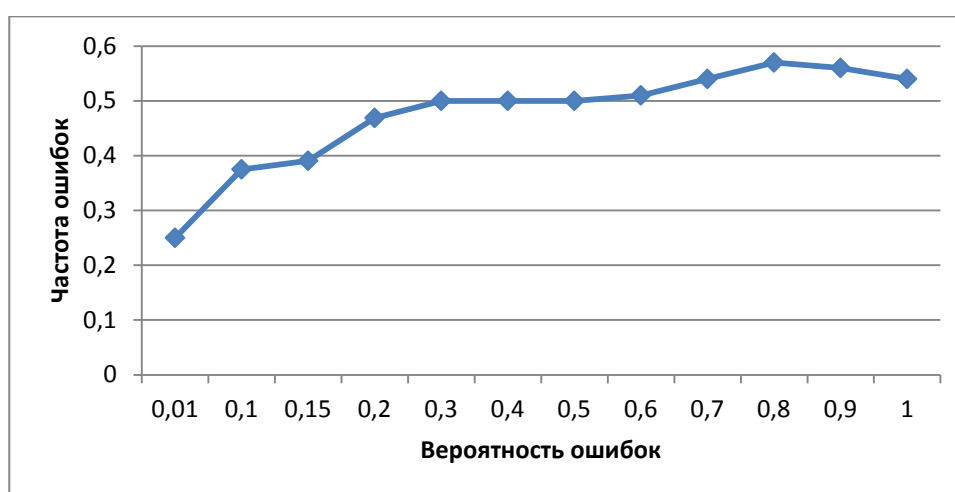


Рис. 3.117. Зависимость вероятности ошибок от количества ошибочных бит

Построим зависимости количества ошибочных битов от вероятности возникновения ошибки (рис. 3.118):

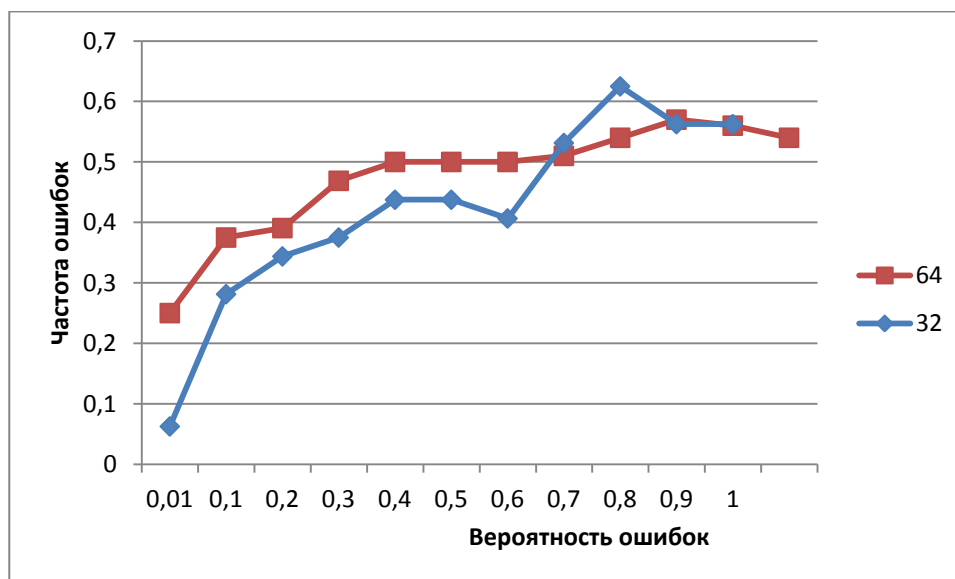


Рис. 3.118. Зависимость количества ошибочных бит

В результате выполнения данной лабораторной работы был проведен теоретический обзор современных алгоритмов использующих для передачи информации по каналам связи турбо-коды.

В ходе работы была разработана модель системы передачи информации с использованием турбо кодов. Кодер представляет собой классический турбо-кодер со скоростью кодирования $R = 1/3$ реализованный путем параллельного каскадирования двух сверточных кодеров. Декодер представляет собой итеративный декодер турбо-кодов. Основным его достоинством его является то, что на каждой итерации с помощью простых процедур декодирования анализируются данные, принадлежащие простым итерационным кодам.

При помощи разработанной модели было установлено, что турбо-коды с большей длиной кодовых комбинацией обладают лучшей помехоустойчивостью, что соответствует теории.

3.6. Низкоплотностные коды. Классификация LDPC-кодов. Методы построения проверочных матриц. Алгоритмы декодирования низкоплотных кодов.

Оценка сложности алгоритмов декодирования на базе MATLAB и LabVIEW

Низкоплотностные коды. Основные положения [16]

В 1948 году Клод Элвуд Шеннон опубликовал свою работу по теории передачи информации. Одним из ключевых результатов работы считается теорема о передаче информации для канала с шумами, которая говорит о возможности свести вероятность ошибки передачи по каналу к минимуму при выборе достаточно большой длины ключевого слова — единицы информации передаваемой по каналу.



Рис. 3.119. Упрощенная схема передачи информации по каналу связи.

При передаче информации её поток разбивается на блоки определённой (чаще всего) длины, которые преобразуются кодером (кодируются) в блоки, называемыми ключевыми словами. Ключевые слова передаются по каналу, возможно с искажениями. На принимающей стороне декодер преобразует ключевые слова в поток информации, исправляя (по возможности) ошибки передачи.

Теорема Шеннона утверждает, что при определённых условиях вероятность ошибки декодирования (то есть невозможность декодером исправить ошибку передачи) можно уменьшить, выбрав большую длину ключевого слова. Однако, данная теорема (и работа вообще) не показывает, как можно выбрать большую длину, а точнее как эффективно организовать процесс кодирования и декодирования информации с большой длиной ключевых слов.

Если предположить, что в кодере и декодере есть некие таблицы соответствия между входным блоком информации и соответствующим кодовым словом, то такие таблицы будут занимать очень много места. Для двоичного симметричного канала без памяти (если говорить упрощённо, то на вход кодера поступает поток из нулей и единиц) количество различных блоков составляет 2^n , где n — количество бит (нулей или единиц) которые будут преобразовываться в одно кодовое слово. Для 8 бит это 256 блоков информации, каждый из которых будет содержать в себе соответствующее кодовое слово. Причём кодовое слово обычно большей длины, так как содержит в себе дополнительные биты для защиты от ошибок передачи данных. Поэтому одним из способов кодирования является использование проверочной матрицы, которые позволяют за одно математическое действие (умножение строки на матрицу) выполнить декодирование кодового слова.

Аналогичным образом каждой проверочной матрице соответствует порождающая матрица, аналогичным способом одной операцией умножения строки на матрицу генерирующей кодовой слово.

Таким образом, для сравнительно коротких кодовых слов кодеры и декодеры могут просто содержать в памяти все возможные варианты, или даже реализовывать их в виде полупроводниковой схемы. Для большего размера кодового слова эффективнее хранить порождающую и проверочную матрицу. Однако, при длинах блоков в несколько тысяч бит хранение матриц размером, соответственно, в мегабиты, уже становится неэффективным.

Одним из способов решения данной проблемы становится использования **кодов с малой плотностью проверок на чётность**, когда в проверяющей матрице количество единиц сравнительно мало, что позволяет эффективнее организовать процесс хранения матрицы или же напрямую реализовать процесс декодирования с помощью полупроводниковой схемы.

Коды с низкой плотностью проверок на чётность (LDPC) — это класс линейных блоковых кодов, позволяющих получить превосходную эффективность с относительно малыми вычислительными затратами на их декодирование. Эти коды были предложены

Робертом Галлагером еще в 1963-ем году, однако были забыты на сорок лет в связи со сложностью реализации алгоритмов их декодирования.

Развитие цифровой техники позволило преодолеть многие проблемы и в конце 20-го столетия исследования в области LDPC – кодов получили новый импульс.

LDPC-коды становятся востребованными в системах передачи информации, требующих максимальной скорости передачи при ограниченной полосе частот. Основным конкурентом LDPC-кодов на данный момент являются турбо-коды, которые нашли свое применение в системах спутниковой связи, ряде стандартов цифрового телевидения и мобильных системах связи третьего поколения. Однако LDPC-коды по сравнению с турбо-кодами имеют ряд преимуществ.

Во-первых, LDPC-коды обгоняют турбо-коды по скорости декодирования.

Во-вторых, LDPC-коды более предпочтительны в каналах с меньшими вероятностями ошибок. С развитием методов передачи информации каналы передачи улучшаются, что дает хорошую перспективу для развития LDPC-кодов.

Применение методов итеративного декодирования к данным кодам позволяет практически вплотную приблизиться к пропускной способности канала при относительно небольшой сложности реализации. В связи с этим во многих новых стандартах передачи различного рода данных (DVB-S2, 802.11n, 802.16e) именно LDPC- коды рекомендованы для исправления ошибок. LDPC - коды представляют собой линейные блочные коды, задаваемые с помощью проверочной матрицы H , характеризуемой относительно малым (<10) числом единиц в строках и столбцах.

В 1981-ом году Р.М. Таннером было предложено использовать двудольные неориентированные графы, впоследствии названные графами Таннера, для описания структуры итеративно декодируемых кодов. В принципе любой блочный код размерности (M,N) , где N – число битов, а M – число проверок в кодовом слове, можно представить в виде двудольного графа Таннера. Например, на рисунке 3.131 изображен такой граф для кода Хэмминга $(7,4)$, проверочная матрица которого имеет вид:

$$H = \begin{pmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}$$

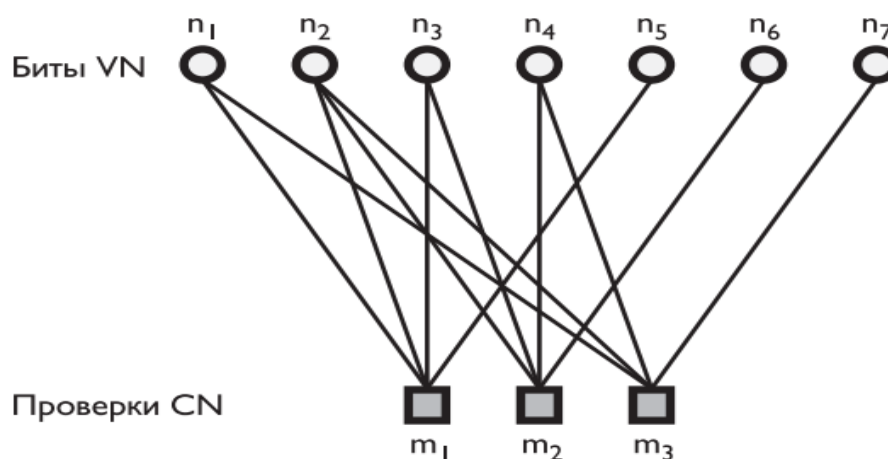


Рис. 3.120. Граф Таннера для кода Хэмминга (7,4)

Вершины графа называются проверочными (check nodes - CN) и битовыми узлами (variable nodes - VN), они обозначены на рисунке m и n соответственно.

При помощи графа Таннера большинство алгоритмов декодирования LDPC-кодов можно представить в виде процессов последовательного обмена сообщениями между соединенными ребрами вершинами. Для проверочных и битовых узлов графа вводится понятие степени – величины, показывающей число рёбер, входящих в рассматриваемый узел. Степени битовых и проверочных узлов обозначаются d_c и d_r соответственно. Если d_c и d_r фиксированы для всех узлов, такой код называют регулярным, а если хотя бы один из этих параметров изменяется от узла к узлу – нерегулярным. Для описания нерегулярных кодов вводится ряд распределения степеней, показывающий долю узлов, имеющих конкретную степень.

В 1996-ом году вышла в свет первая после Р.Галлагера работа, посвященная использованию LDPC – кодов в качестве кодов, способных вплотную приблизиться к границе Шеннона при достаточно большой длине кодового слова. Появление этой статьи породило целую волну исследований, посвященных поиску новых, более эффективных структур LDPC – кодов, а также альтернативных алгоритмов их декодирования с различными соотношениями эффективность/производительность.

В последнее время LDPC – коды получили широкое распространение благодаря превосходной эффективности. Использование LDPC – кодов предусматривает большинство современных стандартов передачи данных (например, стандарты IEEE 802.11, IEEE 802.16), стандартов цифрового вещания (например, стандарты DVB-S2, DVB-T2, DVB-C2).

Классификация LDPC - кодов.

По определению, данному Р. Галлагером, низкоплотный код — это линейный код, проверочная матрица H которого размерности $(M \times N)$ содержит $d_c \ll M$ единиц в каждом столбце и $d_r \ll N$ единиц в каждой строке. Причем распределение единиц по столбцам и строкам в общем случае случайно.

На практике случайное распределение единиц крайне неудобно — для кодирования и декодирования приходится хранить проверочные и генераторные матрицы, что достаточно накладно, особенно при больших длинах кодов.

Очевидным средством борьбы с этой проблемой является переход к низкоплотным кодам, проверочная матрица которых обладает какой-то структурой. Простейший вариант структуризации проверочной матрицы — использование циклических кодов.

Формально проверочная матрица такого кода представляет собой циркулянтную матрицу размерности $N \times N$, в которой каждая строка получается циклическим сдвигом вправо предыдущей строки. Значение влияния цикличности проверочной матрицы на сложность декодера LDPC-кода сложно переоценить, поскольку каждая из строк матрицы однозначно определяется предыдущей строкой, в связи с чем реализация декодера может быть существенно упрощена по сравнению со случайной структурой проверочной матрицы. Как известно, кодер циклического кода достаточно просто реализовать с использованием сдвигового регистра и набора сумматоров.

К недостаткам циклических кодов можно отнести фиксированный, для всех скоростей кодирования, размер проверочной матрицы $N \times N$, что подразумевает более сложный декодер, а также высокий Хэммингов вес строк, что усложняет структуру декодера. Дополнительно стоит заметить, что циклический код — всегда регулярный.

К достоинствам помимо упрощения кодирования/декодирования следует отнести большое минимальное расстояние и очень низкий порог при итеративном декодировании.

Желание преодолеть недостатки циклических LDPC-кодов привело к появлению квазициклических LDPC-кодов. Квазициклические коды также имеют хорошую структуру, позволяющую упростить кодер и декодер. В дополнение к этому они позволяют более гибко подойти к разработке кода, в частности позволяют синтезировать нерегулярные коды. Проверочная матрица такого кода представляет собой не что иное, как набор циркулянтных подматриц:

$$H = \begin{bmatrix} A_{11} & \cdots & A_{1N} \\ \vdots & & \vdots \\ A_{M1} & \cdots & A_{MN} \end{bmatrix}$$

Очевидно, что для получения низкоплотного кода, циркулянтные матрицы должны быть разреженными, что на практике означает использование в качестве циркулянтов единичных матриц. Для того чтобы получить нерегулярный код, какие-то подматрицы просто объявляются нулевыми.

Методы построения проверочных матриц

Методы построения LDPC-кодов также можно разбить на классы. К первому классу относятся все алгоритмические способы и способы, использующие вычислительную технику. А ко второму — способы, основанные на теории графов, математике конечных полей, алгебре и комбинаторике.

При этом стоит заметить, что первый класс методов позволяет получать как случайные, так и структурированные LDPC-коды, в то время как второй нацелен на получение только структурированных LDPC-кодов, хотя бывают и исключения.

В отличие от других линейных блочных кодов, таких как БЧХ или кодов Рида–Соломона, имеющих строгий алгоритм синтеза кодов с заданными параметрами, для LDPC-кодов существует множество способов построения кодов.

Существуют способы построения LDPC-кодов, предложенные Галлагером и МакКеем, о которых ниже и пойдет речь.

Для начала будет рассмотрен метод, предложенный Р. Галлагером.

Р. Галлагер предложил следующий алгоритм построения низкоплотных кодов. Пусть проверочная матрица кода имеет вид:

$$H = \begin{bmatrix} H_1 \\ H_2 \\ \vdots \\ H_{d_c} \end{bmatrix}$$

В ней подматрицы H_a , $a = 1, 2, \dots, d_c$ имеют структуру, которая может быть описана следующим образом.

Для любых двух целых μ и d_r , больших 1, каждая подматрица H_a имеет размерность $(\mu \times \mu \cdot d_r)$, при этом веса строк этой подматрицы - d_r , а столбцов - 1. Подматрица H_1 имеет специфическую форму: для $i=0,1,\dots,\mu-1$ i -ая строка содержит все d_r единиц на позициях с i от d_r до $(i+r) \cdot r - 1$. Очевидно, что результирующая матрица H - регулярная матрица размерности $\mu \cdot d_c \times \mu \cdot d_r$ с весами строк и столбцов d_r и d_c соответственно.

Важной характеристикой матрицы LDPC-кода является отсутствие циклов определенного размера (кратности). Под циклом кратности 4 понимается наличие в двух разных столбцах проверочной матрицы ненулевых элементов на совпадающих позициях. Отсутствие цикла кратности 4 определяется вычислением скалярного произведения столбцов матрицы: если всевозможные скалярные произведения всех столбцов матрицы не превосходят 1, то это означает отсутствие в матрице циклов кратности 4. Цикл кратности 4 является минимально возможным и встречается существенно чаще циклов большей длины (6, 8, 10 и т. д.). Присутствие в матрице LDPC-кода циклов любой кратности свидетельствует о заложенной в структуру матрицы избыточности, не приводящей к улучшению помехоустойчивых свойств кода. Пример циклов кратности 4 приведен на рисунке 3.121

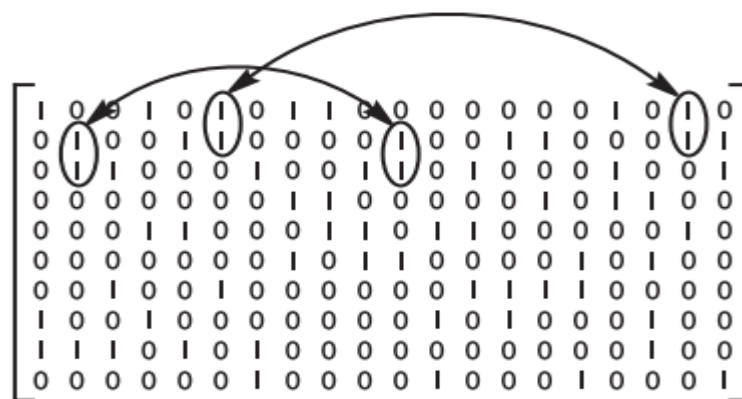


Рис. 3.121. Циклы кратности 4 в матрице LDPC - кода

Рассмотренный выше алгоритм не гарантирует отсутствие циклов кратности 4, однако они могут быть удалены впоследствии. Галлагер показал, что ансамбль таких кодов обладает прекрасными свойствами. Также была показана возможность реализации достаточно простых кодеров, поскольку проверочные биты такого кода могут быть найдены по проверочной матрице кода как функция информационных узлов.

По прошествии тридцати пяти лет МакКей, будучи незнакомым с работой Галлагера, повторно открыл преимущества кодов с разреженными матрицами, и был первым, кто при помощи компьютерного моделирования показал возможность этих кодов вплотную приблизиться к границе Шеннона как для двоичного симметричного канала, так и для канала с аддитивным белым гауссовским шумом. МакКей предложил несколько компьютерных алгоритмов построения проверочных матриц низкоплотностных кодов. Приведем некоторые из них в порядке увеличения сложности реализации.

Проверочная матрица H синтезируется путём случайного генерирования столбцов веса d_c и настолько это возможно, равномерным распределением весов строк d_r .

Проверочная матрица H синтезируется путём случайного генерирования столбцов веса d_c и строк веса d_r , с дополнительной проверкой на отсутствие циклов кратности 4.

Проверочная матрица H синтезируется по алгоритму 2, с дополнительным удалением циклов кратности 4.

Проверочная матрица H синтезируется по алгоритму 3, с дополнительным условием, что проверочная матрица имеет вид $H = [H_1 H_2]$, где H_2 – обратимая матрица.

Недостатком алгоритмов МакКея является отсутствие какой-либо структуры в проверочных матрицах, что усложняет процесс кодирования.

Кодирование осуществляется приведением матрицы H к виду $H = [P^T I]$, из которого можно получить генераторную матрицу в систематической форме $G = [PI]$. Проблема при кодировании по матрице G заключается в том, что подматрица P , в общем случае, не является разреженной. То есть для кодов, представляющих интерес, сложность кодирования оказывается достаточно высокой.

Декодирование LDPC – кодов.

Наибольший интерес для исследователей представляет процедура декодирования, ввиду того что она является более время затратной и ресурсоемкой.

Декодирование – это процедура поиска и исправления ошибки, наложенной каналом на кодовое слово, по принятому из канала вектору или собственно поиск кодового слова по вектору, принятому из канала.

Декодирование по максимуму правдоподобия кода C обозначает нахождение по заданному принятому вектору y такого кодового слова c из C (множества всех кодовых слов), которое максимизирует вероятность того, что передавалось слово c при условии принятия вектора y . Задача декодирования по максимуму правдоподобия является NP-полной.

Для оценки качества работы различных декодеров используется оценка вероятности ошибки декодирования (BER) на информационный бит, вычисляемая как отношение количества ошибочных информационных бит после декодирования к общему количеству переданных информационных бит. итеративные схемы декодирования кодов с низкой плотностью проверок на четность не являются декодерами по максимуму правдоподобия, но

позволяют получить разумный баланс по сложности и вероятности ошибки декодирования по сравнению с декодированием по максимуму правдоподобия. итеративное декодирование подразумевает, что нахождение кодового слова будет производиться не за один проход, а за несколько, с последовательным уточнением результата на каждом шаге. Применяются следующие основные схемы декодирования: «жёсткое» декодирование, быстрое декодирование, многопороговое декодирование.

«Жесткое» декодирование – это схема декодирования для двоичного симметричного канала при небольшом количестве ошибок в канале. «Жесткое» декодирование инвертированием битов – самая простая схема декодирования кодов с низкой плотностью проверок на четность.

Под проверкой понимается любая строка $h = \{h_0, h_1, \dots, h_{N-1}\}$ из проверочной матрицы кода с низкой плотностью проверок на чётность. Будем говорить, что проверка для некоторого вектора $y = \{y_0, y_1, \dots, y_{N-1}\}$ выполняется тогда, когда скалярное произведение вектора y на проверку даёт нуль. Будем говорить, что элемент y_i принятого вектора y участвует в проверке $h = \{h_0, h_1, \dots, h_{N-1}\}$ тогда, когда соответствующий элемент проверки h_i не равен нулю.

Одна итерация «жёсткого» декодирования инвертированием битов производится следующим образом:

1. Для принятого вектора вычисляются все проверки.
2. Если некоторый бит принятого вектора участвовал более чем в половине не выполнившихся проверок, бит инвертируется.
3. После такого анализа всех символов принятого вектора вектор проверяется на принадлежность коду. Если вектор является кодовым словом, декодирование заканчивается, в противном случае выполняется следующая итерация алгоритма.

Такая процедура декодирования применима для кодов с низкой плотностью проверок на четность потому, что большинство проверок в таком случае будут содержать одну ошибку или не будут содержать ошибок вообще и тогда невыполнение большого количества проверок для символа принятого слова будет обозначать наличие в нем ошибки.

Сложность одной итерации «жесткого» декодирования инвертированием бит является линейной, количество итераций декодирования обычно выбирается около $\log_2(N)$, где N – длина кодового слова.

Декодирование по вероятностям является «мягким» декодированием, т.е. декодированием на основе вектора, состоящего не из дискретных значений (0 и 1), а из вещественных величин, полученных на выходе канала путем пересчета вероятностей (англ. belief propagation decoding).

На основе принятого из канала вектора формируются два (для двоичного случая) вектора вероятностей того, что в принятом векторе на данной позиции находился заданный символ.

Каждому ненулевому элементу проверочной матрицы кода с низкой плотностью проверок на чётность приписываются две величины: $q_{i,j}^x$ и $r_{i,j}^x$. Величина $q_{i,j}^x$ является вероятностью того, что j -ый символ принятого вектора имеет значение x по информации, полученной из всех проверок, кроме i -й. Величина $r_{i,j}^x$ является вероятностью того, что проверка i выполняется, если j -ый символ принятого вектора равен x , а все остальные символы проверок имеют распределение вероятностей, заданное величинами $\{q_{i,j}^x: j \text{ из } N(i)/j\}$, где $N(i)$ – множество символов, входящих в i -ую проверку.

Перед началом работы алгоритму требуется инициализация, далее алгоритм работает по принципу пересчета вероятностей символов принятого вектора (belief propagation), используя для пересчета вероятностей правило Байеса для апостериорной вероятности события. Одна итерация алгоритма представляет собой следующую последовательность действий:

1. Для всех проверок вычисляются величины $\Delta r_{i,j}^x$ и пересчитываются вероятности $r_{i,j}^x$ для $x = \{0,1\}$.
2. Для всех символов принятого вектора пересчитываются вероятности $q_{i,j}^x$.
3. Формируются векторы псевдоапостериорной вероятности q_j^0 и q_j^1 .
4. Формируется вектор решения c' по следующему правилу: $c'_j=1$, если $q_j^1 > S$, иначе 0.

Если вектор c' является кодовым словом, декодирование заканчивается, в противном случае выполняется следующая итерация алгоритма.

Сложность данного алгоритма выше, чем сложность «жесткого» декодирования инвертированием битов, но качество декодирования повышается за счет использования дополнительной информации на выходе канала. Однако точность работы такого алгоритма зависит от инициализации: чем точнее она произведена, тем точнее будет конечный результат. Для канала с гауссовским шумом инициализация может быть произведена при

помощи информации о дисперсии шума в канале. Для других распределений шума в канале или при неизвестных характеристиках шума точная инициализация алгоритма может оказаться сложной задачей.

Несмотря на то что декодирование пересчетом вероятностей является эффективным методом для каналов с непрерывным выходом, тот факт, что сложность его значительно выше, чем сложность «жесткого» декодирования, создает предпосылки для поиска более быстрых алгоритмов декодирования, обладающих приемлемым качеством.

Среди известных алгоритмов быстрого декодирования кодов с низкой плотностью проверок на четность для каналов с непрерывным выходом наиболее известен алгоритм «min-sum», являющийся упрощением декодера «belief propagation», а также алгоритм UMP (Uniformly Most Powerful).

Сложность декодера UMP (быстрого декодирования по надежностям) значительно ниже, чем сложность декодера, пересчитывающего вероятности, за счет того, что пересчет надежностей выполняется по упрощенной схеме (схеме «взвешенного» мажоритарного голосования, в качестве «весов» используется надежность проверок), а также за счет возможности использования исключительно целочисленных операций сложения и сложения по модулю два. также к достоинствам быстрого декодера по надежностям можно отнести то, что декодеру не требуется знать характеристики шума в канале (дисперсию и т. д.), следовательно, такой декодер может работать в любом симметричном канале с двоичным входом.

Недостатком быстрого декодера по надежностям является оценка вероятности ошибки декодирования, которая для канала с аддитивным гауссовским шумом оказывается на 0,5 дБ хуже, чем вероятность ошибки декодирования вероятностного декодера.

Далее будет рассмотрено многопороговое декодирование.

Основная идея многопорогового декодирования по надежностям состоит в том, чтобы изменять значения порогов инвертирования символов от одной итерации к другой следующим образом: на первых итерациях порог инвертирования символов выбирается так, чтобы количество инвертированных символов было минимальным (вплоть до инвертирования только одного символа на первой итерации); на последующих итерациях пороги инвертирования постепенно повышаются.

При многопороговом декодировании, если на первой итерации была исправлена хотя бы одна ошибка, декодирование на последующих итерациях становится значительно проще и общее качество декодирования улучшается. По-прежнему для работы декодера не требуется информация о шуме в канале, достаточно лишь задать надежности.

Декодер, работающий по многопороговой схеме, позволяет получить вероятность ошибки декодирования на 0,1–0,4 дБ лучше, чем обеспечивает быстрый декодер по надежности UMP, практически приближаясь к вероятности ошибки, получаемой при вероятностном декодировании кодов с низкой плотностью проверок на четность. Помимо независимости от характеристик канала многопороговый декодер обладает свойством декодеров кодов с низкой плотностью проверок на четность, а именно универсальностью и применимостью для любой конструкции таких кодов.

Следует отметить, что эффективность нерегулярных LDPC-кодов оказывается выше эффективности регулярных кодов. Это объясняется тем, что в нерегулярных кодах из-за различного числа единиц в строках и столбцах информационные символы защищены по-разному. В результате при декодировании проявляется так называемый эффект волны, когда более защищенные биты декодируются быстрее и затем как бы помогают при декодировании менее защищенных бит.

Экспериментальная часть

Задание на лабораторную работу:

1. Изучить работу приложения «LDPC Кодер-Декодер»;
2. Установить параметры матрицы, выставив значения $n=100$, $m=50$, $j=3$, Message Bits = 1000;
3. Меняя значение параметра Noise выписать показания контроллеров BER с использованием кодирования и без него
4. По полученным данным построить соответствующие графики.
5. Привести рисунки, демонстрирующие исходное сообщение, закодированное сообщение, декодированное сообщение и разницу между исходным и декодированным сообщением.

В таблице 3.7 приведены данные, которые были получены в результате эксперимента.

Таблица 3.9. Данные эксперимента

Noise		0	0,0	0,0	0,0	0,1	0,1	0,1	0,1	0,1	0,1
-------	--	---	-----	-----	-----	-----	-----	-----	-----	-----	-----

		,02	4	6	8		2	4	,16	,18
BER(coding)		0,001	0,07064	0,18164	0,30273	0,76832	0,87948	0,25126		
BER(NO_coding)		0,0242	0,565	0,777	0,843	0,129	1	1		

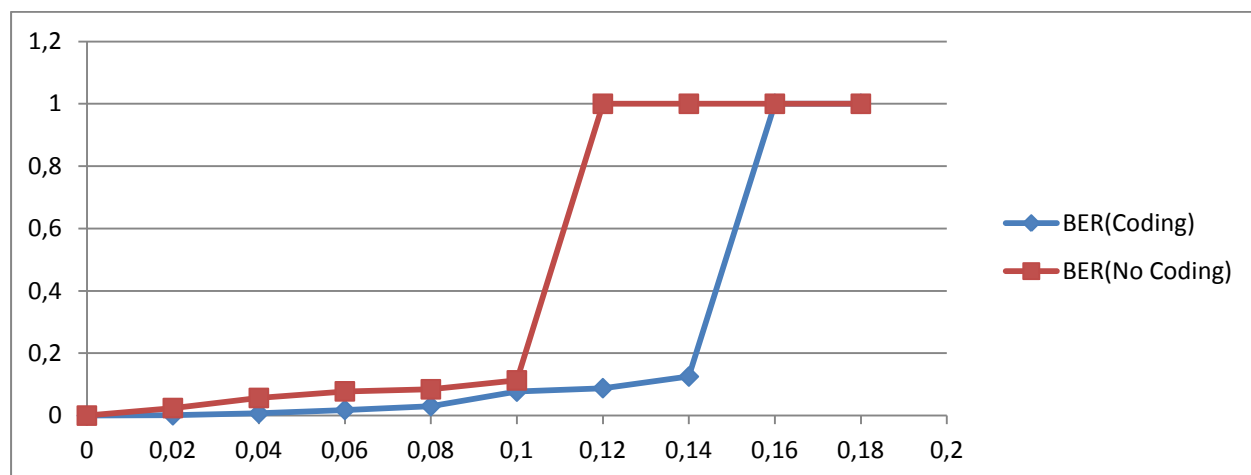


Рис. 3.122. График зависимости BER от Noise

На данном графике видно, что быстрее единицы достиг счётчик, который считал BER в канале без кодирования LDPC.

На рисунках 3.123, 3.124, 3.125 и 3.126 представлены примеры работы LDPC кодера-декодера [17].

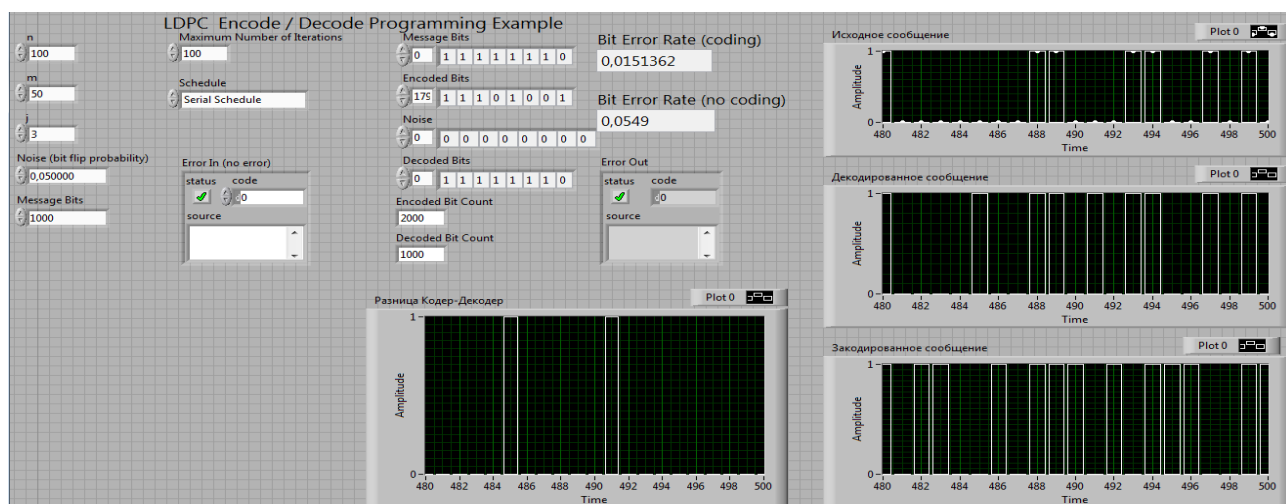


Рис.3.123. Панель управления аппаратно-программного комплекса LDPC кодера-декодера

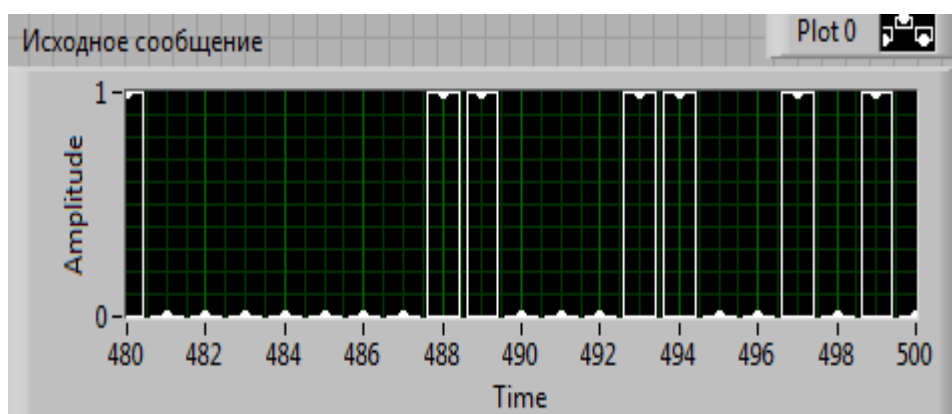


Рис. 3.124. Исходное сообщение

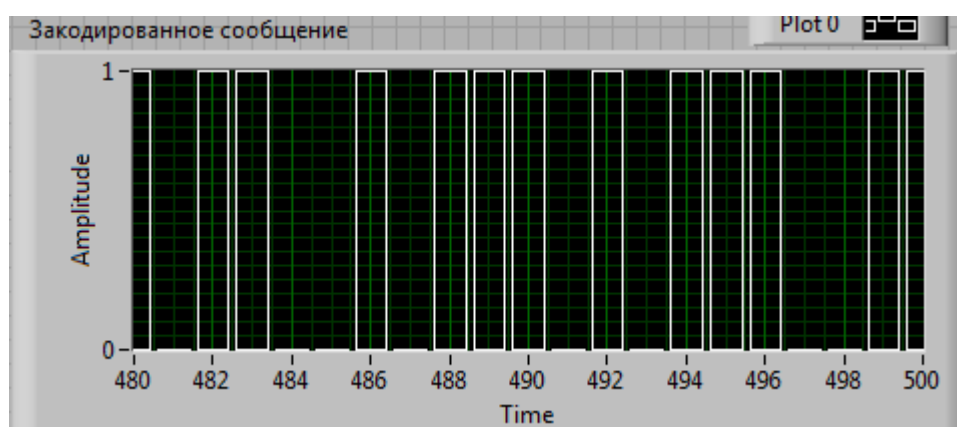


Рис. 3.125. Закодированное сообщение

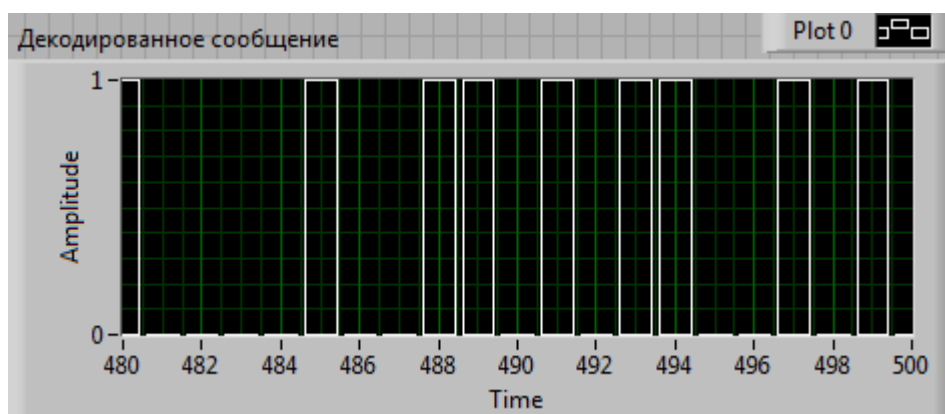


Рис. 3.126. Декодированное сообщение

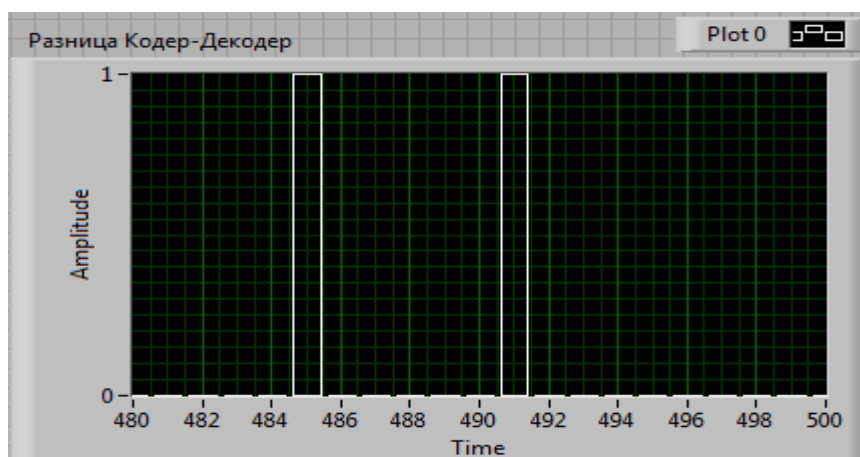


Рис. 3.127. Разница между исходным и декодированным сообщением

В ходе данной лабораторной работы были закреплены навыки работы с ПО LabVIEW, изучен учебный аппаратно-программный комплекс для визуализации и исследования методов канального кодирования/декодирования в беспроводных системах цифрового вещания и связи. Данный комплекс позволяет проводить исследование кодов LDPC. Были сделаны выводы о том, что в канале с кодированием LDPC стопроцентная ошибка возникнет лишь при значении шума, равном 0,16. Тогда как в канале без кодирования стопроцентная ошибка достигается уже при 0,1.

Моделирование LDPC кодов в MATLAB Simulink

В рабочем поле необходимо собрать схему для работы кода LDPC. Схема представлена на рисунке 3.128 [21]

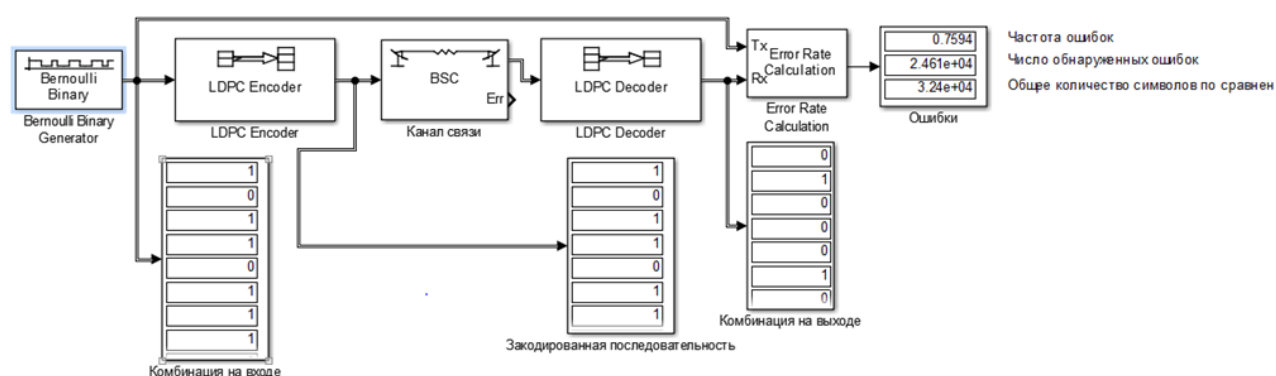


Рис.3.128 Линия передачи с применением код LDPC

В состав линии с кодированием входят:

7. Bernoulli Binary Generator
8. LDPC Encoder
9. Binary Symmetric Channel (канал передачи)

10.LDPC Decoder

11.Error Rate Calculation (анализатор ошибок)

12.Display

Данные блоки необходимо найти в окне «Simulink Library», представленной на рисунке

11. Для упрощения поиска, можно воспользоваться окошком поиска элементов.

Характеристики блоков выставить следующие:

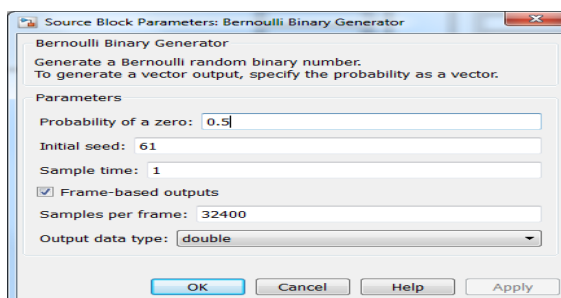


Рис.3.129 Bernoulli Binary Generator

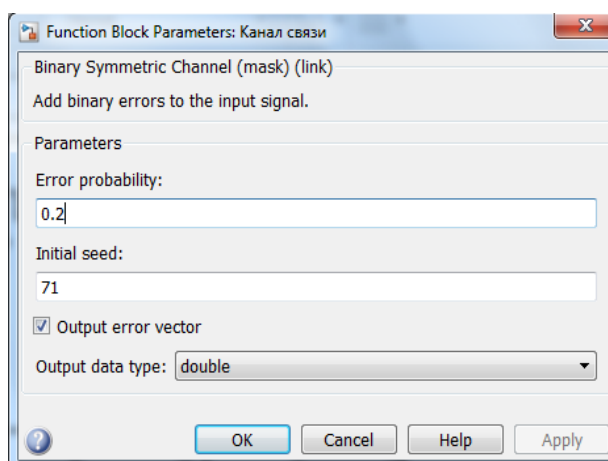


Рис.3.130. LDPC Encoder

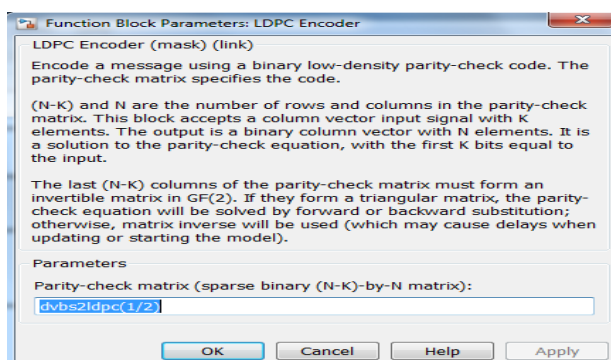


Рис.3.131. Binary Symmetric Channel

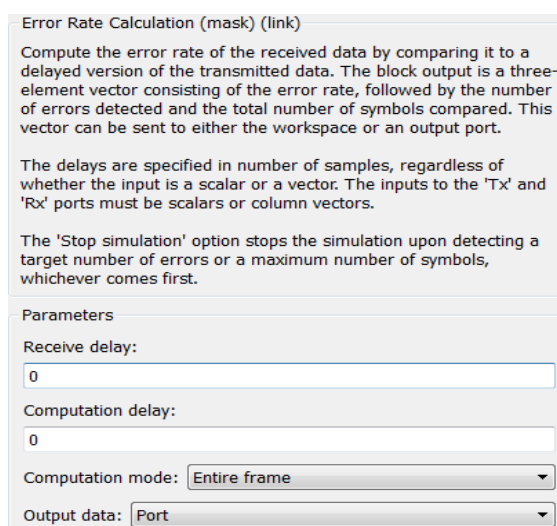


Рис.3.132. BCH Decoder

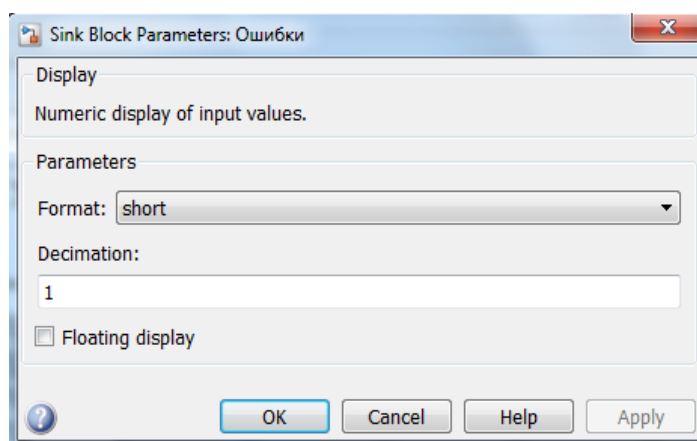


Рис.3.133. Error Rate Calculation (анализатор ошибок)

Представим полученные результаты:

1	0	1	1	0	1	1	1
0	0	0	1	1	1	0	1
0	0	0	0	0	1	0	1
0	1	0	0	0	0	1	1
0	1	0	0	0	0	1	0
0	0	1	0	1	0	1	0
1	0	1	1	0	0	0	0
1	0	0	0	1	1	0	1
1	1	0	1	1	1	0	1
0	1	1	1	0	0	0	1
1	1	1	1	0	0	0	0
0	1	1	0	1	1	0	1
0	1	0	1	1	1	0	1
1	1	0	1	1	0	0	1
0	1	1	0	0	1	0	0
1	1	0	0	1	0	1	0
0	1	1	1	1	1	0	1
1	0	1	1	1	1	1	0
0	0	0	0	0	0	0	1
0	0	1	1	1	1	0	1
1	0	1	1	1	1	1	1
0	0	0	0	0	1	0	1
0	1	0	1	1	1	0	1
0	1	1	1	1	0	1	1
0	1	1	1	0	0	0	1
0	1	1	1	1	0	0	1

Рис. 3.134. Комбинация на входе

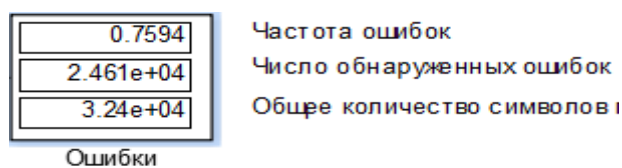


Рис. 3.137. Работа блока оценки ошибки в канале

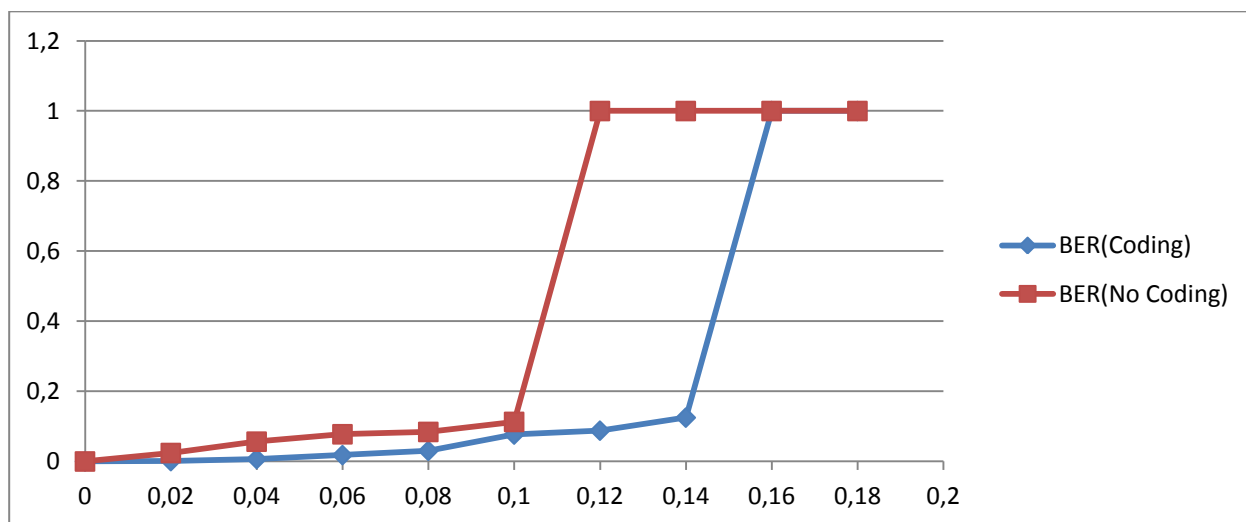


Рис. 3.138 График зависимости BER от Noise

3.7. Исследование каскадных кодов

Каскадными называют коды, в которых кодирование осуществляется в два уровня; имеется внутренний и внешний коды, с помощью которых и достигается желаемая надёжность передачи сообщений. Внутренний код связан с модулятором. Демодулятор, как правило, настраивается для исправления большинства канальных ошибок. Внешний код, чаще всего высокоскоростной (с низкой избыточностью), снижает вероятность появления ошибок до заданного значения. Основной причиной использования каскадного кода является низкая степень кодирования и общая сложность реализации, меньшая той, которая потребовалась бы для осуществления отдельной процедуры кодирования.

В одной из наиболее популярных систем каскадного кодирования для внутреннего кода применяется сверточное кодирование по алгоритму Витерби, а для внешнего — код Рида-Соломона с чередованием между двумя этапами кодирования [1]. Функционирование таких систем при E_b/N_0 , находящемся в пределах от 0,2 до 2,5 дБ, для достижения $P_B = 10^{-5}$ реально достижимо в прикладных задачах. В этой системе демодулятор выдает мягко квантованные кодовые символы на внутренний свёрточный декодер, который, в свою

очередь, выдает жестко квантованные кодовые символы с пакетными ошибками на декодер Рида-Соломина.

Внешний код Рида-Соломона образуется из m -битовых сегментов двоичного потока данных. Производительность такого (недвоичного) кода Рида-Соломона зависит только от числа *символьных ошибок* в блоке. Код не искажается пакетами ошибок внутри m -битового символа. Иными словами, для данной символьной ошибки производительность кода Рида-Соломона такова, как если бы символьная ошибка была вызвана одним битом или m бит. Тем не менее производительность каскадных систем несколько ухудшается за счет коррелирующих ошибок в последовательных символах. Поэтому чередование между кодированиями нужно выполнять на уровне символов (а не битов). Работа [22] представляет собой обзор каскадных кодов, которые были разработаны для дальней космической связи.



Рис. 3.139. Обобщенная структурная схема исследования каскадных кодов

Однако в данной работе исследуется случай, когда в качестве внутреннего кодирования используются турбо-коды.

Турбо-код – параллельный каскадный блочный систематический код, способный исправлять ошибки, возникающие при передаче информации.

Схема каскадного кодирования впервые была предложена Форни [23] как метод получения высокоэффективного кода посредством комбинаций двух или более компонентных кодов (иногда называемых составными). В результате, такие коды могут корректировать

ошибки в значительно более длинных кодах и имеют структуру, которая позволяет относительно легко осуществить декодирование средней сложности. Последовательные каскадные коды часто используются в системах с ограничением мощности, таких как космические зонды. Самая распространенная из этих схем содержит внешний код Рида-Соломона (выполняется первым, убирается последним), который следует за сверточным внутренним кодом (выполняется последним, убирается первым). Турбо-код можно считать обновлением структуры каскадного кодирования с итеративным алгоритмом декодирования связанной кодовой последовательности.

Турбо-коды впервые были введены в 1993 году Берру, Главье и Цитимаджимой. В описываемой схеме достигалась вероятность появления ошибок 10^{-5} при степени кодирования $1/2$ и модуляции BPSK в канале с белым аддитивным гауссовым шумом с E_b/N_0 , равным 0,7 дБ. Коды образуются посредством компоновки двух или более составных кодов, являющихся разными вариантами чередования одной и той же информационной последовательности. Тогда как для сверточных кодов на финальном этапе декодер выдает жестко декодированные биты (или в более общем случае — декодированные символы), в каскадной схеме, такой как турбо-код, для хорошей работы алгоритм декодирования не должен ограничивать себя, подавая на декодеры жесткую схему решений. Для лучшего использования информации, получаемой с каждого декодера, алгоритм декодирования должен применять, в первую очередь, мягкую схему декодирования, вместо жесткой. Для систем с двумя составными кодами концепция, лежащая в основе турбо-декодирования, заключается в том, чтобы передать мягкую схему принятия решений с выхода одного декодера на вход другого и повторять эту процедуру до тех пор, пока не будут получены надежные решения.

Турбокоды представляют собой сравнительно новый тип кодов для исправления ошибок, возникающих при передаче цифровой информации по каналам связи с шумами. Впервые они были введены в рассмотрение в 1993 году и сразу же привлекли к себе пристальное внимание специалистов в области помехоустойчивого кодирования. Причина этому — уникальная способность турбокодов обеспечивать характеристики помехоустойчивости передачи информации, близкие к теоретически достижимым значениям при умеренной сложности реализации кодеков. Разработка турбокодов развивается по двум направлениям: свёрточные турбокоды, образованные путём параллельного соединения двух или более свёрточных кодеров, и блочные турбокоды, образованные путём последовательного соединения двух или более блочных кодеров. Как показали исследования, блочные турбокоды являются более эффективными при относительно высоких кодовых скоростях.

Кодирование

На рисунке 3.140 представлена структурная схема турбо-кодера:

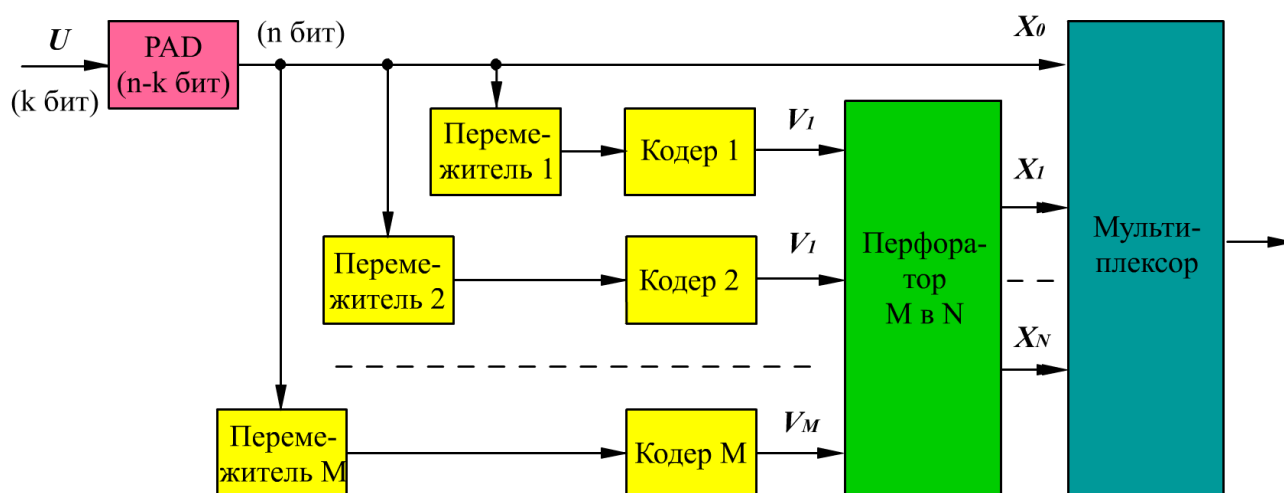


Рис. 3.140. Структурная схема турбо-кодера

Сначала на вход формирователя пакетов (PAD) поступает блок данных U длиной k бит. В формирователе пакетов к данным прибавляется ещё $(n-k)$ дополнительных бит служебной информации, соответствующих используемому стандарту формирования пакета и включающих в себя символы его начала и окончания. То есть получается пакет X_0 , состоящий из n бит.

Далее последовательность бит X_0 поступает параллельно на M ветвей, содержащих последовательно соединённые перемежитель и компонентный кодер. Таким образом X_0 используется в качестве входных данных сразу всеми компонентными кодерами.

В перемежителях по псевдослучайному закону происходит перемешивание поступающих бит. В отличие от посимвольного прямоугольного перемежителя, используемого в кодах Рида-Соломона, в турбо-кодах используется перемежение отдельных бит, которое подобно случайным перестановкам. Причём впоследствии, при операциях декодирования этот закон перемежения будет считаться известным. Полученные последовательности поступают на входы кодеров.

Задача перемежителя — преобразовать входную последовательность так, чтобы комбинации бит X_0 , соответствующие кодовым словам с низким весом (весом называется число ненулевых бит кодового слова) на выходе первого кодера, были преобразованы в комбинации, дающие кодовые слова с высоким весом на выходах остальных кодеров. Таким образом кодеры получают на выходе кодовые слова с различными весами. При кодировании

формируются кодовые слова так, чтобы получалось максимально возможное среднее расстояние между ними (расстоянием между двумя кодовыми словами называется число бит, в которых они различаются). Из-за того, что кодовые блоки формируются из почти независимых частей, на выходе турбо-кодера среднее расстояние между кодовыми словами больше, чем минимальное расстояние для каждого компонентного кодера, а, следовательно, растёт эффективность кодирования.

Кодовая скорость — отношение длины кодового блока на входе к длине преобразованного кодового блока на выходе кодера.

В отсутствие перфоратора исходная последовательность X_0 мультиплексируется с последовательностями проверочных бит, образуя кодовое слово, подлежащее передаче по каналу.

Для увеличения кодовой скорости применяется выкалывание (перфорация) определённых проверочных битов выходной последовательности. Таким образом кодовая скорость возрастает.

Если учесть, что турбо-коды оперируют с блоками большой длины с $k > 10000$, то $k \approx n$.

С помощью перфоратора, выкалывая разное число проверочных бит, возможно регулирование кодовой скорости. То есть можно построить кодер, адаптирующийся к каналу связи. При сильном зашумлении канала перфоратор выкалывает меньше бит, чем вызывает уменьшение кодовой скорости и рост помехоустойчивости кодера. Если же канал связи хорошего качества, то выкалывать можно большое число бит, вызывая рост скорости передачи информации.

Алгоритм декодирования по максимуму апостериорной вероятности

Алгоритм Бала даёт «мягкую» оценку достоверности декодированного бита. То есть предьявляет на выходе степень доверия результату декодирования. В противоположность «жёсткой» структуре, при которой на выходе декодера формируется лишь наиболее вероятное значение декодированного бита («0» или «1»), при вынесении «мягкого» решения используется более подробная дискретизация выходного сигнала, характеризующая вероятность корректного приема бита. Благодаря использованию «мягких» решений в турбо-декодерах оказывается эффективным использование нескольких итераций декодирования. Апостериорная информация, полученная о кодовом слове на выходе первой итерации декодирования, поступает на вход блока следующей итерации и является для него уже априорной вероятностью. Такой подход позволяет улучшать качество декодирования от итерации к итерации. Таким образом, изменяя число итераций декодирования, можно

адаптировать декодер к текущему состоянию канала передачи и достичь требуемой вероятности ошибки на бит.

На рисунке 3.141 представлена структурная схема одной итерации турбо-декодера при двухкаскадном кодировании:

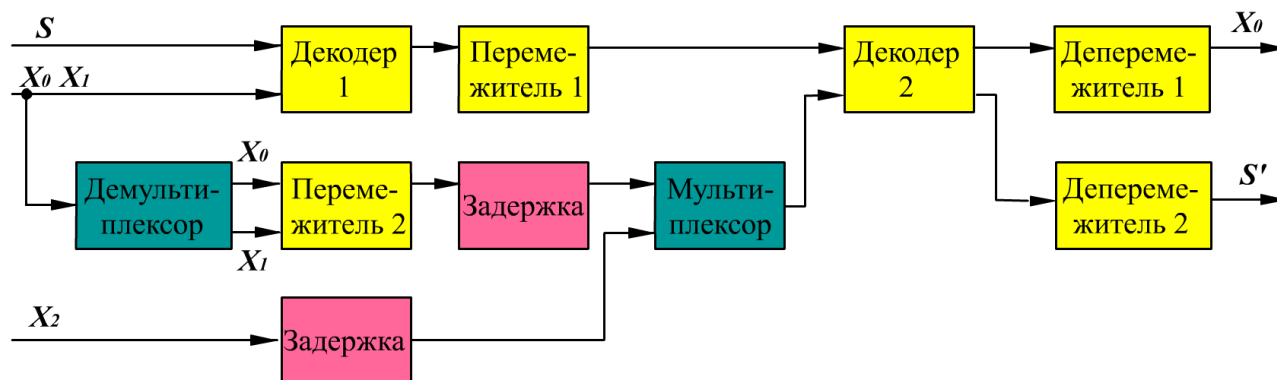


Рисунок 3.141. Структурная схема одной итерации турбо-декодера при двухкаскадном кодировании

На рис. 3.141 для простоты понимания представлен вариант схемы одной итерации турбо-декодирования при двухкаскадном кодировании. Эта схема несложно обобщается на случай произвольного количества каскадов кодирования.

Декодер для одной итерации содержит каскадное соединение двух элементарных декодеров, каждый из которых, основываясь на критерии максимума апостериорной вероятности, выносит «мягкое» решение о переданном символе. На первый декодер первой итерации с выхода демодулятора поступают «мягкие» решения символов последовательностей X_0 и X_1 . Таким образом, на выходе первого декодера появляется оценка информационного символа, которая после последующего перемежения попадает на вход второго декодера и является для него априорной информацией. Используя «мягкое» решение о последовательности X_2 , второй декодер формирует свою оценку.

На рисунке 3.142 представлена структурная схема трёх итераций турбо-декодера при двухкаскадном кодировании:

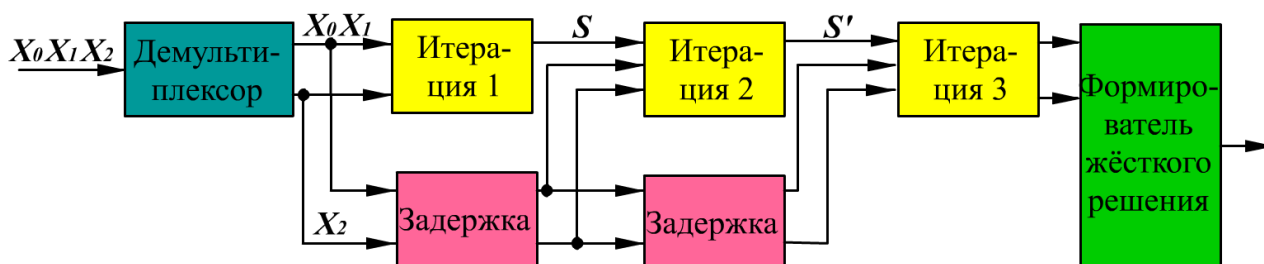


Рис. 3.142. Структурная схема трёх итерации турбо-декодера при двухкаскадном кодировании

С выхода каждой итерации решение переходит на вход следующей. От итерации к итерации происходит уточнение решения. При этом каждая итерация работает с «мягкими» оценками и на выход отдает также «мягкие». Поэтому такие схемы получили название декодеров с мягким входом и мягким выходом. Процесс декодирования прекращается либо после выполнения всех итераций, либо когда вероятность ошибки на бит достигнет требуемого значения. После декодирования из полученного «мягкого» решения производится окончательное «жёсткое».

Преимущества и недостатки турбо-кодов

Преимущества

Среди всех практически используемых современных методов коррекции ошибок турбо-коды и коды с низкой плотностью проверок на чётность наиболее близко подходят к границе Шеннона, теоретическому пределу максимальной пропускной способности зашумленного канала. Турбо-коды позволяют увеличить скорость передачи информации, не требуя увеличения мощности передатчика, или они могут быть использованы для уменьшения требуемой мощности при передаче с заданной скоростью. Важным преимуществом турбо-кодов является независимость сложности декодирования от длины информационного блока, что позволяет снизить вероятность ошибки декодирования путём увеличения его длины.

Недостатки

Основной недостаток турбо-кодов — это относительно высокая сложность декодирования и большая задержка, которые делают их неудобными для некоторых применений. Но, например, для использования в спутниковых каналах этот недостаток не является определяющим, так как длина канала связи сама по себе вносит задержку, вызванную конечностью скорости света.

Ещё один важный недостаток турбо-кодов — сравнительно небольшое кодовое расстояние (то есть минимальное расстояние между двумя кодовыми словами в смысле выбранной метрики). Это приводит к тому, что, хотя при большой входной вероятности ошибки (то есть в плохом канале) эффективность турбо-кода высока, при малой входной вероятности ошибки эффективность турбо-кода крайне ограничена. Поэтому в хороших каналах для дальнейшего уменьшения вероятности ошибки применяют не турбо-коды, а LDPC-коды. Хотя сложность используемых алгоритмов турбо-кодирования и недостаток открытого программного обеспечения препятствуют внедрению турбо-кодов, в настоящее время многие современные системы используют турбо-коды.

Применение турбо-кодов

Компании France Telecom и Telediffusion de France запатентовали широкий класс турбо-кодов, что ограничивает возможность их свободного применения и, в то же время, стимулирует развитие новых методов кодирования таких, как, например, LDPC.

Турбо-коды активно применяются в системах спутниковой и мобильной связи, беспроводного широкополосного доступа и цифрового телевидения. Турбо-коды утверждены в стандарте спутниковой связи DVB-RCS. Турбо-коды также нашли широкое применение в мобильных системах связи третьего поколения (стандарты CDMA2000 и UMTS).

Моделирование каскадных кодов в MATLAB Simulink [21]

Виртуальная модель передачи данных с исправлением ошибок при помощи каскадного кода была реализована в среде Simulink Matlab. Модель демонстрирует работу кодера Рида-Соломона (внешний код) и кодера Турбо-кодов (внутренний код), позволяет исследовать исправляющую способность кодов для разных видов модуляции и сравнить её характеристики с работой указанных выше кодеров в отдельности.

На рисунке 3.143 приведена разработанная модель:

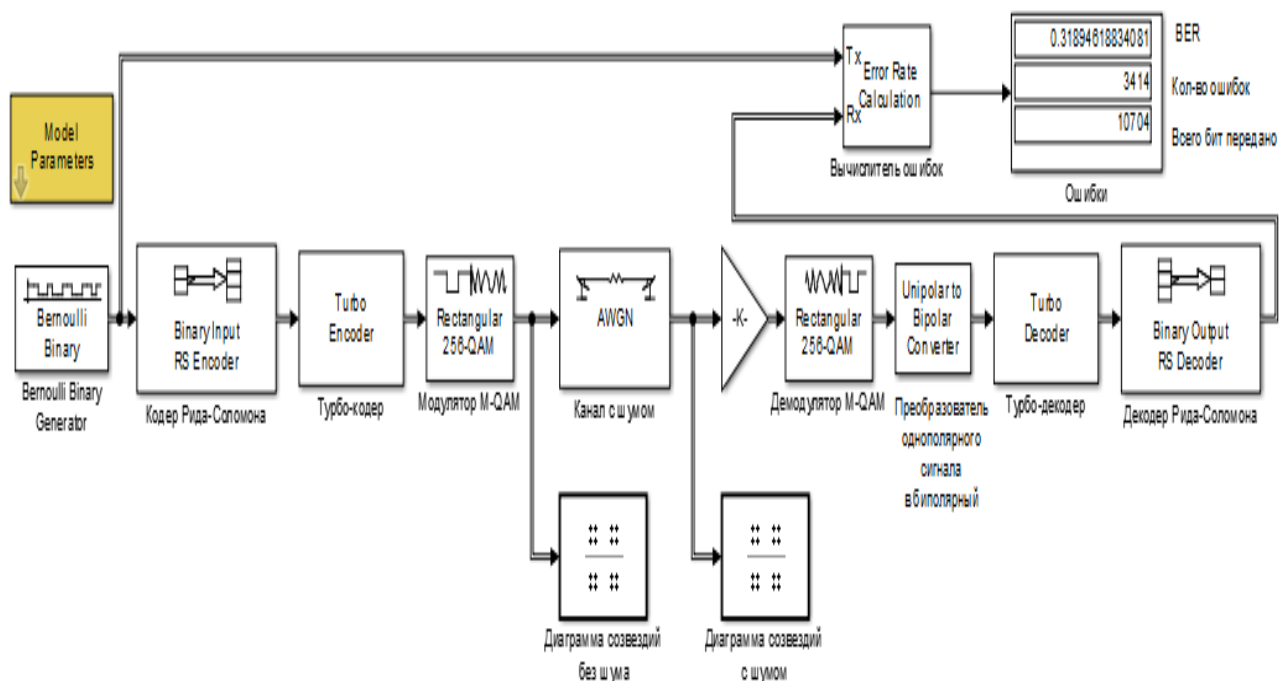


Рис. 3.143. Разработанная модель исследования каскадных кодов

В её основу положены следующие элементы, встроенные в библиотеку Simulink:

- Bernoulli Binary Generator
- Binary Input RS Encoder
- Turbo Encoder
- Rectangular QAM Modulator Baseband
- AWGN Channel
- Rectangular QAM Demodulator Baseband
- Unipolar to Bipolar Converter
- Turbo Decoder
- Binary Output RS Decoder
- Error Rate Calculation
- Discrete Time Scatter Plot Scope
- Gain
- Display (Дисплей, отражающий ошибки)

Далее представлено описание основных блоков:

Bernoulli Binary Generator (генератор псевдослучайной последовательности) – генерирует случайную бинарную последовательность (рисунок 3.144).

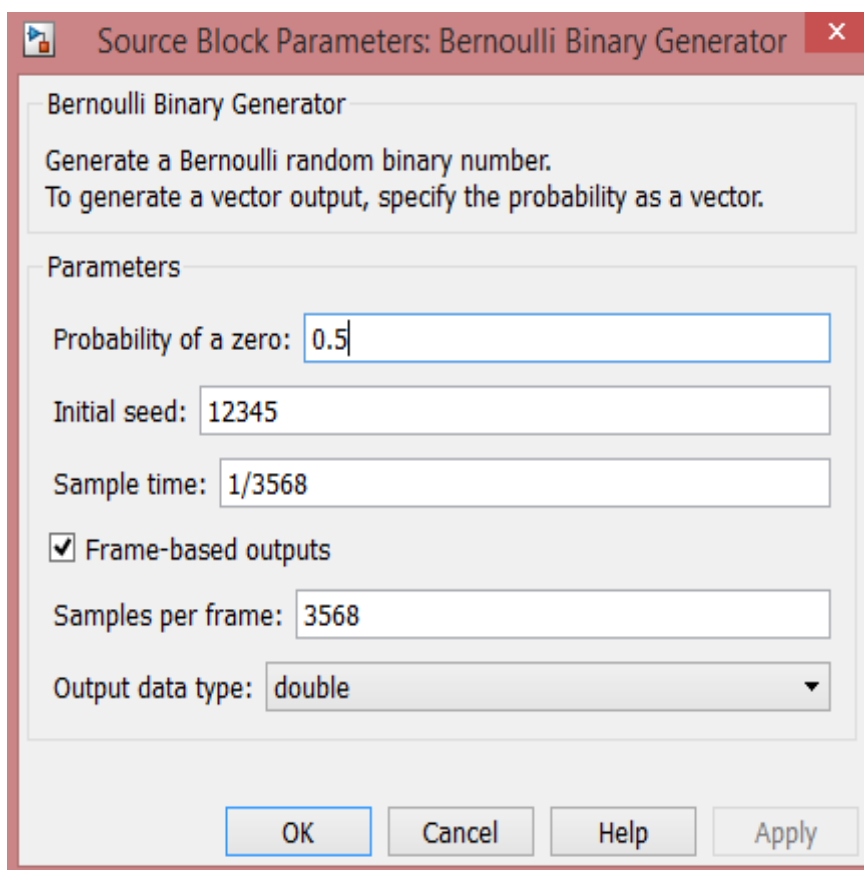


Рис. 3.144. Параметры блока «Bernoulli Binary Generator»

«Probability of a zero» - вероятность появления нуля;

«Initial seed» - начальное значение для генерации;

«Sample time» - длительность сэмпла;

«Samples per frame» - размер фрейма.

Binary Input RS Encoder – кодер Рида-Соломона (рисунок 3.7).

«Codeword length N» - общее количество бит;

«Message length K» - количество информационных бит.

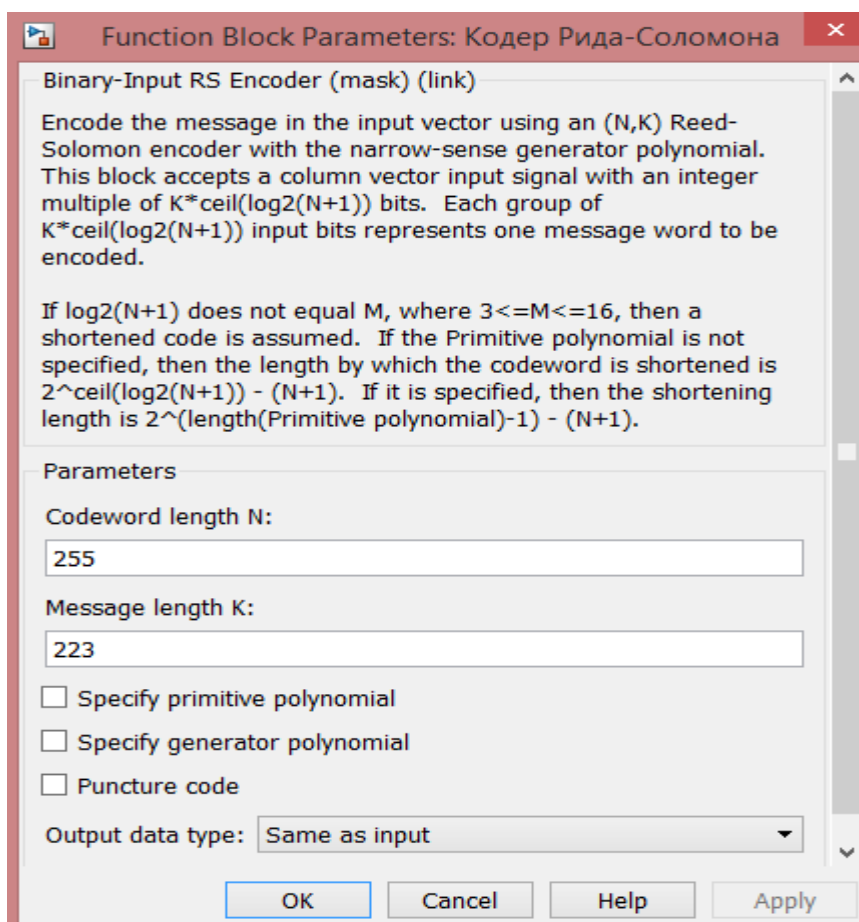


Рис. 3.145. Параметры блока «Binary Input RS Encoder»

Данный блок имеет следующую структуру внутри (рисунок 3.8):

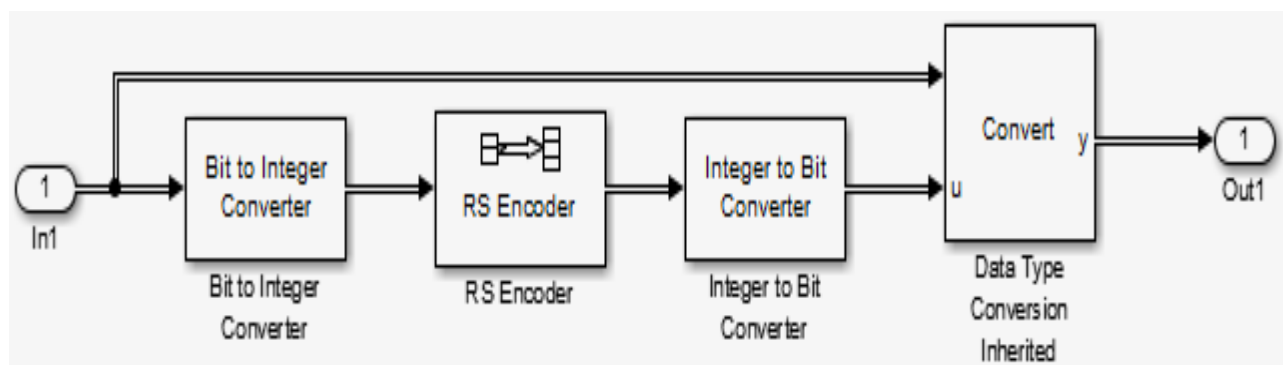


Рис. 3.146. Состав блока «Binary Input RS Encoder»

Таким образом, информационные биты, поступающие со входа генератора преобразуются в тип «Integer», кодируются и преобразуются обратно, затем полученные биты конвертируются в тот тип данных, который изначально был на входе кодера.

Turbo Encoder – Турбо-кодер (рисунок 3.147).

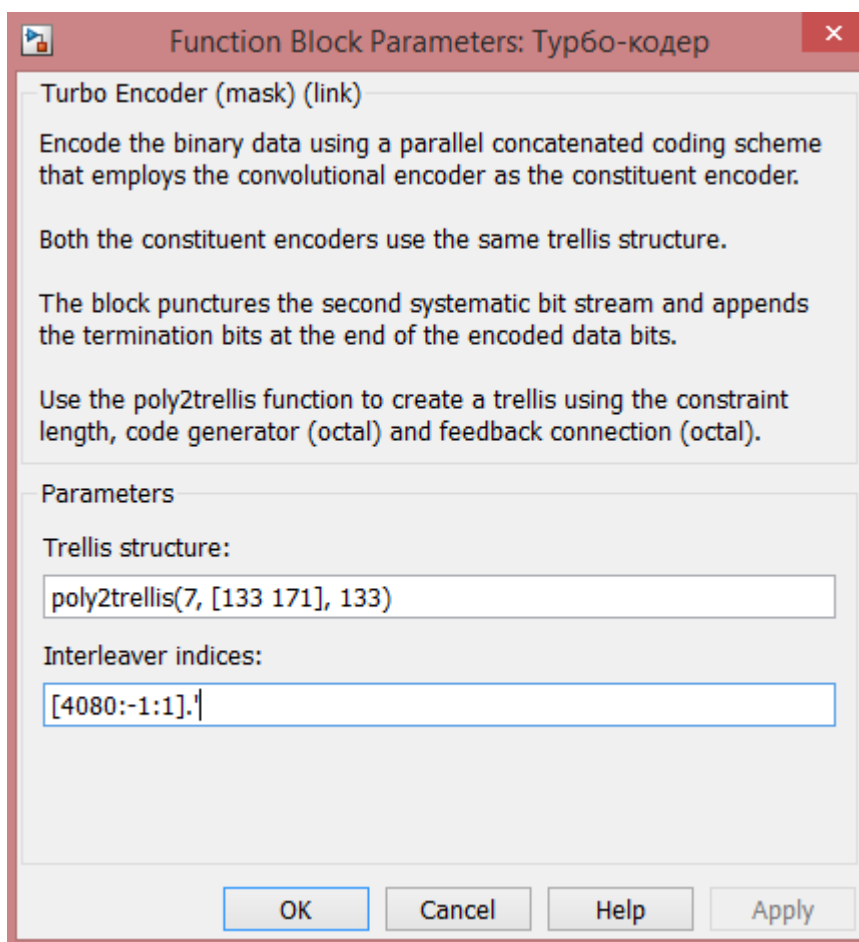


Рис. 3.147. Параметры блока «Turbo Encoder»

«Trellis structure» - структура треллис-модуляции.

Треллис-модуляция (TCM – Trellis Coded Modulation) представляет собой способ, который позволяет обеспечить повысить скорость передачи сообщения с сохранением уровня помехоустойчивости. Этот способ отличается тем, что помехоустойчивое кодирование и тип модуляции используются совместно. Выбранная соответствующим образом пара помехоустойчивый код – способ модуляции часто также носит название сигнально-кодовая конструкция (СКК).

Для кодирования использован один из наиболее часто употребляемых свёрточных кодов – код (171,133,7), который кодирует последовательность со скоростью $1/2$.

«Interleaver indices» - входные параметры перемежителя.

Схема турбокодера имеет следующую структуру (рисунок 3.148):

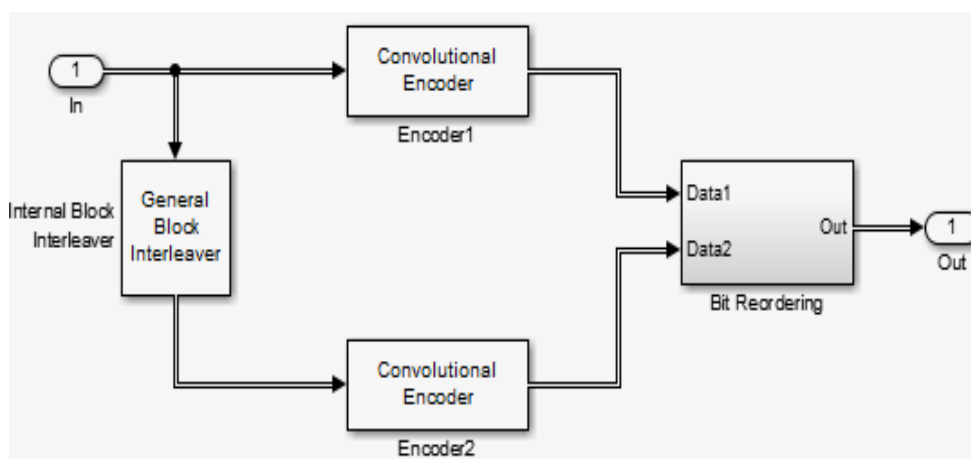


Рис. 3.148. Структура блока «Turbo Encoder»

Поток бит распараллеливается на два. В первом случае биты поступают на свёрточный кодер (133, 171), а во втором потоке биты сначала проходят перемежитель, затем поступают на аналогичный свёрточный кодер. Блок «Bit Reordering» выстраивает биты в последовательный поток.

Rectangular QAM Modulator Baseband – модулятор QAM-M (рисунок 3.149).

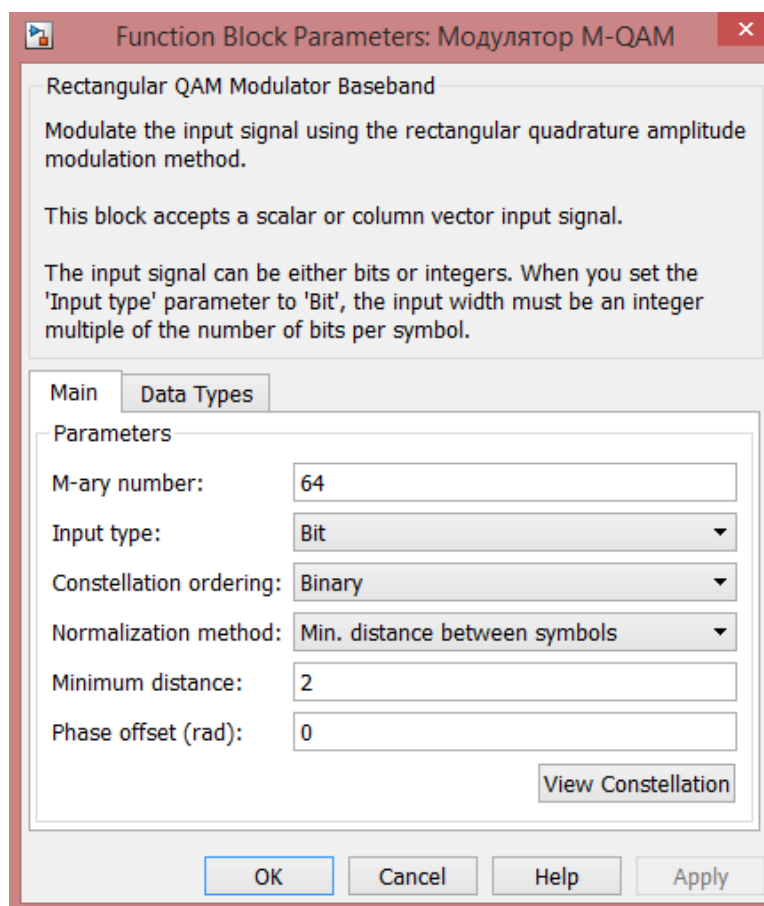


Рис. 3.149. Параметры блока «Turbo Encoder»

«M-ary number» - количество позиций в QAM-M;

«Input type» - тип входных данных;

«Constellation ordering» - порядок построения созвездия (рисунок 3.150).

Остальные параметры по умолчанию.

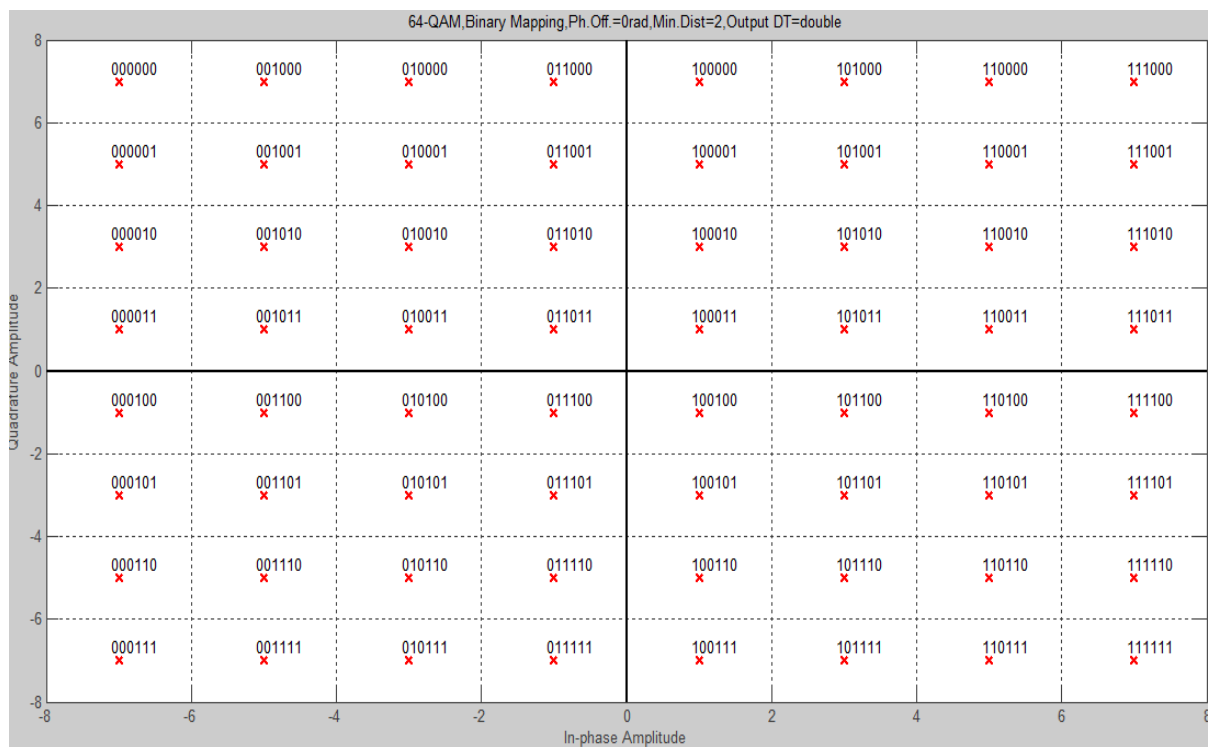


Рис. 3.150. Диаграмма построения созвездий

AWGN Channel (Канал связи) – добавляет «белый» гауссовский шум в канал (рисунок 3.13).

«Variance» - считывает параметр E_b/N_0 из блока Model Parameters.

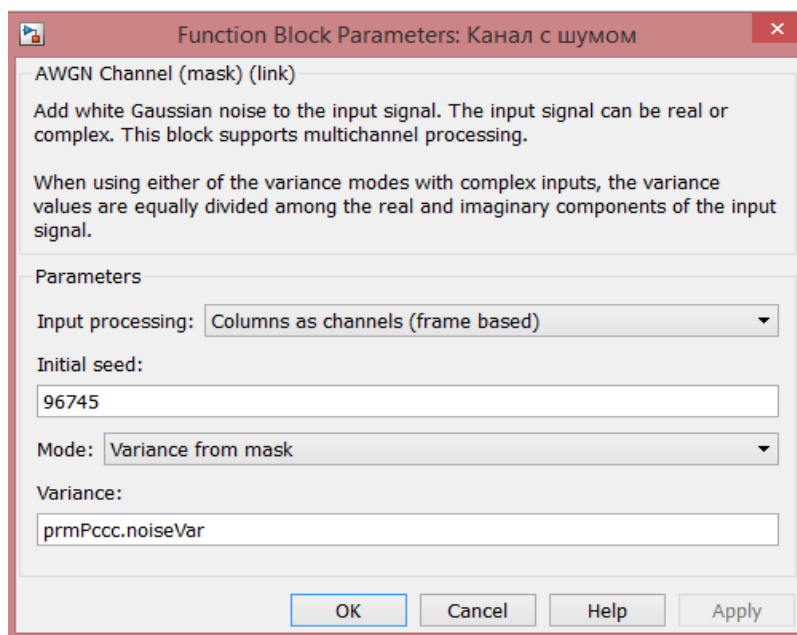


Рис. 3.151. Параметры блока «AWGN»

Discrete Time Scatter Plot Scope – Блок для отражения диаграммы созвездий.

Rectangular QAM Demodulator Baseband – демодулятор QAM-M, обладает теми же параметрами, что и модулятор.

Unipolar to Bipolar Converter – Преобразователь сигнала из однополярного в биполярный. На вход турбо-декодера необходимо подавать биполярный сигнал.

Turbo Decoder – Турбо-декодер (рисунок 3.152).

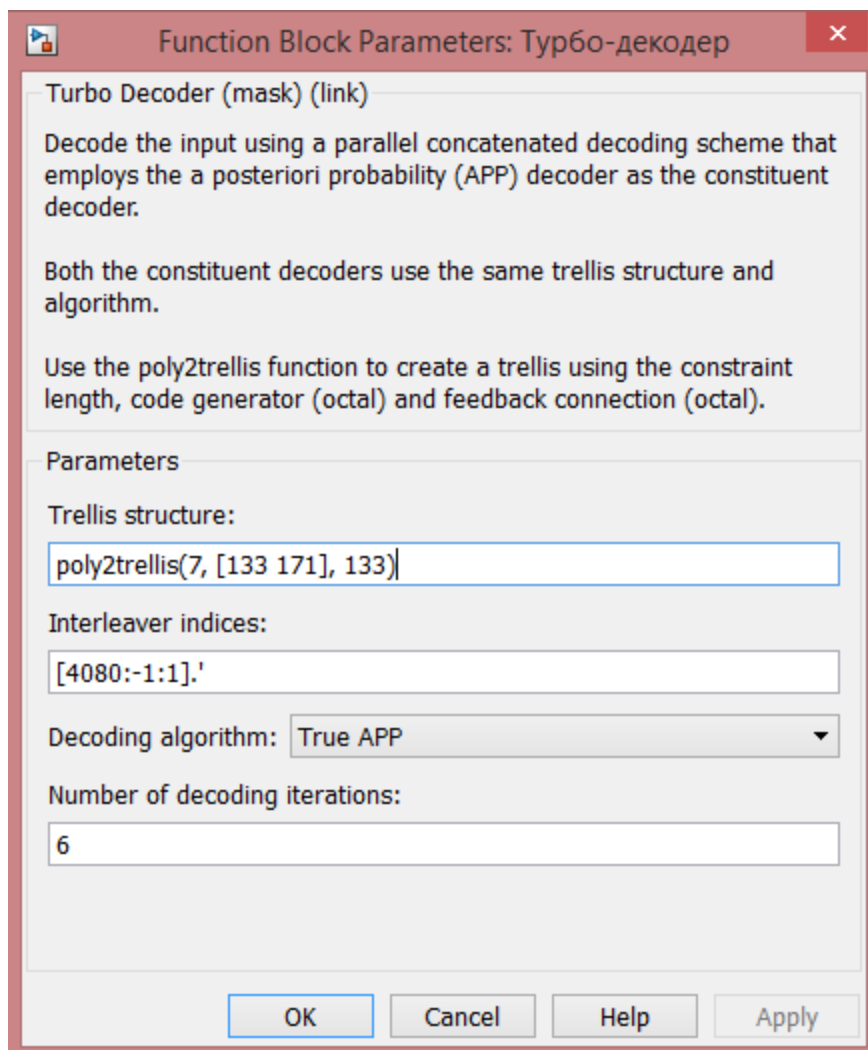


Рис. 3.152. Параметры блока «Turbo decoder»

Первые два параметра задаются аналогично параметрам кодера.

«Number of decoding iterations» – количество итераций декодирования. Декодирование в турбо-декодере происходит в несколько итераций. Чем больше итераций, тем точнее декодирование. Однако, большое количество не даёт результата, а лишь увеличивает длительность вычислений и может даже ухудшить помехоустойчивость.

Схема блока имеет следующую структуру (рисунок 3.153):

Binary Input RS Decoder – декодер Рида-Соломона. Параметры и структура аналогична блоку кодера.

Error Rate Calculation – вычислитель ошибок между переданной и принятой последовательностью.

Display - дисплей, отражающий ошибки.

Исследование каскадных кодов

1. Спроектированная модель передачи данных демонстрирует работу каскадного кодирования (рисунок 3.143). Данная модель позволяет исследовать применение последовательно-параллельного кодирования на примере использования кодера Рида-Соломона (внешний код) и Турбо-кодера (внутренний код), а также позволяет исследовать исправляющую способность кодов для разных видов модуляции и сравнить её характеристики с работой указанных выше кодеров в отдельности.

2. В качестве турбо-кода используются два параллельных свёрточных кодера с треллис-модуляцией (для ускорения передачи данных).

На рисунке 3.155 представлен график зависимости битовой вероятности ошибки (BER) от отношения сигнал/шум в канале (SNR) для разных видов модуляции:

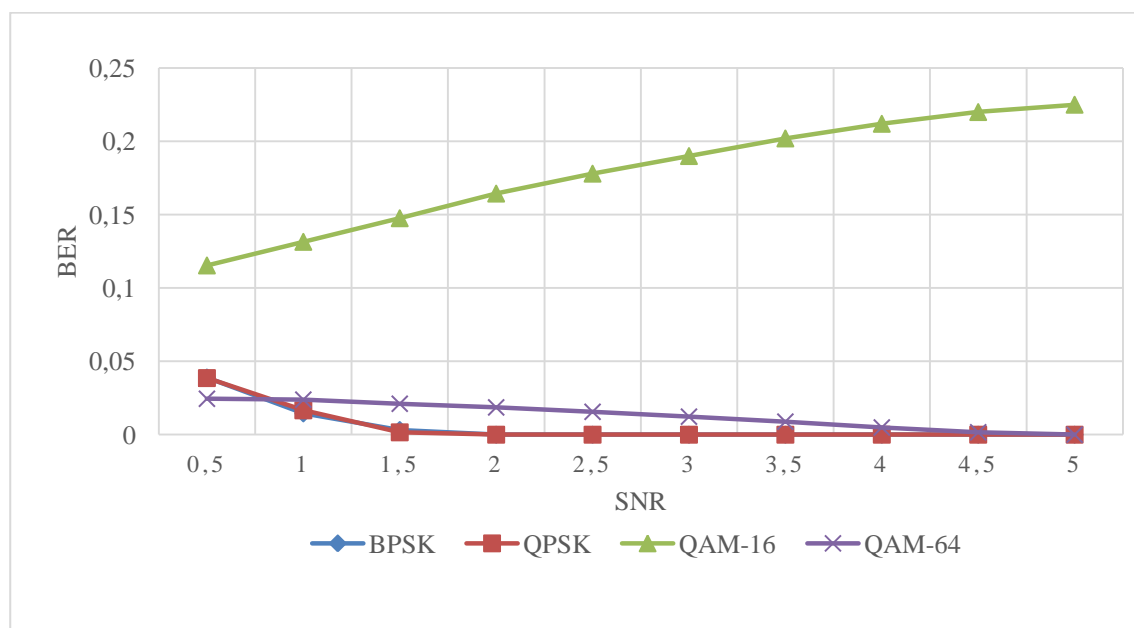


Рис. 3.155. График зависимости BER от SNR для разных видов модуляции

Из рисунка 3.155 можно заметить, что зависимость BER от SNR является неправильной в явном виде. Предположительно, данное явление связано с ошибочным программным кодом самого блока QAM-Modulator, поэтому данную зависимость рассматривать не будем. На

рисунке 3.156 представлен график зависимости BER от SNR в каскадных кодах для видов модуляции BPSK, QPSK и QAM-64:

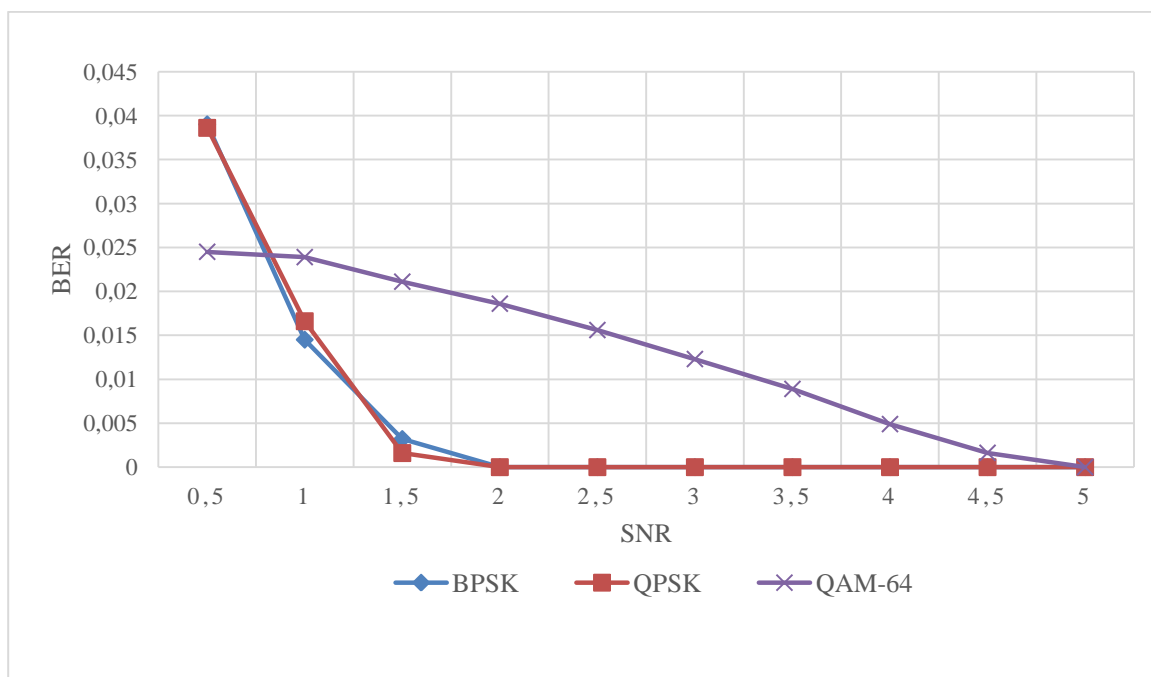


Рис. 3.156 – График зависимости BER от SNR для разных видов модуляции

На рисунке 3.157 представлена диаграмма созвездий QAM-64 сигнала на выходе передатчика:

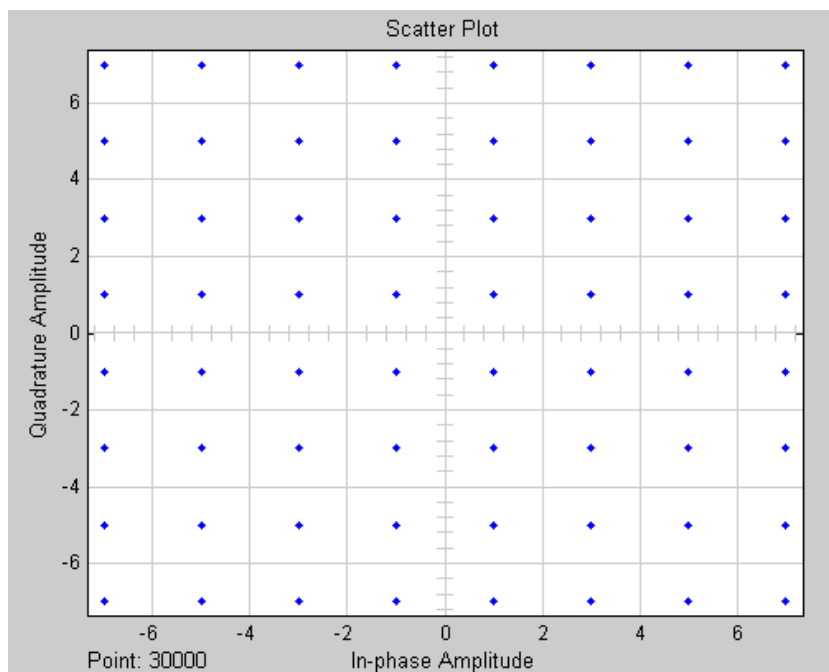


Рис. 3.158. Диаграмма созвездий QAM-64 сигнала на выходе передатчика

На рисунке 3.159 представлена диаграмма созвездий QAM-64 сигнала на приёмном конце после канала с шумом ($\text{SNR} = 3$ дБ) :

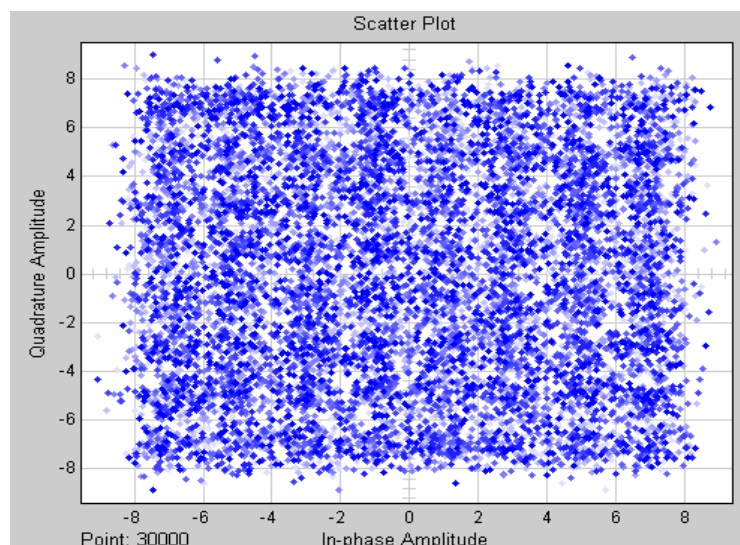


Рис. 3.160. Диаграмма созвездий QAM-64 сигнала на входе приёмника

На рисунке 4.7 представлена диаграмма созвездий QAM-64 сигнала после исправления ошибок каскадным декодером ($\text{SNR} = 3$ дБ) :

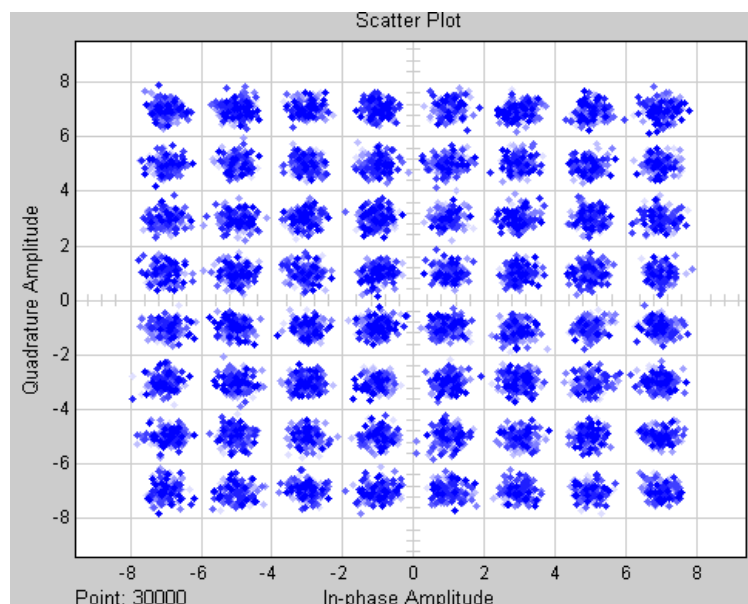


Рис. 3.161. Диаграмма созвездий QAM-64 сигнала после декодирования

На рисунке 3.162 представлены временные формы сигнала QAM-64 с каскадным кодированием ($\text{SNR} = 3$ дБ) :

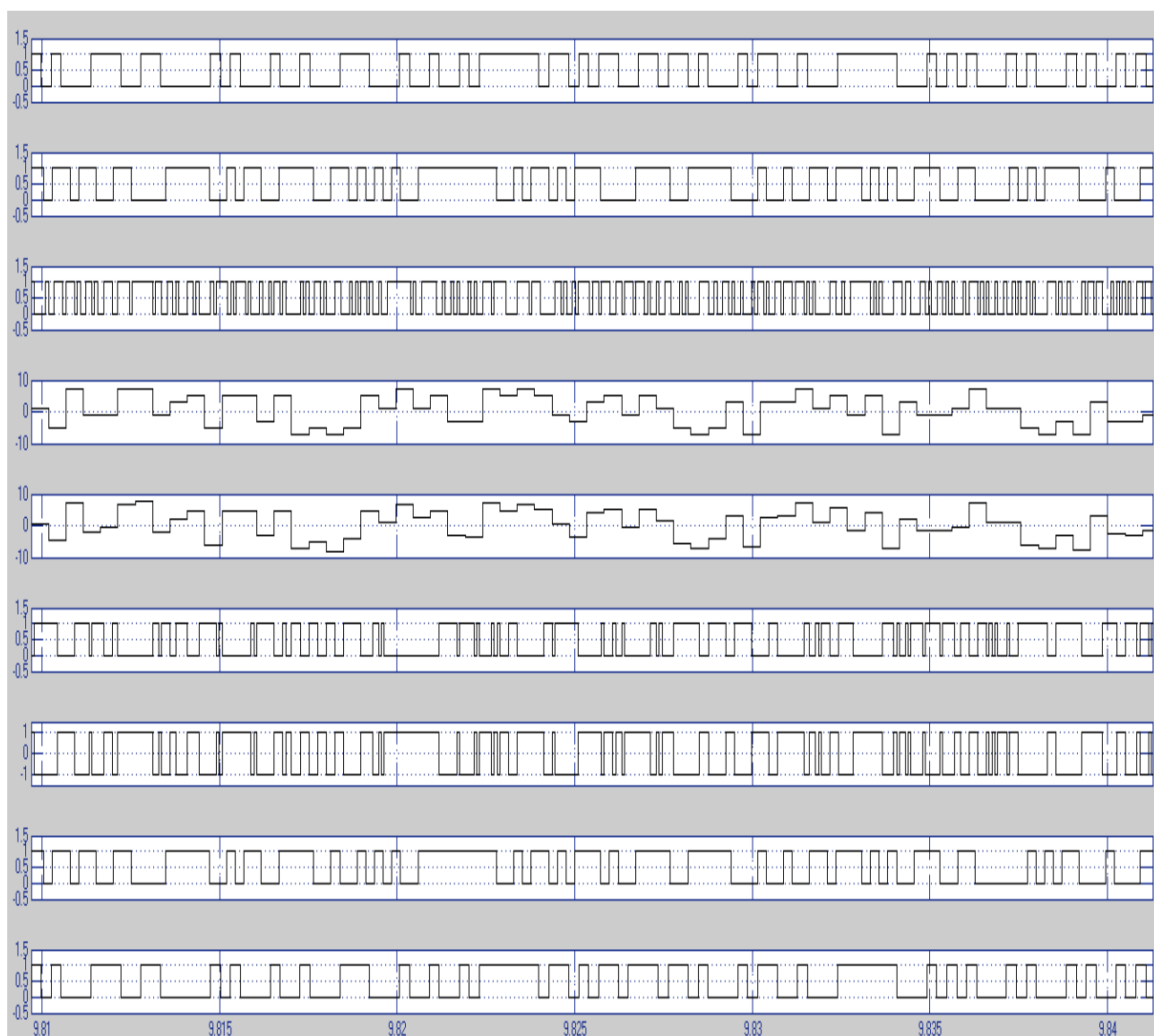


Рис. 3.162. Временные формы сигналов (сверху вниз):

- 1) на выходе генератора псевдослучайной последовательности;
- 2) на выходе внешнего кодера (РС);
- 3) на выходе внутреннего кодера (турбо-кодера);
- 4) на выходе QAM-модулятора;
- 5) на входе QAM-демодулятора;
- 6) на выходе QAM-демодулятора;
- 7) на выходе преобразователя сигнала из однополярного в биполярный;
- 8) на выходе внутреннего декодера (турбо-декодера);
- 9) на выходе внешнего декодера (РС).

На рисунке 3.163 представлен график зависимости BER от SNR при модуляции QAM-64 для каскадного кода, внешнего кода и внутреннего кода:

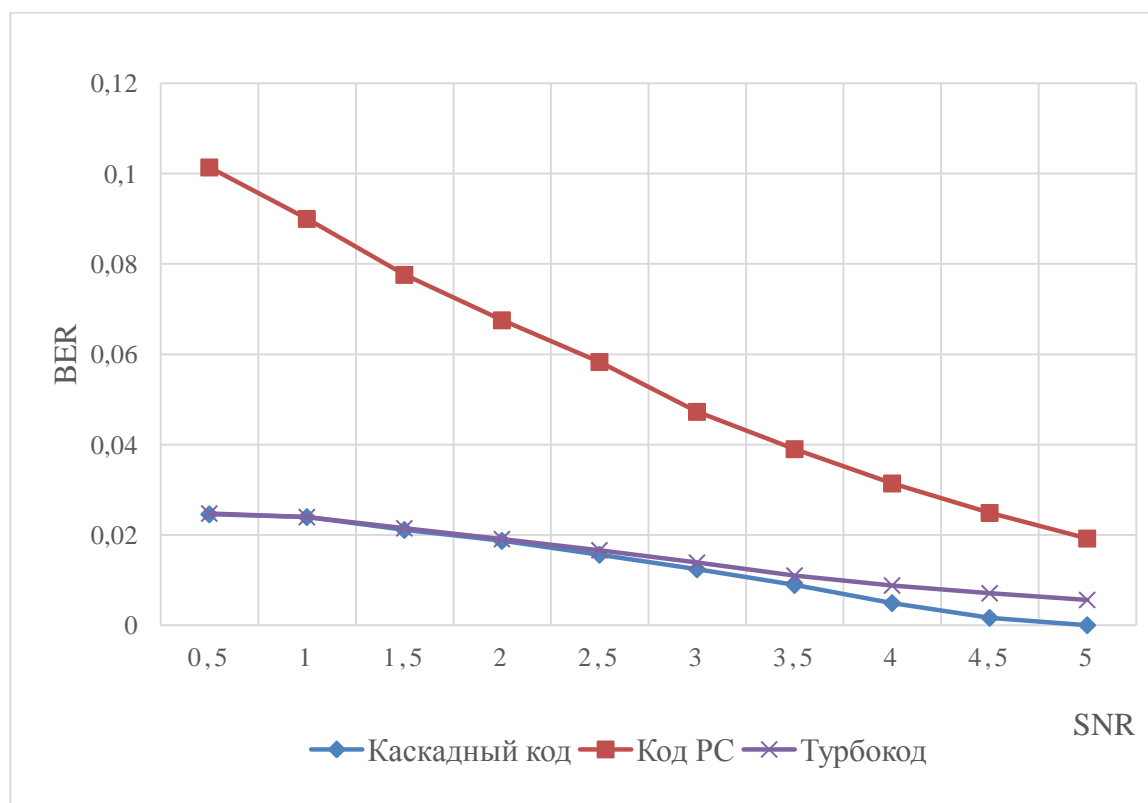


Рис. 3.163. График зависимости BER от SNR для каскадного кода, кода Рида-Соломона и турбокода

Как видно из рисунка 4.9, применение каскадного кодирования неоправданно по сравнению с применением простого турбокодирования, однако он имеет гораздо лучшую характеристику, чем применение простого помехоустойчивого кодирования (РС).

Также данная модель позволяет исследовать исправляющую способность каскадного кода в зависимости от количества итераций декодирования (рисунок 4.10).

Как видно из графика, повышение количества итераций декодирования не улучшает помехоустойчивость, а даже делает её чуть хуже, и, к тому же, приводит к повышению времени декодирования каждой посылки, кратное количеству этих итераций.

Можно сделать предположение, что повышение количества итераций необходимо при увеличении размера фрейма.

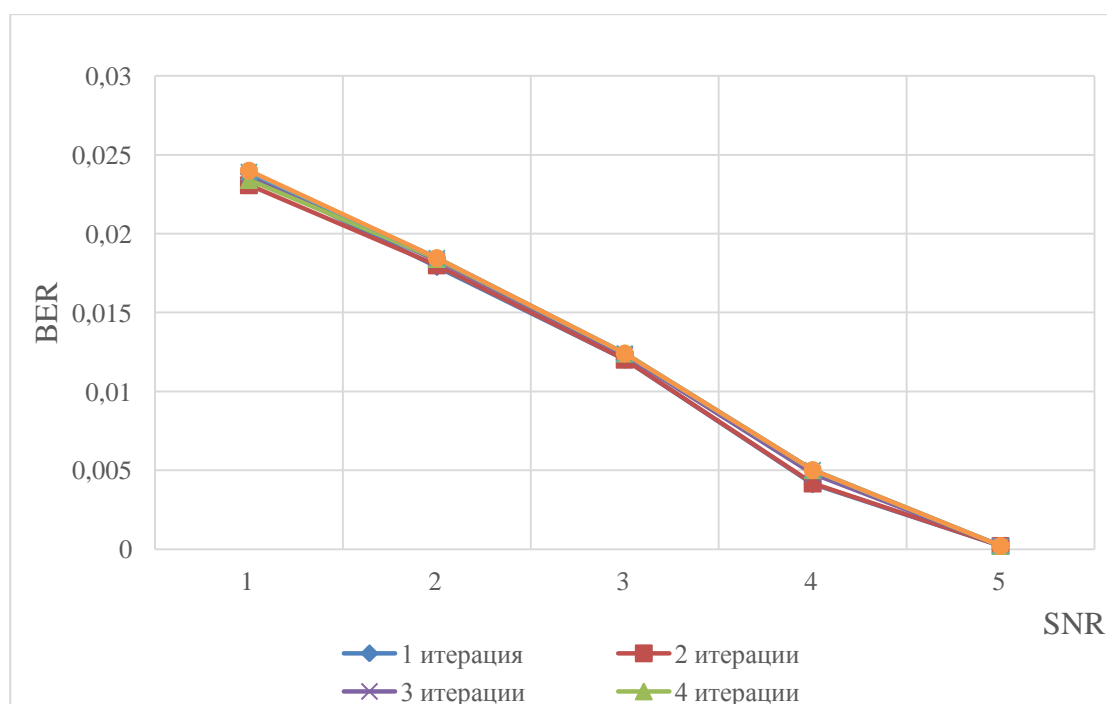


Рис. 3.164. График зависимости BER от SNR для разных значений количества итераций декодирования при количестве переданных символов $\sim 75\ 000$.

В результате работы спроектированы модель исследования каскадных кодов.

Модель позволяет исследовать работу каскадных кодов. В качестве внешнего кода используется код Рида-Соломона, в качестве внутреннего – Турбо-код на базе свёрточного кодирования и треллис-модуляции. Данная модель позволяет исследовать зависимость битовой вероятности ошибки (BER) системы от отношения сигнал/шум (SNR) в канале.

Получены следующие результаты и выводы:

Как видно из рисунка 4.9, применение каскадного кодирования неоправданно по сравнению с применением простого турбо-кодирования.

Повышение количества итераций декодирования не улучшает помехоустойчивость при одинаковой характеристике канала, и, к тому же, приводит к повышению времени декодирования каждой посылки, кратное количеству этих итераций.

Повышение количества итераций необходимо при изменении параметров канала.

Эффективность от применения каскадного кодирования заметна лишь при значительном размере фрейма ($k > 10000$).

Методические указания позволяют использовать данные модели на лабораторных работах студентами для исследования помехоустойчивых кодов.

4. СИГНАЛЬНО-КODOVЫЕ КОНСТРУКЦИИ В ТЕЛЕКОММУНИКАЦИОННЫХ СИСТЕМАХ

4.1. Сигнально-кодovые конструкции на основе Треллис кодovой модуляции (TCM) и их анализ с использованием MATLAB [16]

В цифровых системах связи обычно повышают скорость передачи данных путем уменьшения энергетической емкости бита, т.е. количества энергии сигнала, приходящейся на один бит информации. Но чем меньше энергия, тем больше вероятность того, что бит будет искажен в канале при передаче. Поэтому при повышении скорости передачи разработчики всегда сталкиваются со снижением уровня помехоустойчивости.

Для повышения помехоустойчивости канала передачи данных в цифровых системах применяются коды, исправляющие ошибки. Однако действие таких кодов не всегда эффективно, так как снижается скорость передачи данных.

Треллис-модуляция (TCM – Trellis Coded Modulation) представляет собой способ, который позволяет обеспечить повысить скорость передачи сообщения с сохранением уровня помехоустойчивости. Этот способ отличается тем, что помехоустойчивое кодирование и тип модуляции используются совместно. Выбранная соответствующим образом пара помехоустойчивый код – способ модуляции часто также носит название сигнально-кодovая конструкция (СКК).

В данной работе описан способ включения сверточного кодека, используемого для передачи данных с помощью радиорелейных систем связи, в режим треллис-модуляции. Такой способ повышает скорость передачи в два раза при сохранении уровня помехоустойчивости.

В цифровых системах связи обычно повышают скорость передачи данных путем уменьшения энергетической емкости бита, т.е. количества энергии сигнала, приходящейся на один бит информации. Но чем меньше энергия, тем больше вероятность того, что бит будет искажен в канале при передаче. Поэтому при повышении скорости передачи разработчики всегда сталкиваются со снижением уровня помехоустойчивости.

Для повышения помехоустойчивости канала передачи данных в цифровых системах применяются коды, исправляющие ошибки. Однако действие таких кодов не всегда эффективно, так как снижается скорость передачи данных.

Треллис-модуляция (TCM – Trellis Coded Modulation) представляет собой способ, который позволяет обеспечить повысить скорость передачи сообщения с сохранением уровня помехоустойчивости. Этот способ отличается тем, что помехоустойчивое кодирование и тип модуляции используются совместно. Выбранная соответствующим образом пара помехоустойчивый код – способ модуляции часто также носит название

сигнально-кодовая конструкция (СКК).

В данной работе описан способ включения сверточного кодека, используемого для передачи данных с помощью радиорелейных систем связи, в режим треллис-модуляции. Такой способ повышает скорость передачи в два раза при сохранении уровня помехоустойчивости.

Корректирующие коды

Наряду с многопозиционными сигналами для повышения эффективности системы электрической связи (СЭС) широко используются помехоустойчивые коды. Применение корректирующих кодов позволяет повысить верность передачи сообщений или при заданной верности повысить энергетическую эффективность системы. Это особенно важно для систем с малой энергетикой, например, систем спутниковой связи.

На практике используются как блочные, так и непрерывные коды. На рис. 4.1 приведены кривые эффективности для циклического кода Боуза-Чоудхури-Хоквингема (БЧХ) и для сверточного кода (СК) с декодированием по алгоритму Витерби.

Применение циклического кода позволяет получить энергетический выигрыш $\Delta\beta = 2...4$ дБ, а сверточного кода $\Delta\beta = 4...6$ дБ в обмен на снижение частотной эффективности примерно в 2 раза (3 дБ).

Энергетический выигрыш $\Delta\beta$ от применения помехоустойчивого кодирования тем больше, чем выше требуемая верность передачи. Для непрерывного канала с белым гауссовским шумом при требуемой вероятности ошибки 10^{-5} предельный энергетический выигрыш кодирования по сравнению с ФМн-2 без кодирования при оптимальном когерентном приеме составляет примерно 10 дБ. Расчетные кривые на рис. 9.2 показывают, что применение циклического кода в канале с фазовой манипуляцией (ФМн) или сверточного кода в канале с АФМ позволяет повысить одновременно энергетическую, так и частотную эффективность. Построение таких высокоэффективных систем на основе сигнально-кодовых конструкций ведет к неизбежному увеличению сложности системы. Не пропускная способность, а сложность является ограничивающим фактором при построении высокоэффективных систем. Задача состоит в том, чтобы построить систему, удовлетворяющую высоким показателям эффективности, при допустимой сложности.

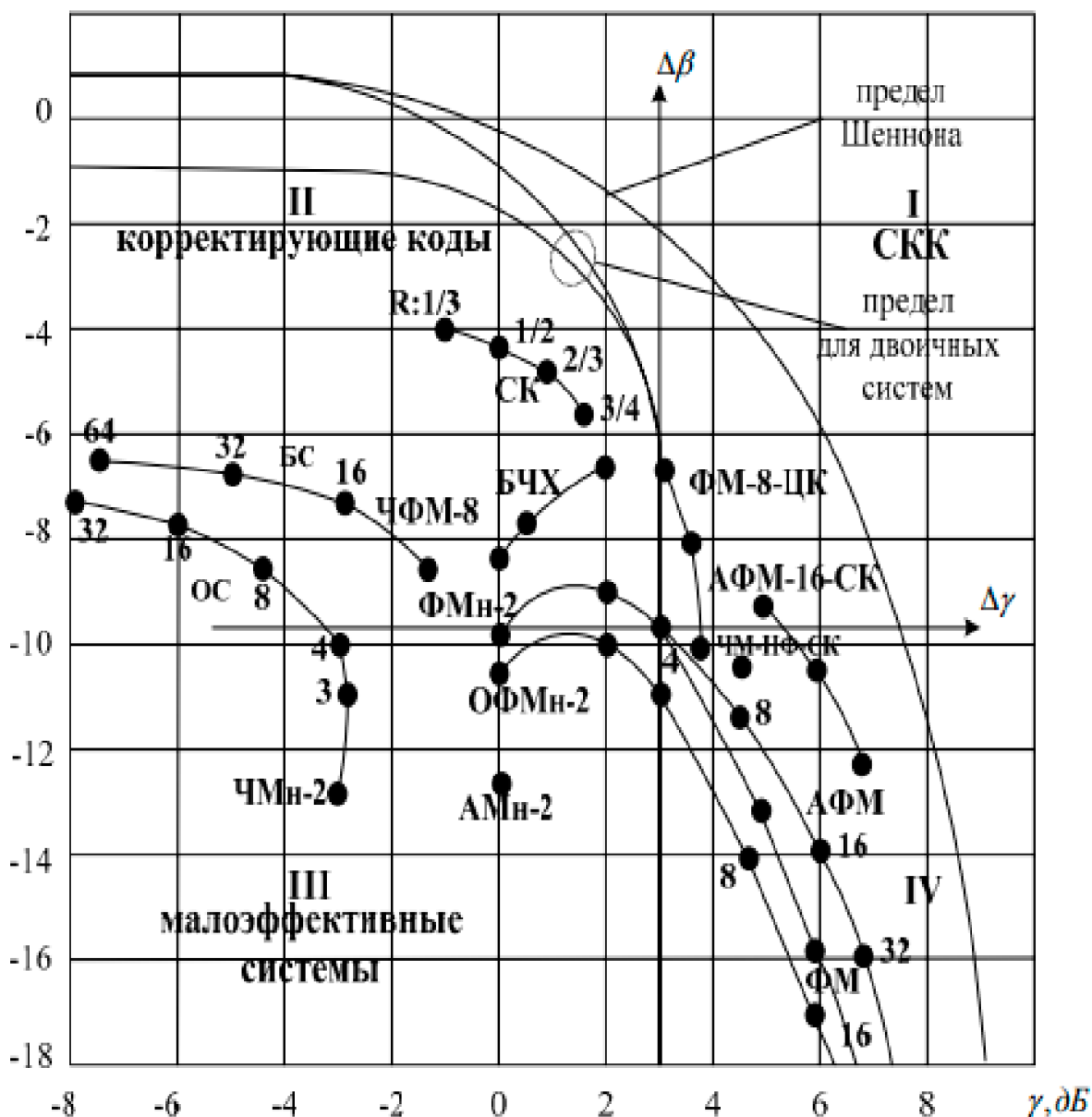
$\beta, \text{дБ}$ 

Рис.4.1. Графики зависимости энергетической и частотной эффективности систем связи

При современной элементной базе затраты на реализацию кодирующих и декодирующих устройств значительно сократились. В то же время стоимость энергетики канала практически не изменилась. Таким образом, «цена» выигрыша $\Delta\beta$ за счет кодирования может быть существенно меньше цены того же выигрыша, полученного за счет увеличения энергетики канала (мощности сигнала или размеров антенн).

Отметим, что выбор способов кодирования и модуляции зависит от характеристик

канала. Улучшение этих характеристик, например, путем адаптации к помехам и оценивания искажений сигнала и их последующей компенсации, снижает потери в канале и создает лучшие условия для применения корректирующих кодов.

Оптимизация систем связи

Повышение таких важнейших показателей систем электрической связи, как скорость и верность передачи, связано со значительными частотными и энергетическими затратами. Сравнение между собой различных СЭС осуществляется по степени использования ими основных ресурсов канала связи (пропускной способности, мощности, занимаемой полосы частот), выражаемой через показатели информационной, энергетической и частотной эффективности. Создание СЭС, в которых достигаются близкие к предельным показатели эффективности, требует совместного согласования кодека и модема с учетом статистических свойств непрерывного канала.

Согласование методов модуляции и кодирования

Эффективный путь повышения удельной скорости передачи информации заключается в увеличении числа используемых сигналов m на интервале T . Однако увеличение m приводит к уменьшению расстояния между ближайшими сигналами ансамбля и снижению энергетической эффективности.

При высоких требованиях к верности передачи целесообразным становится применение помехоустойчивых кодов, которые позволяют повысить энергетическую эффективность за счет снижения частотной. Помехоустойчивое кодирование позволяет снизить необходимую величину мощности сигнала, поскольку расстояние между кодовыми комбинациями увеличивается. Одновременное требование большой скорости и верности передачи в условиях ограниченного частотного и энергетического ресурса может быть выполнено при использовании многопозиционных сигналов и помехоустойчивых кодов.

При многопозиционной модуляции, когда по каналам связи передается блок из n кодовых символов, важно также правильно выбрать манипуляционный код, определяющий правило сопоставления с каждым передаваемым сигналом определенного блока кодовых символов. Общий принцип заключается в том, что большему расстоянию по Хэммингу между кодовыми блоками должно соответствовать большее расстояние по Евклиду между отображающими их сигналами.

Создание СЭС, в которых достигаются близкие к предельным показатели эффективности, требует совместного согласования кодека и модема с учетом статистических свойств непрерывного канала. Это означает, что кодирование и модуляцию необходимо рассматривать как единый процесс формирования сигнала, а демодуляцию и декодирование – как процесс оптимального приема сигнально-кодированного блока в целом.

Согласование модуляции и кодирования сводится к поиску такого заполнения сигнального пространства, при котором обеспечивается высокая удельная скорость (сигналы расположены достаточно плотно) и одновременно высокая помехоустойчивость (сигналы достаточно далеко друг от друга).

Комбинирование различных ансамблей m -ичных сигналов, помехоустойчивых и манипуляционных кодов порождает множество конструкций. Однако только согласованные варианты обеспечивают повышение частотно - энергетической эффективности СЭС.

Эти варианты называют **сигнально - кодовыми конструкциями (СКК)**. Рассмотрим обобщенную схему передачи дискретных сообщений, приведенную на рис.4.2.

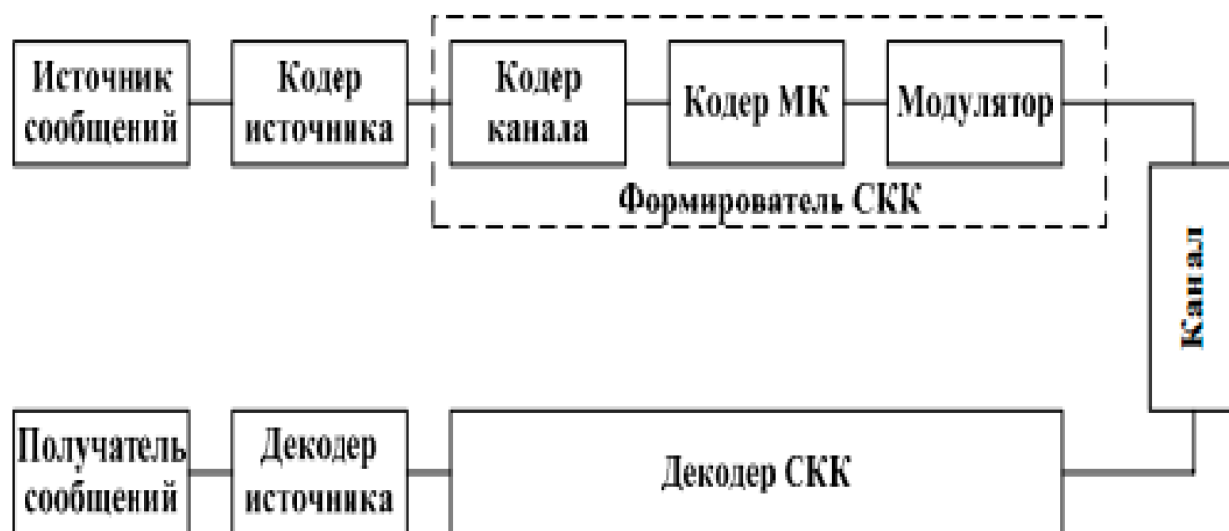


Рис. 4.2. Обобщенная схема передачи дискретных сообщений

Последовательность символов с выхода кодера канала разбивается на блоки по n символов. Отображение блоков в сигналы, формируемые модулятором, осуществляется по правилу манипуляционного кодирования, т.е. манипуляционный код определяет правило соответствия блоков кодовых символов m -ичным сигналам. Например, в случае двоичного канального кода каждому из $m = 2^n$ кодовых блоков ставится в соответствие один из 2^n сигналов.

Оптимальная процедура приема СКК заключается в обработке сигнально - кодового блока в целом. Поэтому демодулятор, декодеры канала и манипуляционный декодер рассматриваются как единое устройство – декодер СКК (рис.4.2).

Декодер СКК строится так, чтобы минимизировалась вероятность ошибки приема. Оптимальный декодер реализует принцип максимального правдоподобия. При белом

гауссовском шуме выбирается кодовое слово, находящееся на минимальном евклидовом расстоянии от принятого.

Декодирование МК можно рассматривать как последний этап обработки сигнально – кодового блока оптимальным декодером СКК. При этом декодер канала работает в метрике Евклида с сигналами, а не с их двоичными представлениями по правилу манипуляционного кода. Схема поэлементного приема, наоборот, ориентирована на применение декодера канала в метрике Хемминга, т.е. обработку двоичных величин после декодера манипуляционного кода.

Следовательно, достижение наибольшей эффективности возможно при декодировании по алгоритму максимального правдоподобия сигнально - кодового блока в целом». Необходимо отметить, что в принципе любой сигнальный ансамбль на выходе последовательно соединенных кодека и модема может быть отнесен к СКК. Введение понятия СКК отражает подход к модуляции и кодированию как процессу объединения сигналов и кодов в единую эффективную конструкцию.

Классификация сигнально – кодовых конструкций

В основе формирования СКК лежат операции отображения информационной последовательности в кодовую, путем внесения избыточности и кодовой последовательности в канальную заданием манипуляционного кода. Помехоустойчивое кодирование, повышающее энергетическую эффективность СЭС, является одной из важнейших операций формирования СКК. Получаемый при этом энергетический выигрыш от кодирования зависит от степени увеличения минимального сигнального расстояния между разрешенными кодовыми блоками. В качестве сигнального для канала АБГШ используется расстояние Евклида. Асимптотический энергетический выигрыш определяется формулой [21]:

$$\text{ЭВК[ДБ]} = 20 \lg \left(\frac{d_{ef}}{d_e} \right), \quad (4.1)$$

где, d_{ef} – минимальное евклидово расстояние между разрешенными кодовыми блоками; d_e – минимальное евклидово расстояние между различными некодированными последовательностями канальных символов одинаковой мощности с кодированными символами.

Согласно (4.1), для получения больших величин энергетического выигрыша при построении СКК необходимо подбирать коды, максимизирующие минимальное евклидово расстояние между разрешенными кодовыми комбинациями. В основу классификации СКК можно положить отличительные особенности по типам

помехоустойчивого кода, по типам ансамблей сигналов и по способам согласования модуляции и кодирования.

По типу помехоустойчивых кодов все СКК могут быть поделены на два больших класса: СКК на основе блочных кодов и СКК на основе непрерывных кодов. Кроме того, отдельный класс составляют СКК на основе каскадных кодов, в которых применяются одновременно блочные и непрерывные коды. Каждый из классов делится на группы по конкретным видам кода.

Среди блочных наиболее употребляемыми являются коды Хэмминга, Голея, БЧХ, Рида - Соломона, Рида-Маллера и др. Непрерывные коды на практике представлены сверточными кодами, которые обладают дополнительными свойствами линейности, и постоянства во времени.

При использовании сверточного кода практически удобным является случай, когда при объеме ансамбля сигналов $m = 2^{k+1}$ скорость сверточного кода выбирается равной $R_{СКК} = k/k+1$. Тогда частотная эффективность у системы с кодированием и без него одна и та же. Поскольку каждый кодовый блок длиной $(k + 1)$ переносится одним двумерным сигналом, то и СКК считается также двумерной.

Декодирование СКК ведется обычно по алгоритму Витерби, реализующему принцип максимального правдоподобия. Одно из важнейших преимуществ СКК заключается в простоте применения алгоритма Витерби для мягкого решения на выходе демодулятора.

Любая СКК вне зависимости от способа согласования модуляции и кодирования представляет собой каскадный код с ансамблем сигналов на внутренней ступени и одним или несколькими помехоустойчивыми кодами на внешней. При использовании нескольких помехоустойчивых кодов говорят о построении СКК на основе обобщенного каскадного кода.

По типу ансамблей сигналов СКК делятся на конструкции с одномерными, двумерными и многомерными сигналами. Многомерные сигналы состоят из более простых (одномерных, двумерных) сигналов. При использовании в качестве составляющих двумерных сигналов число позиций M , соответствующих каждому n -мерному сигналу, определяется выражением $M = m^n/2$, где m – позиционность двумерного сигнала.

Каждый n -мерный сигнал в этом случае образуется последовательностью $n/2$ двумерных сигналов. Например, для получения многомерного сигнала с $n = 6$ требуется последовательность из трех двумерных сигналов, например ФМн-4.

Способы согласования модуляции и кодирования условно можно разделить на две группы: согласование кодом Грея и согласование на основе разбиения ансамбля на вложенные подансамбли.

Сигнально-кодовые конструкции, принадлежащие первой группе, представляют собой результат согласования известных двоичных помехоустойчивых кодов с многопозиционным ансамблем сигналов путем использования кода Грея в качестве манипуляционного кода. Поскольку ошибки происходят за счет переходов в области соседних сигналов, то кодовые блоки, соответствующие соседним сигналам, должны различаться наименьшим числом двоичных символов.

Вторая группа включает в себя достаточно большое число типов СКК, различающихся модификациями методов согласования. Разбиение осуществляется таким образом, что подансамбли содержат равное количество сигналов, расстояния между соседними сигналами подансамблей одинаковы, а минимальные расстояния между сигналами подансамблей увеличиваются с каждым шагом разбиения. Широкое практическое применение получило согласование путем разбиения ансамбля на вложенные подансамбли, когда внешними кодами являются сверточные коды. В основе синтеза СКК со сверточными кодами лежит поиск кодов, максимизирующих евклидово расстояние, причем обычно эти коды не являются оптимальными в метрике Хэмминга. У решетчатой диаграммы, описывающей сверточные коды в метрике Евклида, переходы между состояниями промаркированы не двоичными блоками, а сигнальными точками.

Таким образом, достижение близких к предельным показателей частотно - энергетической эффективности цифровых систем связи предполагает согласование кодека и модема с учетом статистических свойств непрерывного канала. Одно из решений подобного согласования представляют сигнально-кодовые конструкции сверточного кодирования. Мягкое декодирование по алгоритму Витерби обеспечивает энергетический выигрыш порядка 3...7 дБ без расширения занимаемой полосы частот.

Характеристики основных типов СКК. Согласование канала кодом Грея
Рассмотрим СКК, представляющие собой результат согласования известных двоичных помехоустойчивых кодов с многопозиционным ансамблем сигналов путем использования в качестве манипуляционного кода Грея (табл. 4.1).

Комбинации кода в табл. 4.1 получены по следующему правилу. Кодовая комбинация натурального кода складывается по модулю 2 с такой же комбинацией, сдвинутой на один разряд вправо, при этом младший разряд сдвинутой комбинации отбрасывается.

Таблица 4.1.

Результат согласования двоичных помехоустойчивых кодов с кодом Грея

Десятичное число	Натуральный двоичный код	Код Грея
0	000	000
1	001	001
2	010	011
3	011	010
4	100	110
5	101	111
6	110	101
7	111	100

Поскольку ошибки чаще происходят за счет переходов в области соседних сигналов, то кодовые блоки, соответствующие соседним сигналам, должны различаться наименьшим числом двоичных символов.

На рис. 4.3 приведены примеры кода Грея для ансамблей одномерных (АМ-4) и двумерных (ФМ-4, КАМ-16) сигналов.

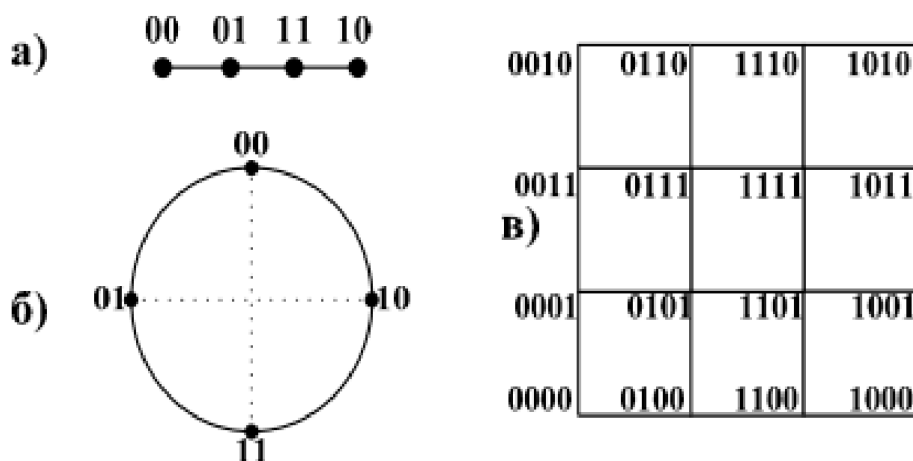


Рис. 4.3. Коды Грея ансамблей: одномерных: а) АМ-4; б) ФМ-4; двумерных: КАМ-16

Несмотря на достаточно высокие показатели энергетической эффективности при мягком решении в демодуляторе и декодировании алгоритмом Витерби, согласование кодом Грея не является оптимальным.

Двоичные коды, оптимальные по критерию максимума хэммингова расстояния, будут оптимальны и по критерию максимума свободного евклидова расстояния, если при отображении двоичных подблоков в сигнальные точки ансамбля выполняется принцип:

большему расстоянию Хэмминга $dh \max$, соответствует большее расстояние по Евклиду $de \max$.

Простейшие ансамбли сигналов АМн-2, ФМн-2, ФМн-4 этому условию для кода Грея удовлетворяют. В табл. 4.4 показаны комбинации (подблоки) двоичного кода длиной $n = 3$, а также расстояния dh и de при использовании кода Грея для ФМн-8 (см. рис. 4.4).

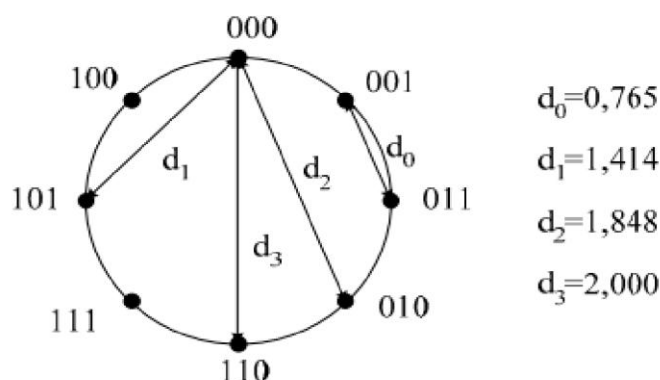


Рис. 4.4. Расстояние между сигнальными точками ФМн-8

Как следует из таблицы, сформулированный принцип соответствия большому расстоянию Хэмминга большому расстоянию Евклида для всех вариантов отображения не выполняется. Например, для комбинации 111 большому расстоянию Хэмминга $dh = 3$ соответствует не самое большое расстояние Евклида $dh = 1,848$, и т.д.

Соответствие расстояний Хэмминга и Евклида для сигналов ФМн-8

Кодовые комбинации	000	001	011	010	110	111	101	100
d_h	0	1	2	1	2	3	2	1
d_e	0	0,765	1,414	1,848	2,000	1,848	1,414	0,765

Таким образом поскольку манипуляционный код Грея для сложных сигналов не обеспечивает оптимального согласования кодека и модема, необходимо найти методы дальнейшего повышения свободного евклидова расстояния def и, соответственно, энергетической эффективности β .

Согласование на основе разбиения ансамбля на вложенные подансамбли

В начале 80 х гг. Унгербок (Ungerboeck G.) опубликовал статью, в которой, анализируя СКК на базе ансамбля ФМн-8 и сверточного кода со скоростью $R_{кк} = k / k + 1$,

сформулировал ряд правил построения СКК. Поэтому СКК построенные по этим правилам (Trellis-Coded Modulation – TCM), часто называют СКК Унгербоека.

По способу согласования модуляции и кодирования СКК Унгербоека относятся к конструкциям, полученным на основе разбиения ансамбля сигналов на вложенные подансамбли. Разбиение осуществляется таким образом, что подансамбли содержат равное количество сигналов, расстояния de между соседними сигналами подансамблей одинаковы, минимальные расстояния de_{\min} между сигналами подансамблей увеличиваются с каждым шагом разбиения; при этом левая ветвь разбиения кодируется символом «0», а правая «1». Считывание кодовой комбинации, соответствующей сигнальной точке на амплитудно-фазовой плоскости, осуществляется снизу вверх. Разбиение для ансамбля сигналов ФМн-8 представлено на рис. 4.5.

Как следует из рис. 4.5, исходный ансамбль разбивается на подансамбли при максимальном увеличении наименьших расстояний de_{\min} между сигналами внутри подансамблей $d_0 < d_1 < d_2 < d_3$. Разбиение осуществляется поэтапно.

В данном примере три этапа, заключающихся в разбиении каждого из подансамблей пределы предыдущего этапа на 2 равноэлементных подансамбля.

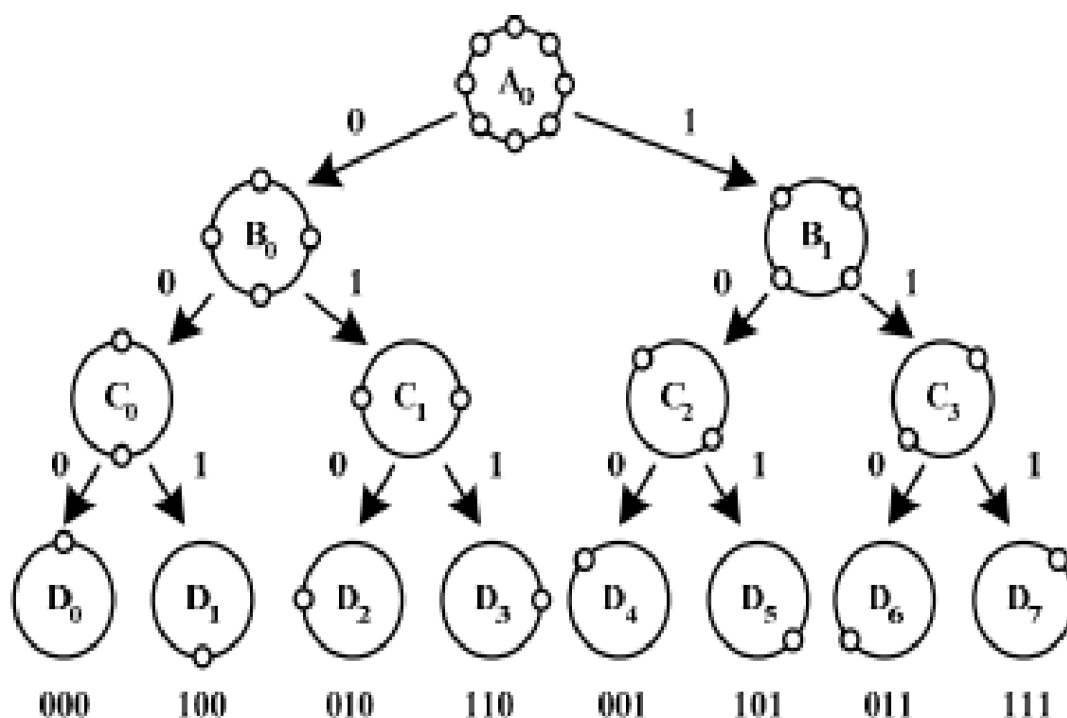


Рис.4.5. Разбиение ансамбля сигналов ФМн-8

В общем случае количество этапов i полного разбиения ансамбля из m сигналов на вложенные подансамбли определяется выражением: $i_{\max} = \log_2 m$, (4.6), т. е. совпадает кратностью ансамбля n . В ансамбле из m сигналов кратности n каждой сигнальной точке соответствует блок двоичных символов $b = [b_{n-1}, b_{n-2}, \dots, b_0]$. Соответствие между

кодовым блоком b и сигнальной точкой на плоскости определяет манипуляционный код.

Достижение наибольшей помехоустойчивости непосредственно связано с увеличением евклидова расстояния между передаваемыми сигнальными последовательностями. Решетчатая диаграмма сверточного кода (5.6.3), ребра которой промаркированы сигнальными точками, полностью отображает весь набор разрешенных сигнальных последовательностей.

Таким образом, величина свободного евклидова расстояния def зависит от маркировки ребер решетчатой диаграммы сигнальными точками (канальными символами).

Унгербок на примере ансамбля сигналов ФМн-8 (см. рис. 4.6) сформулировал четыре необходимых правила маркировки ребер сигнальными точками: все сигнальные точки используемого ансамбля сигналов должны встречаться с одинаковой частотой и с определенной степенью регулярности и симметричности;

переходы из одного и того же состояния соответствуют сигналам из подансамблей B_0 или B_1 ;

переходы в одно и то же состояние соответствуют сигналам из подансамблей B_0 или B_1 ;

параллельные переходы между состояниями соответствуют сигналам из подансамблей C_0 или C_1 , или C_2 , или C_3 . Как показывает анализ, СКК Унгербока имеют несколько более высокие частотно-энергетические характеристики по сравнению с традиционными СКК, при той же сложности реализации. Это определило их бурное внедрение в технике связи. Но известные правила построения СКК Унгербока, хотя и снижают размерность переборной задачи синтеза, но не обеспечивают гарантированное построение СКК с максимальными частотно-энергетическими характеристиками. В то же время, основной целью работ в области синтеза систем сигналов и СКК является поиск таких способов их формирования и обработки, которые при заданных ограничениях на сложность устройств формирования и приема, временные задержки, позволяли бы приблизиться к известной шенноновской границе.

При построении многомерных СКК возникает проблема выбора манипуляционного кода, поскольку известные методы его построения (правила построения кодов Грея и разбиения ансамбля на вложенные подансамбли Унгербока) не всегда позволяют согласовать евклидовы и хемминговы расстояния. Именно с этим связаны многие проблемы построения многомерных СКК.

Синтез многопозиционных ансамблей сигналов и СКК, построенных на их основе, является одним из направлений решения более общей задачи статистического согласования

вероятностных характеристик передаваемого информационного сигнала и вероятностных характеристик канала. В рамках этих традиционных задач, такое согласование осуществляется на уровне канальных символов или их блоков (супербукв канала). При этом подходы к построению алфавита таких супербукв (ансамблей сигналов и СКК) могут существенно отличаться между собой, но направлены на решение этой общей проблемы.

Известно, что ансамбль сигналов, соответствующий полному двоичному коду длины n в пространстве соответствующей размерности n , построенный заменой «1» на «-1», а «0» на «+1», соответственно, обладает практически идеальным манипуляционным кодом. Минимальным хемминговым расстояниям таких ансамблей соответствуют ребра n -мерного куба, которые характеризуются и минимальными евклидовыми расстояниями.

Кодовые комбинации и соответствующие им координаты сигнальных векторов приведены в табл. 4.2; графическое изображение ансамбля представлено на рис. 9.13. При приеме сигналов такого ансамбля минимальная ошибка (ошибочный прием одной координаты сигнальной точки) приводит к неправильному приему одного бита информации. Ошибочный прием двух координат сигнальной точки приводит к искажению двух бит информации и так далее. Однако, если рассмотреть зависимость между хемминговыми $dh(i, j)$ и евклидовыми $de(i, j)$ расстояниями для такого ансамбля, то можно выявить следующую закономерность, связывающую эти две величины :

$$de(i, j) = 2r \sqrt{dh(i, j)}, \quad (4.2)$$

где r – радиус сферы.

Таблица 4.2

Взаимосвязь кодовых комбинаций манипуляционного кода и координат сигнальных векторов

Манипуляционный код	Координаты сигнальных векторов
000	+1,+1,+1
001	+1,+1,-1
010	+1,-1,+1
011	+1,-1,-1
100	-1,+1,+1
101	-1,+1,-1
110	-1,-1,+1
111	-1,-1,-1

Таким образом, взаимосвязь между евклидовыми и хемминговыми расстояниями в многомерном ансамбле сигналов нелинейная, хотя большему хемминговому расстоянию

будет соответствовать большее евклидово расстояние. Если мощность и энергия сигналов являются постоянными величинами, не зависящими от номера, то ансамбли таких сигналов считают сигналами поверхностно-сферической упаковки.

В противном случае ансамбли сигналов рассматривают как объемные упаковки. Сохранение манипуляционного кода, принятого для простого трехмерного куба, в значительной мере сохраняет пропорциональность между евклидовыми и хемминговыми расстояниями и поэтому будет наилучшим и для наиболее плотного ансамбля. Для других комбинаций манипуляционных кодов для сигнальных векторов изначально не будет соблюдаться взаимная пропорциональность между евклидовыми и хемминговыми расстояниями.

Таким образом, практически невозможно создать идеальный манипуляционный код и, следовательно, целесообразно строить манипуляционные коды, у которых хотя бы частично выполняется взаимосвязь между евклидовыми и хемминговыми расстояниями.

Практическая часть работы в MATLAB 2015 [21]

Перед началом выполнения работы, необходимо с папки TCM скопировать код в командную строку, который подгрузит необходимую схему для треллис модуляции.

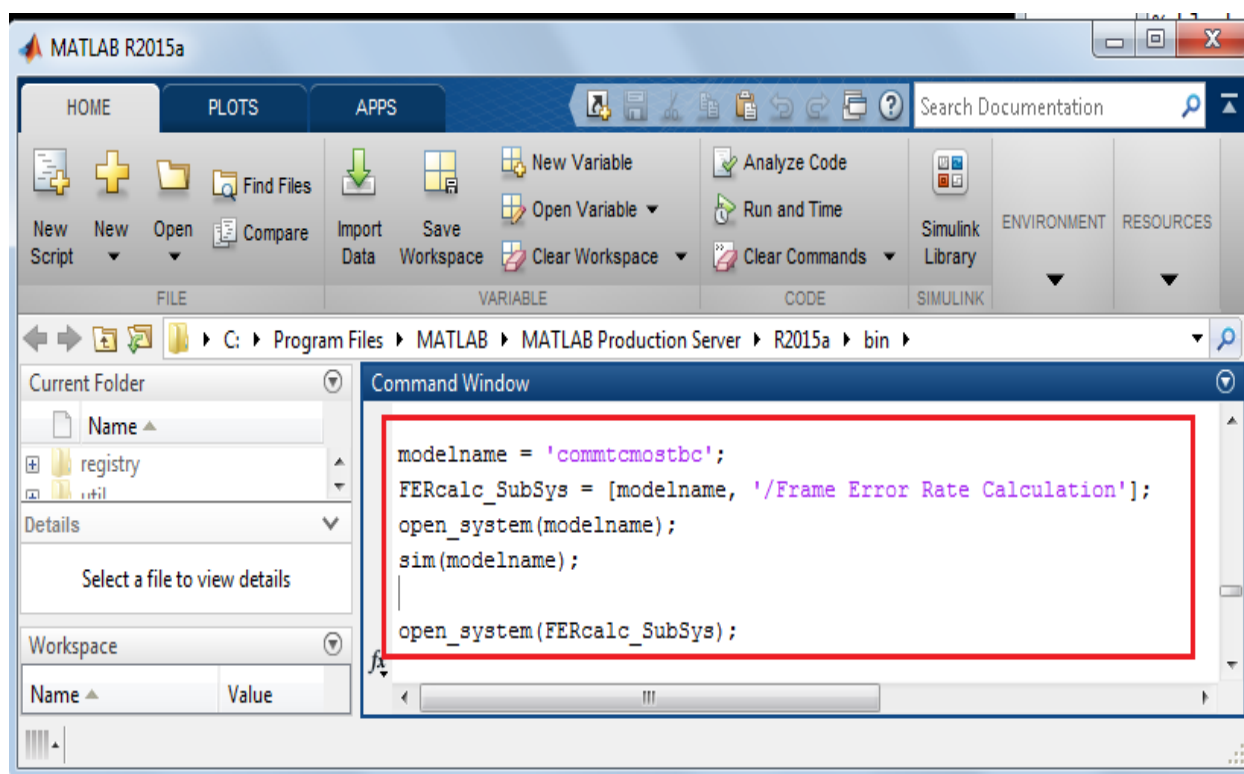


Рис. 4.6. Панель MATLAB 2015

После, откроется следующая схема:

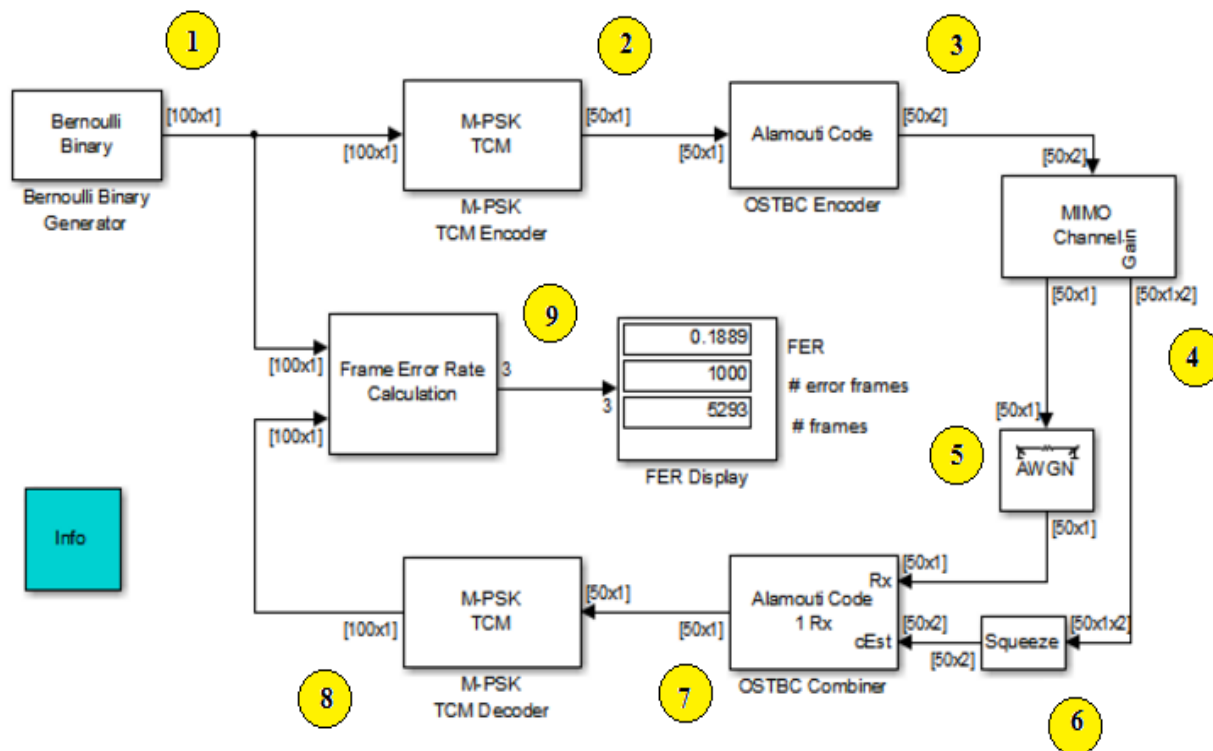


Рис. 4.7. Схема для исследования TCM в Simulink MATLAB 2015

1. Блок Двоичный генератор Бернулли генерирует случайные двоичные числа, используя распределение Бернулли. Выставьте ниже представленные параметры:

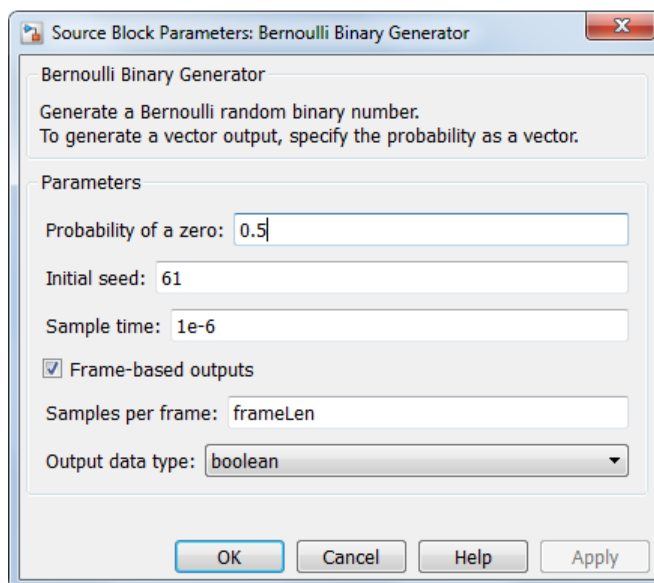


Рис. 4.8. Параметры источника (Source Block Parameters)

2. Блок M-PSK TCM кодер преобразовывает данные, полученные от блока Бернулли в PSK созвездия с заданной его средней энергией. В этом примере мы используем схему TCM 8 - PSK созвездия для 8 решетчатых состояний. Выставьте ниже представленные параметры:

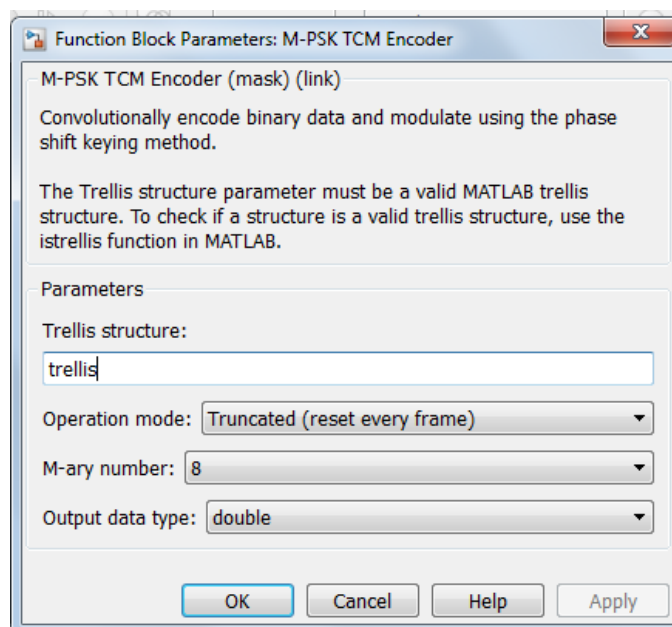


Рис. 4.9. Параметры блока: M-PSK TCM Encoder

После 2го блока нужно добавить диаграмму созвездий, который расположен в панели программы Matlab. Выставьте ниже представленные параметры:

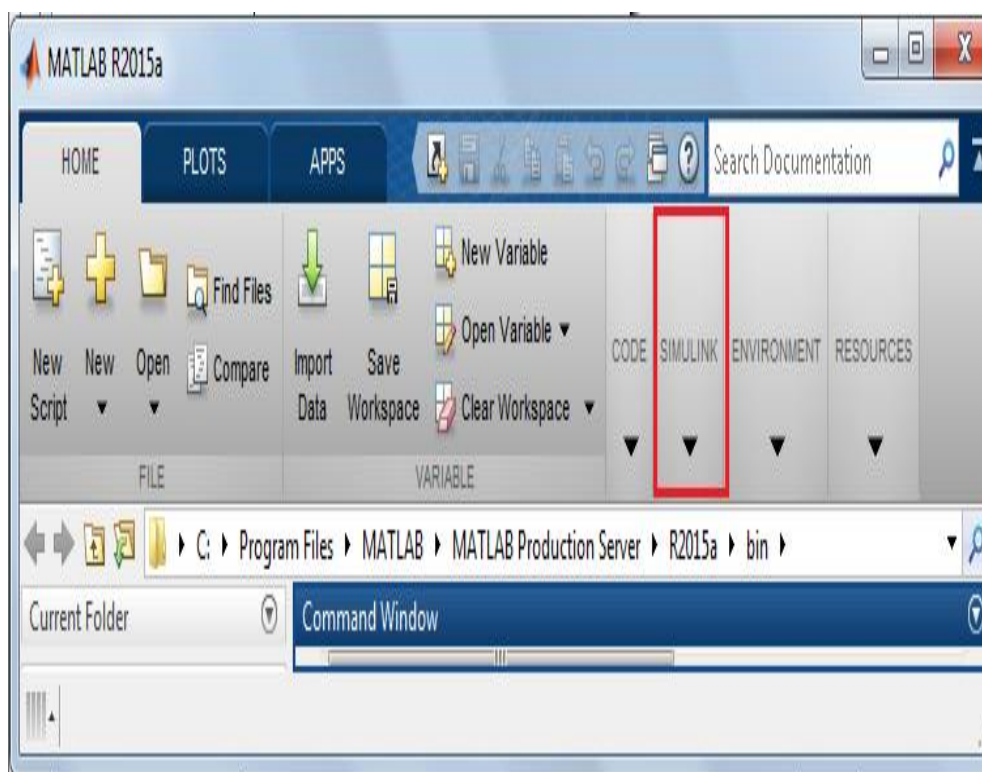


Рис. 4.10. Панель MATLAB

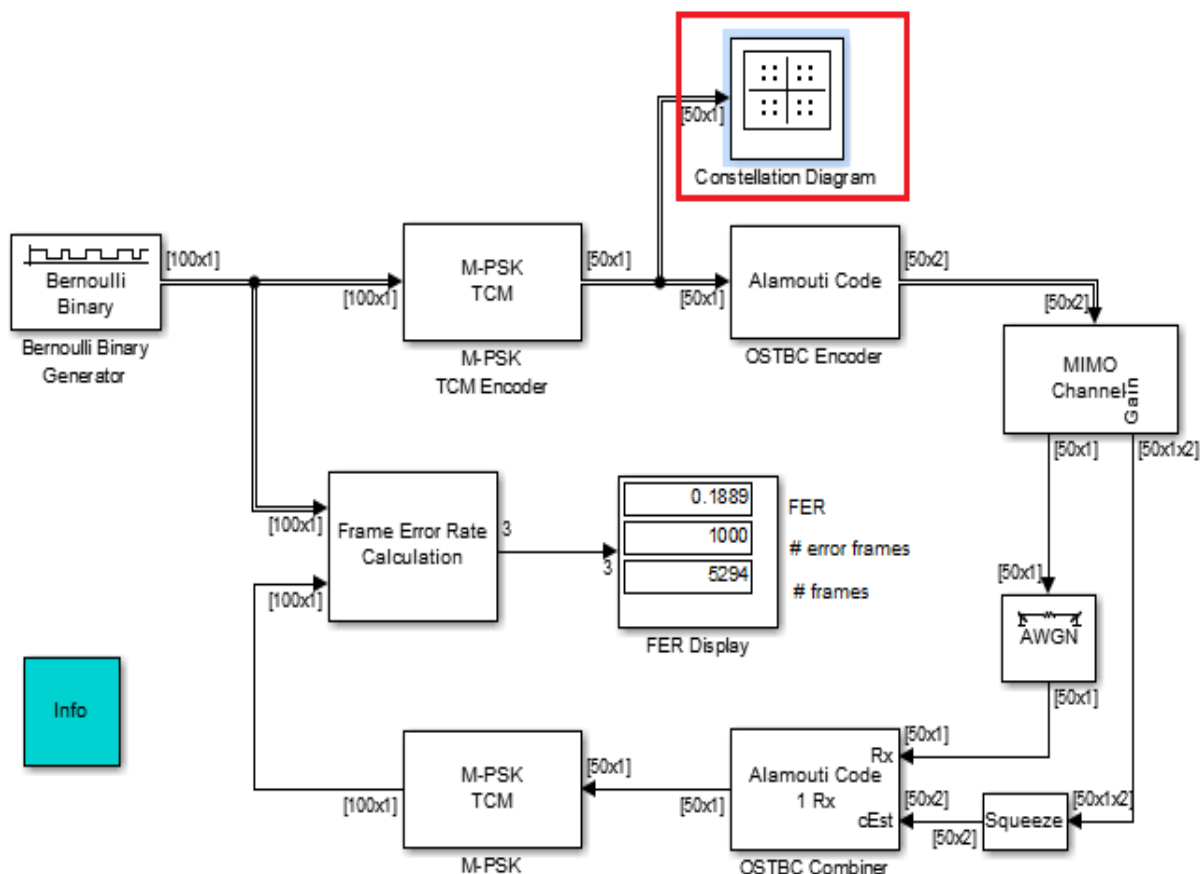


Рис. 4.11. Схема для исследования TCM в Simulink MATLAB 2015

3. Блок OSTBC (ортогонально пространственно-временные блочные коды) кодирует информационные символы из TCM кодировщика, используя код Alamouti для 2х передающих антенн. Выходом этого блока является матрица размером 50×2 , элементы которой соответствуют данным колонки, передаваемых по одной антенне.

Блок OSTBC сочетает в себе полученные сигналы от приемной антенны с информацией о состоянии канала (CSI), которые затем подают в M-PSK TCM декодера. CSI известно на стороне приемника. Выставьте ниже представленные параметры:

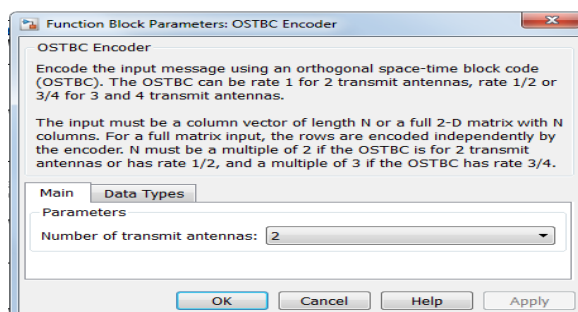


Рис. 4.12. Параметры блока: OSTBC Encoder

4. MIMO канал разносит передающие и приёмные антенны так, чтобы корреляция между соседними антеннами была слабой. Выставьте ниже представленные параметры:

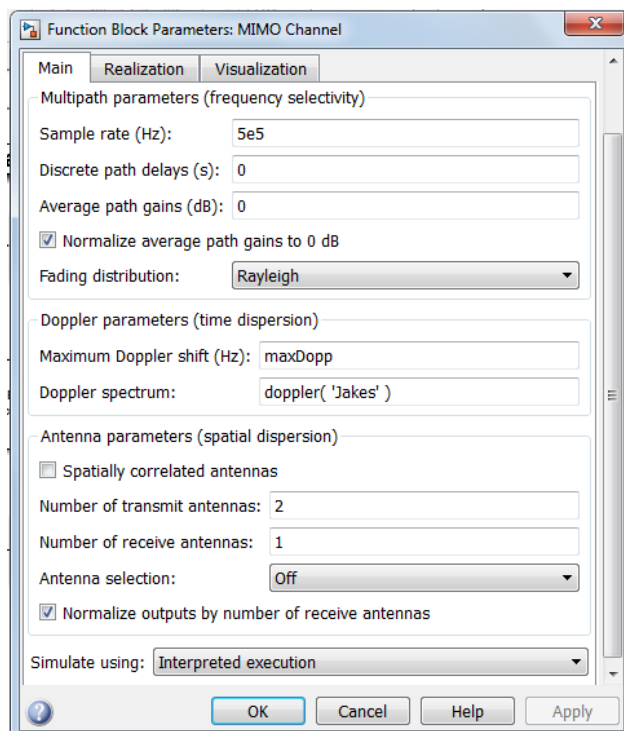


Рис. 4.13. Параметры блока: MIMO Channel

5. Блок AWGN добавляет белый гауссовский шум на приемной стороне. Выставьте ниже представленные параметры:

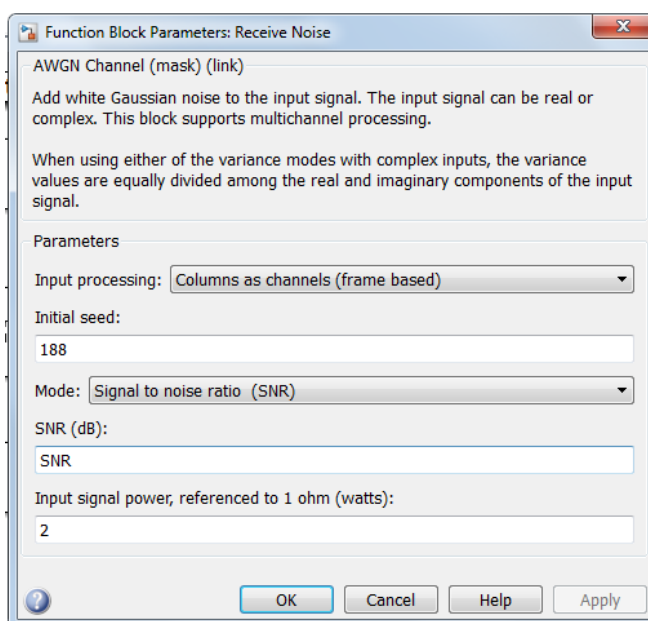


Рис. 4.14. Параметры блока: Receive Noise

6. Блок Сжатия. Выставьте ниже представленные параметры:

7. Блок OSTBC дает оценку канала сигнала. Выставьте ниже представленные параметры:

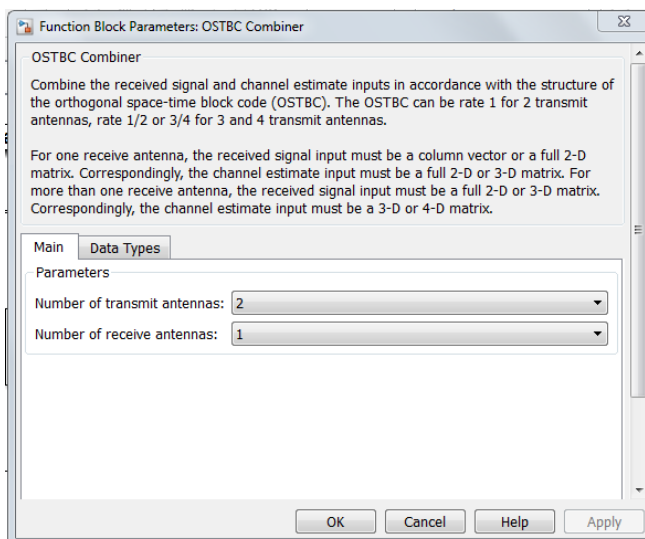


Рис. 4.15. Параметры блока: OSTBC Combiner

8. Блок M-PSK TCM декодер выполняет декодирование сигнала ранее модулированного с использованием PSK созвездия на входе. Для декодирования используется алгоритм Витерби. Выставьте ниже представленные параметры:

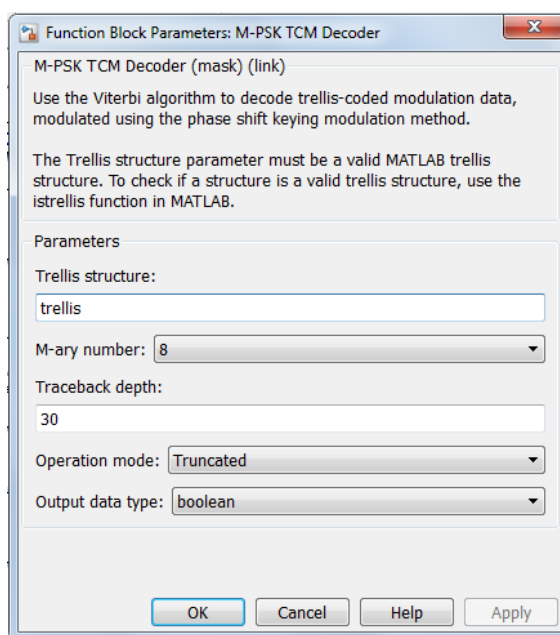


Рис. 4.16. Параметры блока: M-PSK TCM Decoder

9. Блок FER Display сравнивает декодированные биты с исходными битами в кадре для обнаружения ошибок и в онлайн режиме обновляет показатели. Этот блок состоит из трех показателей: количество ошибок в кадре, количество наблюдаемых кадров и количество обработанных ошибок в кадре. Выставьте ниже представленные параметры:

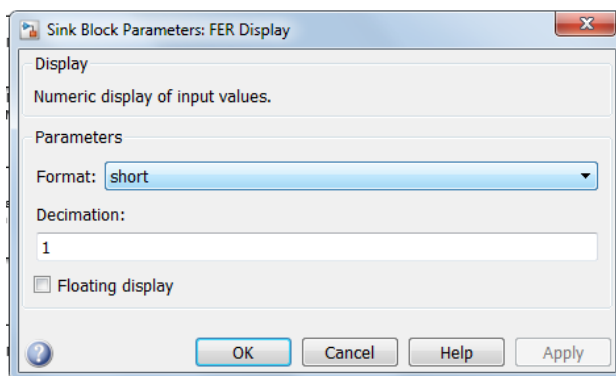


Рис. 4.17. Параметры блока: FER Display

Таблица 4.3 – Результаты работы

		0,9	0,8	0,6	0,2	0,1	0,0	
FER	1	94	873	289	813	88	8	0,038
SNR	0	2	4	6	8	10	12	14

Таблица 4.4 – Результаты работы

		0,9	0,8	0,6	0,2	0,1	0,0	0,0
FER	1	95	85	1	7	89	8	45
SNR	0	2	4	6	8	10	12	14

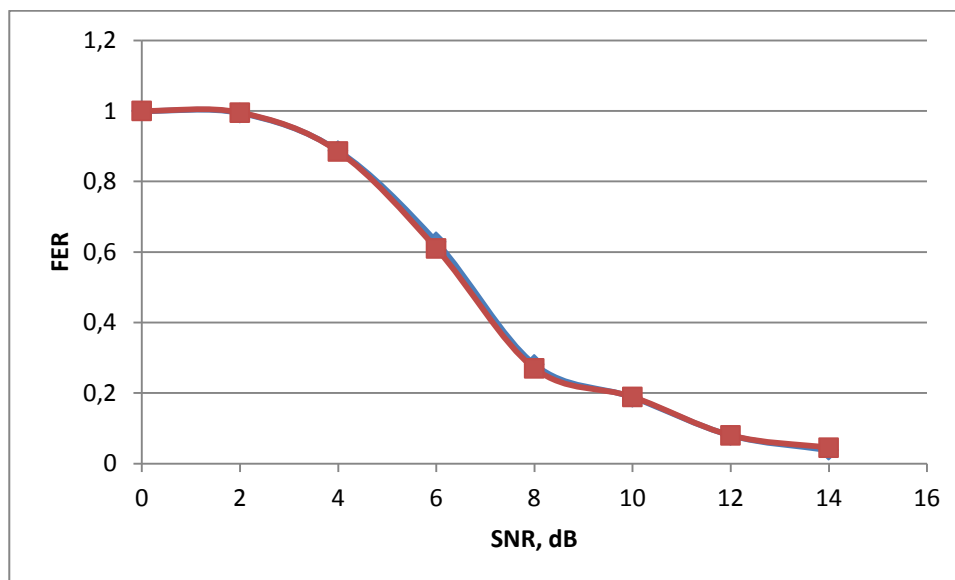


Рис. 4.18. Зависимость FER от отношения сигнал/шум

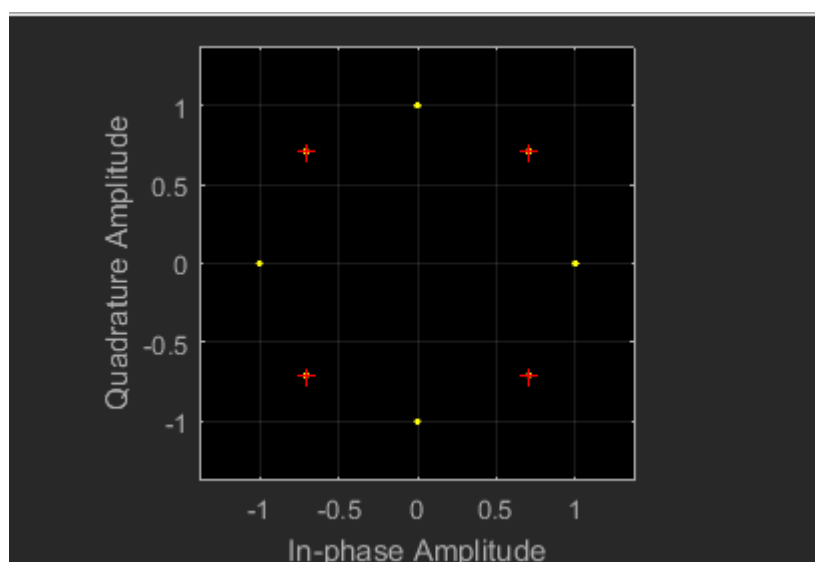


Рис. 4.19. Диаграмма созвездия 8-PSK

Таким образом, в данной работе на примере сверточного кодирования были рассмотрены преимущество и технология применения треллис-модуляции для повышения скорости передачи данных.

Переход в режим треллис-модуляции позволил сохранить тот же уровень помехоустойчивости (3 дБ) при увеличении скорости передачи информации в 2 раза. Это достигается с помощью совместного использования сверточного кода со скоростью $2/3$ и фазовой манипуляции 8PSK. Использование треллис-модуляции является простым и эффективным решением для повышения скорости передачи в цифровых высокоскоростных системах связи.

4.2. Исследование сигнально-кодовой конструкции на базе системы с ортогональным частотным мультиплексированием и пространственно-временным кодированием OFDM - MIMO

В последнее время все активнее ощущается рост беспроводных систем связи. Развитие технологий мобильных устройств, беспроводных локальных сетей (WLAN) и стремительный рост Интернет вызывают все возрастающую потребность в увеличении емкости мобильных сетей. Также наблюдается все большая интеграция сотовых сетей с сетями передачи данных, например GPRS в GSM сетях, а также сети 3G. Однако существующие технологии не могут удовлетворить новых потребностей по емкости сети, скорости передачи и стоимости услуг.

OFDM - ортогональное частотное мультиплексирование – это схема модуляции, которая позволяет быстро и эффективно передавать данные даже в каналах с многолучевым распространением сигнала. Передача ведется одновременно на большом количестве несущих частот. Эти несущие имеют небольшое разнесение по частоте и их спектры образуют групповой спектр OFDM сигнала. Частотное разнесение и синхронизация подобраны так,

чтобы несущие были ортогональны между собой, то есть не оказывали влияния друг на друга, несмотря на перекрытие по спектру. Другой важный подход, позволяющий существенно улучшить, по сравнению с традиционными системами, спектральную эффективность и помехоустойчивость системы - применение разнесения как на передающей, так и на приемной стороне путем использования нескольких передающих и нескольких приемных антенн (MIMO). Эта технология уже применяется в некоторых современных стандартах, например IEEE 802.16 и LTE.

Особенности распространения сигнала

В мобильной радиосвязи параметры канала изменяются во времени, поскольку перемещение абонента в пространстве приводит к изменению условий распространения сигнала. Скорость изменения условий распространения определяет скорость замираний. В большинстве случаев мобильный канал связи характеризуется отсутствием прямой видимости между передатчиком и приемником, особенно в условиях плотной городской застройки. Если имеется большое число многократно отраженных лучей и отсутствует прямая видимость, огибающая полученного сигнала может быть статистически описана с помощью релевской функции плотности вероятности. Характер замираний в этом случае называют релевским. На Рис. 4.20 показан многолучевой характер распространения сигнала вследствие отражений.

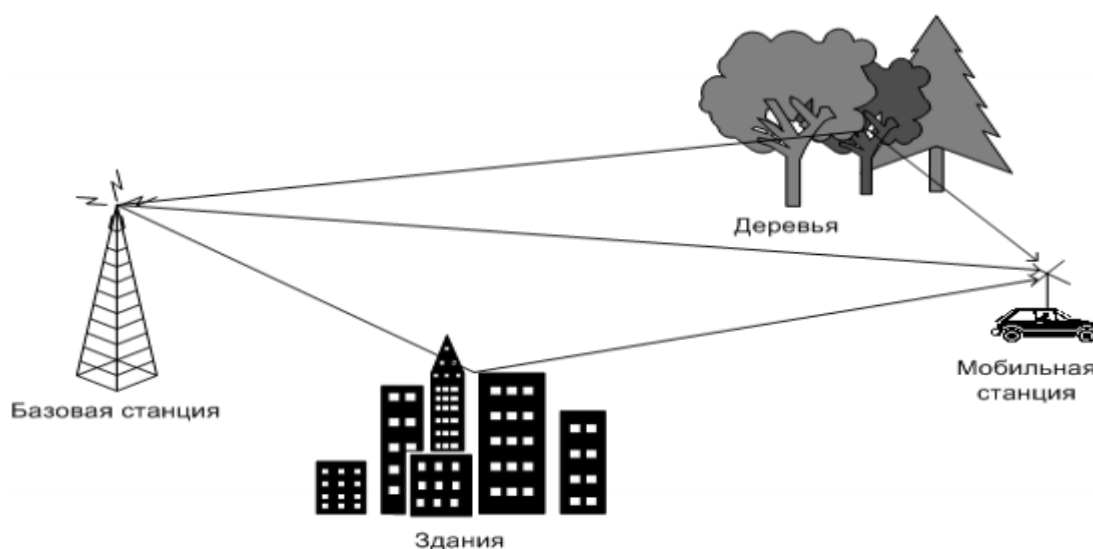


Рис. 4.20. Многолучевой характер распространения сигнала в мобильном канале связи.

В мобильном канале связи сигнал подвергается воздействию следующих основных типов искажений:

- Многолучевое распространение возникает в результате отражения, рассеяния и дифракции электромагнитных волн при взаимодействии с различными объектами в

пространстве. Таким образом, сигнал в приемной антенне содержит сумму волн с различными задержками, амплитудами и фазами. Суперпозиция этих волн приводит к

изменению амплитуды и фазы принимаемого сигнала.

- Допплеровский сдвиг возникает из-за перемещения абонента в пространстве. Возникает изменение амплитуды и фазы принимаемого сигнала во времени. Даже небольшие перемещения на расстояния, соизмеримые с длиной волны передаваемого сигнала, могут вызывать существенные изменения параметров принимаемого сигнала. Изменения параметров сигнала во времени, вызванные движением абонента и многолучевым распространением радиоволн называются быстрыми замираниями. Величина доплеровского сдвига пропорциональна частоте передачи и скорости движения. Чем меньше разнесение между несущими в сигнале OFDM, тем более восприимчива система к доплеровскому сдвигу частоты. Различные режимы передачи позволяют получить компромисс между уровнем восприимчивости к межсимвольной интерференции и доплеровскому сдвигу частоты.

- Затенение вызывается объектами, такими как здания, холмы, деревья и т.п., оказывающимися на пути сигнала и ограничивающими прямую видимость между передатчиком и приемником. Изменения параметров сигнала во времени, вызванные затенением обычно относят к медленным замираниям.

- Потери в тракте характеризуются как зависимость падения средней мощности сигнала от расстояния между передатчиком и приемником. В открытом пространстве средняя мощность уменьшается пропорционально квадрату расстояния между передатчиком и приемником. В мобильном радиоканале, где прямая видимость часто отсутствует, эта характеристика часто представляется как функция степени от 3 до 5.

Искажения, вызванные потерями в тракте и затенением (медленные замирания) обычно компенсируются с помощью систем управления мощностью. Борьба с быстрыми замираниями, вызванными движением и многолучевым распространением волн, является более сложной задачей и требует применения сложной обработки сигнала как на приемной, так и на передающей стороне.

Следствием многолучевого распространения радиоволн является искажение формы принимаемого сигнала. Многолучевая интерференция присуща любому типу сигналов, но особенно негативно она сказывается на широкополосных сигналах. Дело в том, что при использовании широкополосного сигнала в результате интерференции определенные частоты складываются синфазно, что приводит к увеличению уровня сигнала, а некоторые, наоборот, противофазно, вызывая ослабление сигнала на данной частоте.

Говоря о многолучевой интерференции, возникающей при передаче сигналов, различают два крайних случая. В первом из них максимальная задержка между различными сигналами не превышает времени длительности одного символа и интерференция возникает в пределах одного передаваемого символа. Во втором случае максимальная задержка между различными сигналами больше длительности одного символа, в результате интерференции складываются сигналы, представляющие разные символы, и возникает так называемая межсимвольная интерференция (InterSymbol Interference, ISI).

Межсимвольная интерференция наиболее существенно влияет на искажение сигнала. Поскольку символ - это дискретное состояние сигнала, характеризующееся значениями частоты несущей, амплитуды и фазы, то для различных символов меняются амплитуда и фаза сигнала, а значит, восстановить исходный сигнал крайне сложно. Чтобы хотя бы частично компенсировать эффект многолучевого распространения, используются частотные эквалайзеры, однако по мере роста скорости передачи данных либо вследствие увеличения символьной скорости, либо из-за усложнения схемы кодирования эффективность использования эквалайзеров падает.

В традиционных системах с одной несущей борьба с межсимвольной интерференцией обычно ведется путем адаптивного выравнивания. Этот процесс использует адаптивную фильтрацию для аппроксимации импульсного отклика канала. Затем инверсный фильтр используется для воссоздания копий искаженных символов. Этот процесс, однако, довольно сложный, ввиду высокой сложности адаптивного эквалайзера. В случаях, когда межсимвольная интерференция становится довольно высокой, процесс также теряет эффективность.

Модель системы связи с ортогональным частотным мультиплексированием (OFDM)

При высоких скоростях передачи применяется метод передачи данных, который состоит в том, что поток передаваемых данных распределяется по множеству частотных подканалов и передача ведется параллельно на всех этих подканалах. При этом высокая скорость передачи достигается именно за счет одновременной передачи данных по всем каналам, а скорость передачи в отдельном подканале вполне может быть невысокой. Поскольку в каждом из частотных подканалов скорость передачи можно сделать не слишком высокой, это создает предпосылки для эффективного подавления межсимвольной интерференции.

При частотном разделении каналов необходимо, чтобы ширина отдельного канала была, с одной стороны, достаточно узкой для минимизации искажения сигнала в пределах

отдельного канала, а с другой - достаточно широкой для обеспечения требуемой скорости передачи. Кроме того, для экономного использования всей полосы канала, разделяемого на подканалы, желательно как можно более плотно расположить частотные подканалы, но при этом избежать межканальной интерференции, чтобы обеспечить полную независимость каналов друг от друга. Частотные каналы, удовлетворяющие перечисленным требованиям, называются ортогональными. Несущие сигналы всех частотных подканалов (а точнее, функции, описывающие эти сигналы) ортогональны друг другу. Важно, что хотя сами частотные подканалы могут частично перекрывать друг друга, ортогональность несущих сигналов гарантирует независимость каналов друг от друга, а следовательно, и отсутствие межканальной интерференции.

Рассмотренный способ деления широкополосного канала на ортогональные частотные подканалы называется ортогональным частотным мультиплексированием (OFDM). Сигнал в системе с ортогональным частотным мультиплексированием имеет разбиение на множество несущих, что обеспечивает небольшое количество символов на одну несущую и снижает межсимвольную интерференцию. На рисунке 4.21 представлен спектр сигнала при модуляции OFDM.

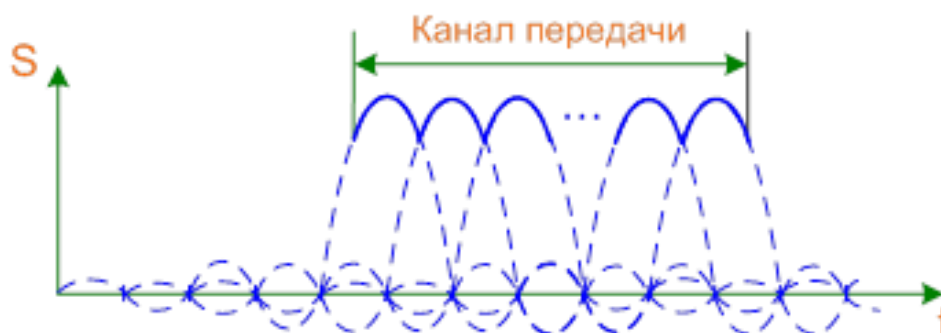


Рис. 4.21. Спектр сигнала при модуляции OFDM

Дополнительно применяется защитный интервал - циклический префикс, добавляемый в начало каждого символа. Для эффективной работы системы, использующей такой подход, максимальная задержка в канале не должна превышать длину циклического префикса. Высокая эффективность систем OFDM при работе в каналах с многократными отражениями делает их пригодными для высокоскоростных систем передачи данных в наземных системах связи.

На рисунке 4.22 представлена структура OFDM символа.

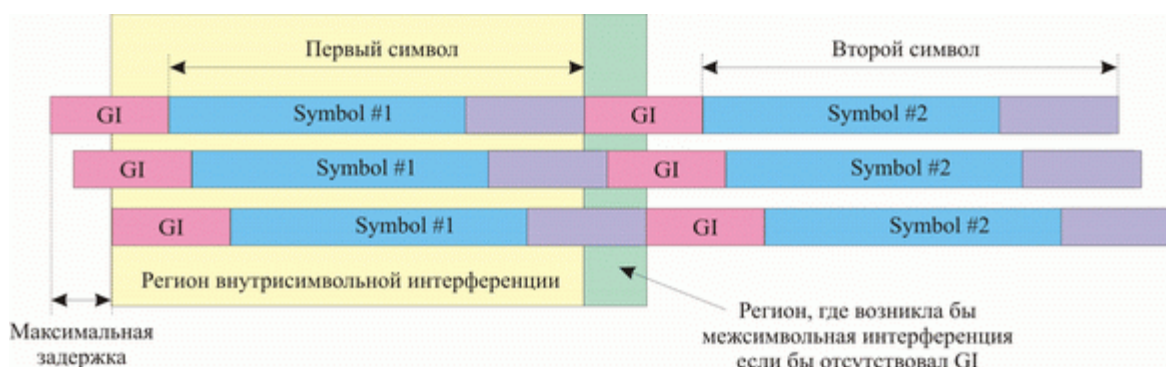


Рис. 4.22. Структура OFDM символа.

На рисунке 4.24 показана обобщенная структура системы связи с ортогональным частотным мультиплексированием. Данные пользователя вначале поступают в блок помехоустойчивого кодирования, а затем на модулятор. После модуляции данные в виде комплексных символов поступают в блок Обратного быстрого преобразования Фурье (ОБПФ), где происходит формирование сигнала OFDM с использованием M поднесущих. На рисунке 4.23 представлена временная форма сигнала после прохождения ОБПФ.

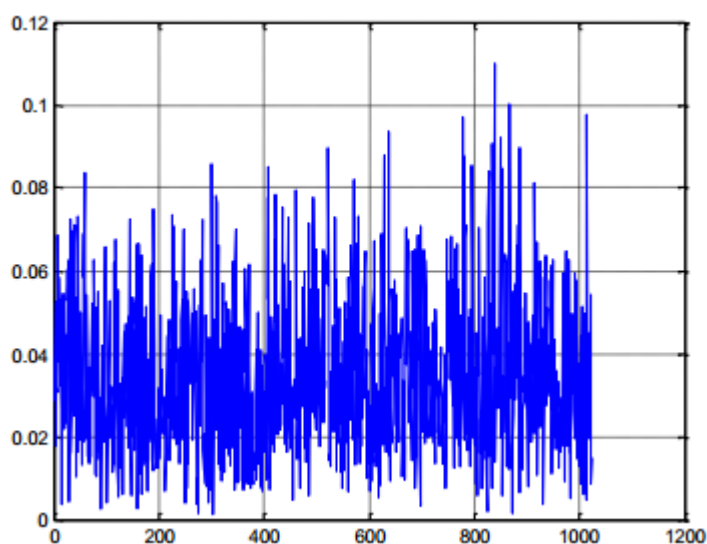


Рис. 4.23. Временная форма сигнала

Символ OFDM состоит из основной информационной части и циклического префикса, который формируется путем копирования последних L отсчетов в начало кадра. Далее последовательность символов OFDM преобразуется в аналоговый сигнал и передается по каналу связи. Длительность циклического префикса выбирается таким образом, чтобы быть больше, чем длительность импульсного отклика канала связи. Таким образом, символ OFDM имеет длительность $(M+L)T_0$, где T_0 - период дискретизации в системе.

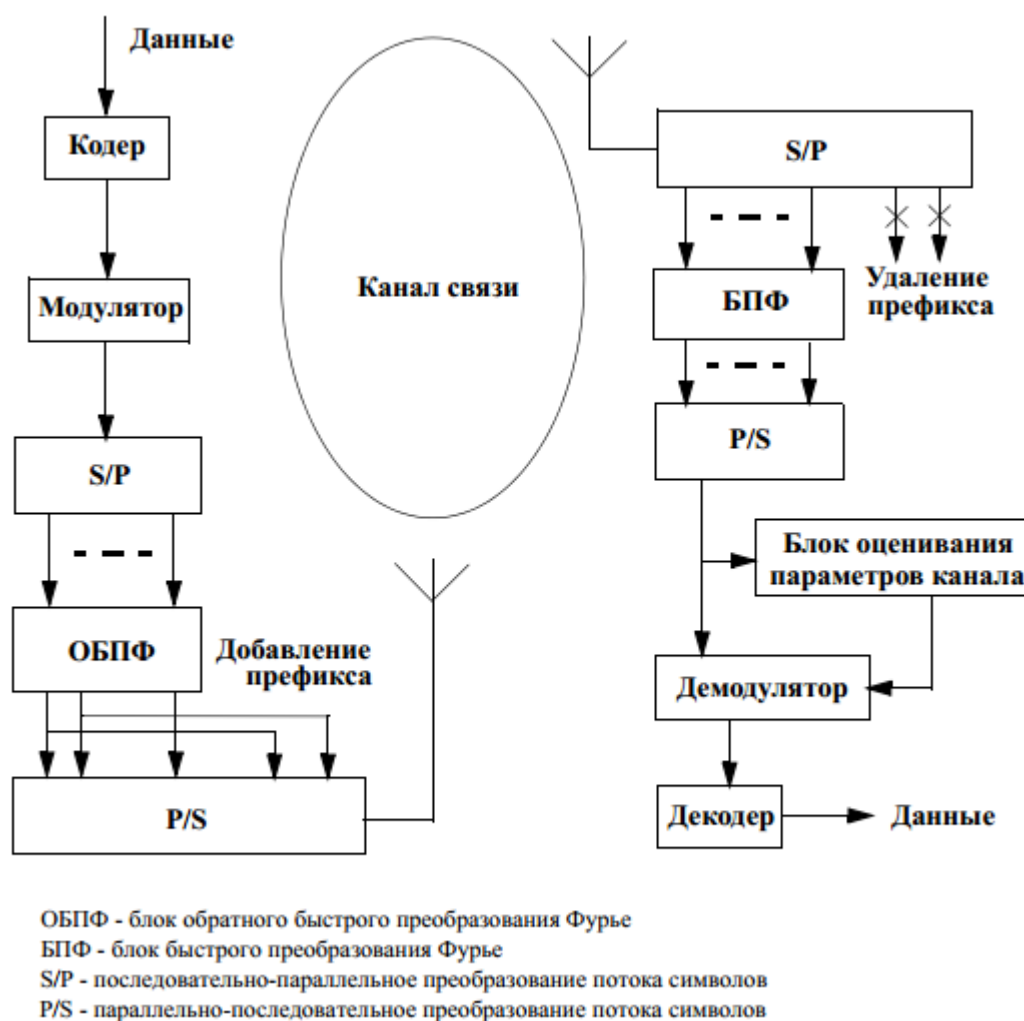


Рис. 4.24. Структурная схема системы связи с ортогональным

частотным мультиплексированием

На приемной стороне, сигнал после дискретизации и удаления префикса поступает в блок Быстрого преобразования Фурье (БПФ), где осуществляется демодуляция сигнала OFDM. После преобразования параллельного потока символов в последовательный, символы поступают в демодулятор, а также в блок оценивания параметров канала. Оценки параметров канала необходимы для работы демодулятора, поэтому выход блока оценивания параметров канала соединен с демодулятором. Дальнейшие преобразования сигнала зависят от структуры конкретной системы и не относятся к основной концепции OFDM. В реальных условиях работы системам беспроводной подвижной связи с FDM приходится работать в условиях, когда в канале связи сигнал подвергается воздействию замираний. Поэтому для получения приемлемой помехоустойчивости требуется использование помехоустойчивого кодирования.

В процессе развития сетей передачи данных постоянно увеличиваются требования к скорости передачи информации и качеству предоставляемых сервисов. Данный механизм можно наблюдать и для сетей передачи данных, где средой передачи является радиоканал.

Из-за этого в значительной степени усложняются сигналы, которые используют для передачи информации. Одним из перспективных видов является ортогональное частотное мультиплексирование OFDM.

В современных системах связи, например, в сотовых системах связи, высокоскоростных локальных вычислительных сетях и др., существует необходимость повышения пропускной способности. Пропускная способность может быть увеличена путём расширения полосы частот или повышения излучаемой мощности. Тем не менее, применимость этих методов ограничена из-за требований биологической защиты, ограниченной мощности источника питания (в мобильных устройствах) и электромагнитной совместимости. Поэтому если в системах связи эти подходы не обеспечивают необходимую скорость передачи данных, то эффективным может оказаться применение адаптивных антенных решёток со слабо коррелированными антенными элементами. Системы связи с такими антеннами получили название систем MIMO.

Технология OFDM

OFDM основан на разделении потока входных данных на множество параллельных потоков, каждый из которых передается на своей несущей (ортогональной) частоте. Это обеспечивает высокую скорость и помехоустойчивость передачи информации, в частности, по отношению к провалам в спектре передаваемых сигналов, так как узкополосное затухание может исключить только одну или несколько несущих частот из их большого числа (сотни - тысячи). Поскольку модуляция OFDM использует для передачи ортогональные несущие колебания, то возможна демодуляция модулированных сигналов даже в условиях частичного перекрытия полос отдельных несущих. Наличие большого числа несущих не позволяет реализовать модуляцию OFDM непосредственно, т.е. с использованием нескольких тысяч синтезаторов несущих колебаний и нескольких тысяч модуляторов. Поэтому для уменьшения объема оборудования учитывают, что модуляция OFDM представляет собой обратное преобразование Фурье, а демодуляция прямое преобразование Фурье, и применяют быстрые алгоритмы двух этих преобразований, допускающие более простую аппаратную реализацию по сравнению с непосредственной реализацией алгоритмов модуляции OFDM. Модуляция OFDM используется в системах цифрового телевидения, системах сотовой связи WiMAX, MobileWiMAX, MBWA, автоматизированных системах контроля и учета электроэнергии, системах типа "интеллектуальный дом" и др. На ней базируются стандарты беспроводной связи IEEE 802.11a,e,g,n; 802.16a,d,e; 802.20.

Ортогональное частотное разделение каналов

При беспроводной передаче сигналов один и тот же сигнал в результате многократных отражений может поступать в приемник различными путями. Поэтому в точке приема результирующий сигнал представляет собой суперпозицию (интерференцию) многих сигналов с различными амплитудами и начальными фазами. Применительно к многолучевой интерференции, возникающей при передаче сигналов, различают два крайних случая. В первом случае максимальная задержка между различными сигналами не превышает длительности одного символа, и интерференция возникает в пределах одного передаваемого символа. Во втором случае максимальная задержка между различными сигналами больше длительности одного символа, и в результате интерференции складываются сигналы, представляющие разные символы. Вследствие этого возникает межсимвольная интерференция, которая наиболее сильно сказывается на искажении сигнала. Для того, чтобы частично компенсировать эффект многолучевого распространения, применяют частотные эквалайзеры, однако по мере роста скорости передачи данных либо за счет увеличения символьной скорости, либо за счет усложнения схемы кодирования, эффективность их применения падает. Поэтому для достижения высокой скорости передачи данных используют другой подход, состоящий в том, что поток передаваемых данных распределяется по множеству частотных подканалов и передача ведется параллельно на всех этих подканалах. При этом достигается высокая скорость передачи за счет одновременной передачи данных по всем каналам, причем скорость передачи в отдельном подканале может быть и невысокой. Это создает предпосылки для эффективного подавления межсимвольной интерференции. При частотном разделении каналов необходимо, чтобы ширина каждого канала была, с одной стороны, достаточно узкой для минимизации искажения сигнала β в его пределах, а с другой - достаточно широкой для обеспечения требуемой скорости передачи. Кроме того, для экономного использования всей полосы канала, разделяемого на подканалы, желательно как можно плотнее расположить частотные подканалы, но при этом избежать межканальной интерференции для обеспечения полной независимости каналов друг от друга. Перечисленным требованиям удовлетворяют ортогональные частотные каналы. Функции, описывающие несущие сигналы всех этих каналов, ортогональны друг другу, т.е. для них выполняется условие:

$$\int_0^T \sin 2\pi f_l(t) \cdot \sin 2\pi f_m(t) dt = 0, \quad k \neq l$$

где T - длительность передаваемого символа,
 f_l и f_k - частоты l -го и k -го несущих сигналов соответственно.

Ортогональность несущих сигналов обеспечивает частотную независимость каналов друг от друга и, следовательно, отсутствие межканальной интерференции. Рассмотренный способ деления широкополосного канала на ортогональные частотные подканалы называется ортогональным частотным разделением с мультиплексированием или OFDM-модуляцией.

Ортогональность несущих сигналов обеспечивается только тогда, когда за время длительности одного символа T несущий сигнал будет совершать целое число колебаний. Так как каждый символ длительности T передается ограниченной по времени синусоидальной функцией (рис. 2.1.1), то ее спектр описывается функцией вида

$$\frac{\sin 2\pi(f - f_i)}{2\pi(f - f_i)}$$

где f_i - центральная (несущая) частота i -го канала.

Такой же функцией описывается и форма частотного подканала.

Несмотря на частичное перекрытие частотными подканалами друг друга (рис.1.1), ортогональность несущих сигналов обеспечивает их частотную независимость каналов друг от друга и, следовательно, отсутствие межканальной интерференции.

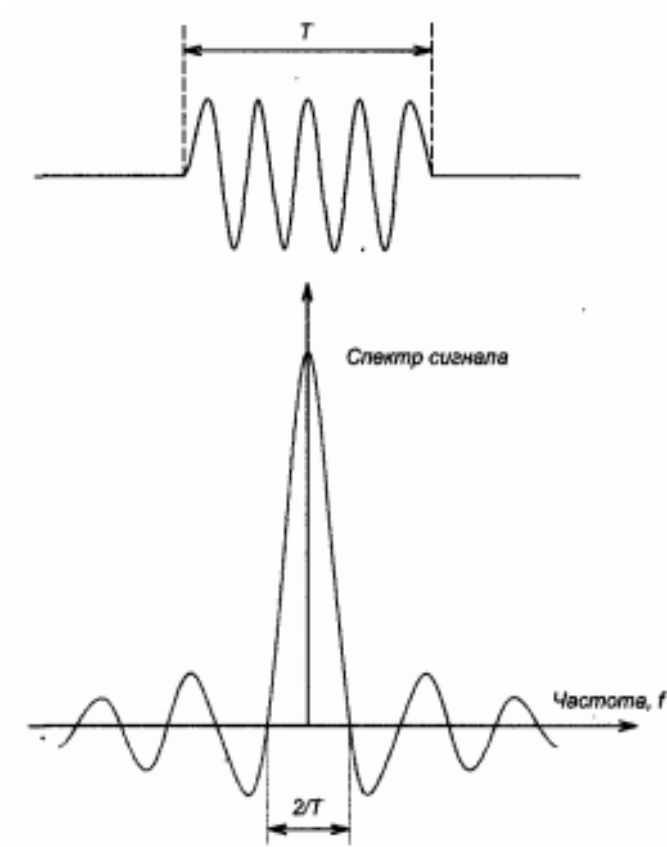


Рисунок 4.5 - Символ длительностью T и его спектральное представление.

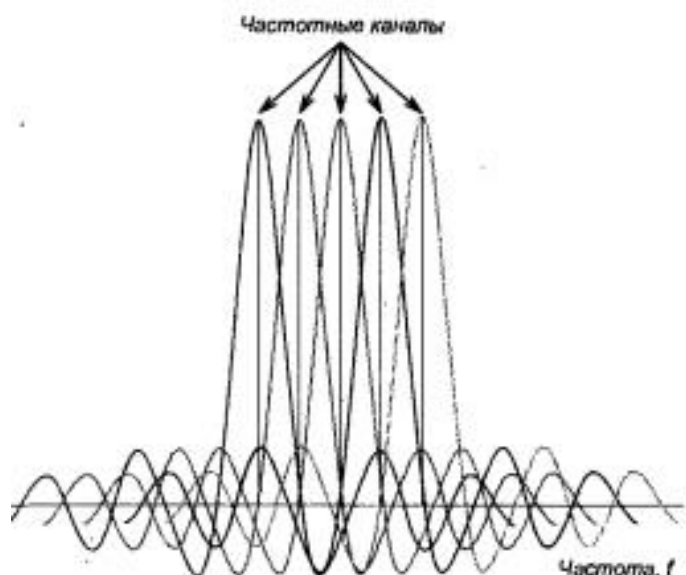


Рисунок 4.26 - Частотное разделение каналов с ортогональными несущими сигналами.

Принцип организации канала OFDM на примере DVB-T

Характеристики канала передачи, к сожалению, не остаются постоянными во времени, но в течение короткого промежутка времени эти характеристики для наземного канала можно считать постоянными. Используя эту особенность, в системе OFDM имеется возможность применить расщепление наземного канала передачи во времени и по частоте (см. рис. 1.3). В результате радиочастотный канал организуется в виде набора узких частотных полос и в виде коротких во времени смежных «временных сегментов»

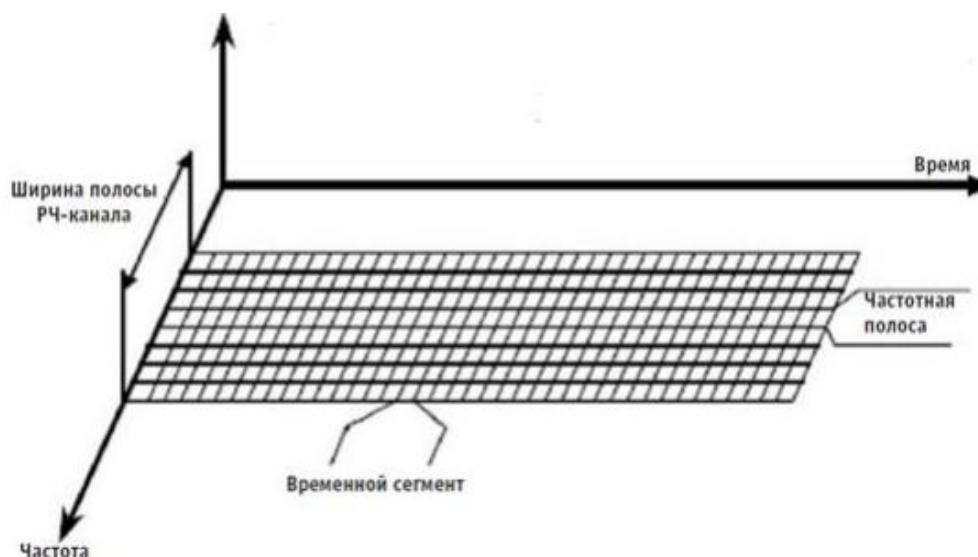


Рисунок 4.27 - Расщепление канала

Каждая частотно-временная ячейка имеет свою собственную поднесущую (см. рис. 4.28). Набор поднесущих в определенном временном сегменте называется символом OFDM. Для устранения взаимных помех между поднесущими расстояние (промежуток) между ними

выбирается равным обратной величине длительности символа: в этом случае поднесущие являются ортогональными.

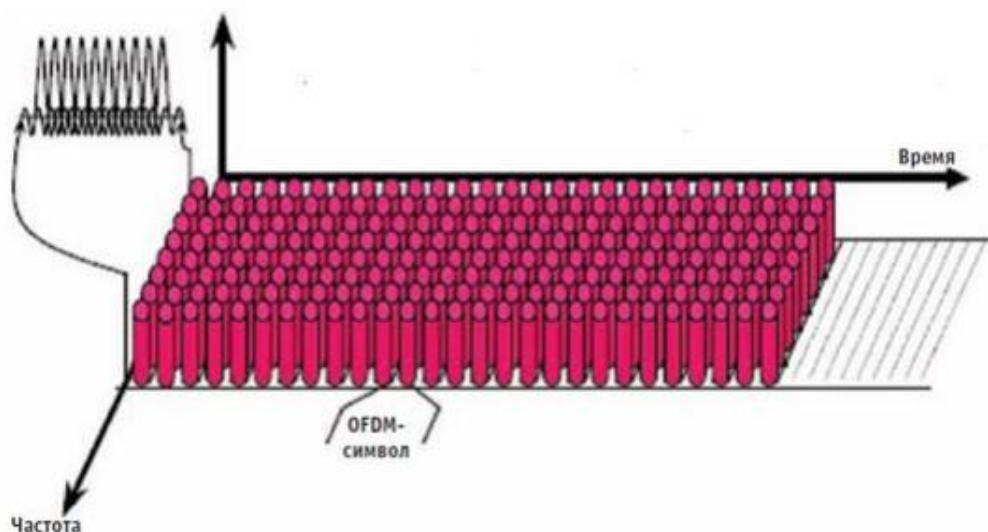


Рисунок 4.28 - Ввод поднесущих частот.

Поскольку эхо-сигналы представляют собой задержанные во времени копии основного сигнала, начало данного символа OFDM подвергается «загрязнению» задержанным окончанием предыдущего (взаимные помехи между символами). Для устранения этого эффекта между двумя соседними символами OFDM вводится защитный интервал (см. рис. 4.29). Во время защитного интервала приемные устройства игнорируют поступающий сигнал, что приводит к снижению пропускной способности канала передачи.

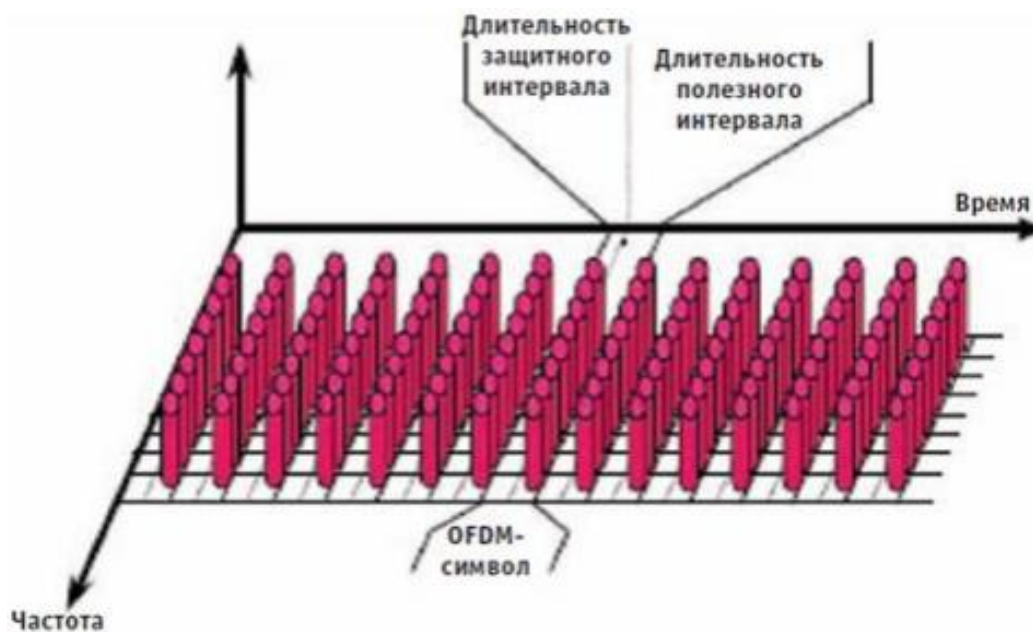


Рисунок 4.30 - Ввод защитного интервала.

Чтобы осуществить надлежащим образом демодуляцию сигнала, приемные устройства должны произвести его выборку во время полезного периода символа OFDM (но не во время защитного интервала). Но тогда необходимо ввести временное окно по отношению к моменту, когда передается в эфир каждый символ OFDM. В системе DVB-T используются «пилотные» поднесущие, равномерно распределенные в канале передачи в виде маркеров синхронизации (см. рис. 4.31)

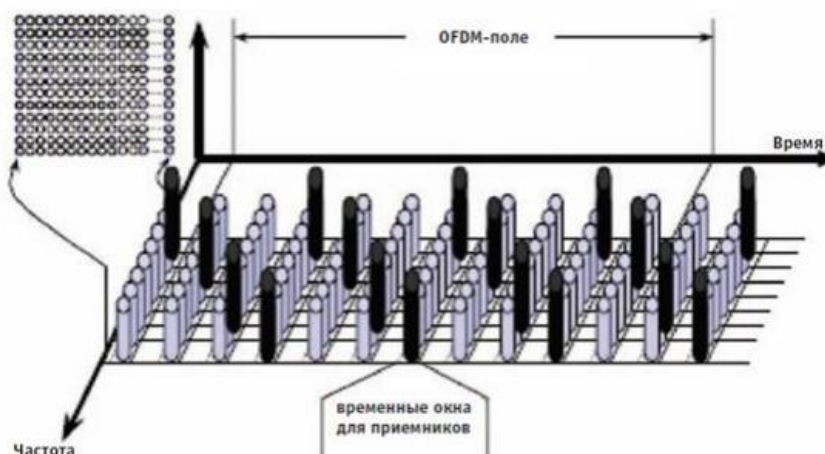


Рисунок 4.31 - Маркеры синхронизации.

Итак, общая длительность T_s OFDM-символа представляет собой сумму длительностей полезной части T_u и защитного интервала T_g . Расстояние по частоте между соседними несущими частотами OFDM-сигнала равно $1/T_u$ (поскольку чтобы осуществить надлежащим образом демодуляцию сигнала, приемные устройства должны произвести его выборку во время полезного периода символа OFDM, но не во время защитного интервала). Защитный интервал располагается перед полезной частью OFDM-символа.

Структурная схема модулятора

Принцип OFDM-модуляции заключается в следующем. В полосе канала связи передается множество несущих, каждая из которых модулируется, например, с использованием QAM-модуляции, частью общего цифрового потока. До преобразования спектра такого сигнала цифровой поток разбивается на последовательности, каждая из которых соответствует передаче kR_a битов информации, где R_a — число активных несущих, k — коэффициент используемой QAM-модуляции (или число битов информации, передаваемой на каждой активной несущей). Длительность T_0 интервала, на котором передаются все указанные kR_a битов информации, определяет минимальную частоту несущей $f_U = 1/T_U$ и интервал между несущими, т. е. частотный интервал $(R_a + R_n)f_U$, где R_n - число пассивных несущих, характеризует групповой спектр мощности радиосигнала.

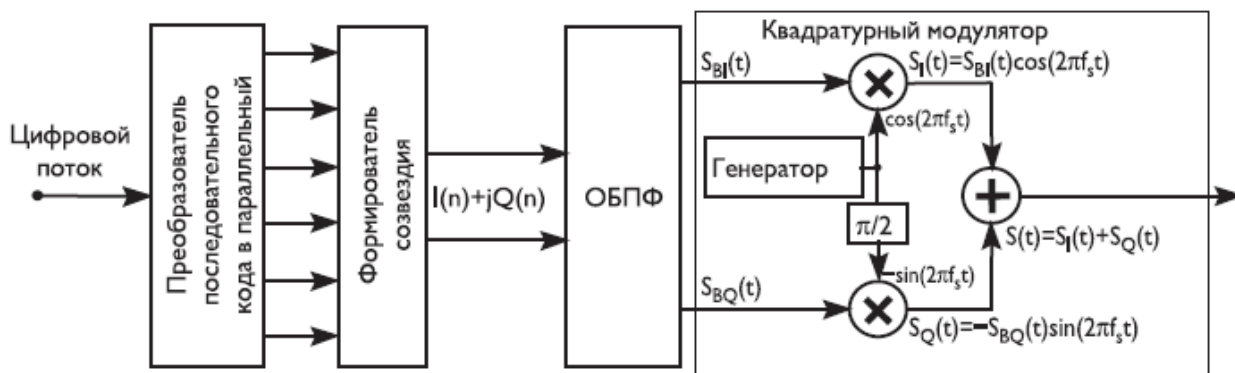


Рисунок 4.32 - Структурная схема идеального OFDM-модулятора.

Цифровой поток поступает на вход преобразователя последовательной информации в параллельную. На выходе этого преобразователя формируется код, состоящий из k битов и соответствующий используемой QAM-модуляции несущих ($k = 2$ при QPSK, $k = 4$ при QAM-16, $k = 6$ при QAM-64 и т.д.). Последовательно каждые k битов преобразуются в параллельный код, который подается на формирователь созвездия, преобразующий этот код в значения соответствующих векторов звездной диаграммы, как показано, например, на рис. 4.33.

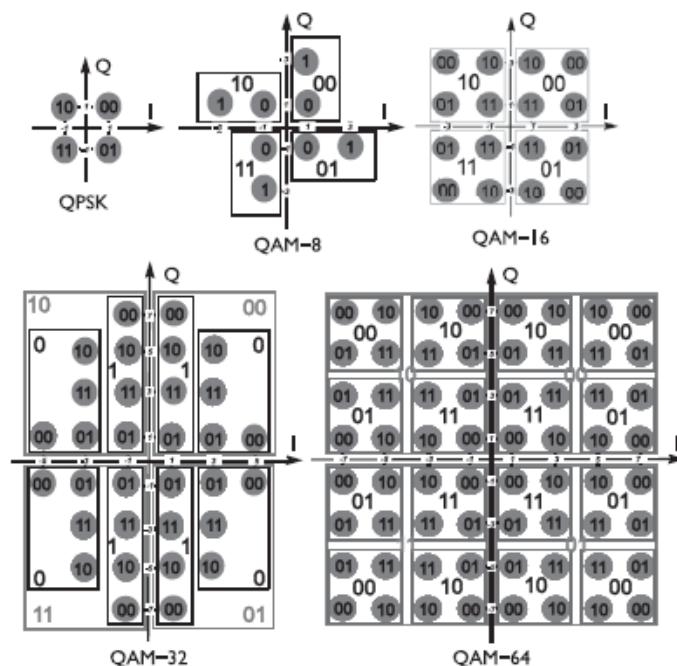


Рисунок 4.33 - Варианты сигнальных созвездий: QPSK, QAM-8, QAM-16, QAM-32 и QAM-64.

Поток битов трансформируется в формирователе созвездия в поток I и Q сигналов: $S(n) = I(n) + jQ(n)$, $1 \leq n \leq 2N + 1$, где $2N + 1 = R_a + R_n$. Блок обратного быстрого преобразования Фурье (ОБПФ) преобразует последовательности во временную комплексную функцию: $S_B(t) = \sum_{n=1}^{2N+1} S(n) e^{-j2\pi f_s t n}$

$$[I(n) + jQ(n)] \exp(2\pi jnfUt) = SBI(t) + jSBQ(t),$$

где $SBI(t) = \sum_{n=1}^{2N+1} I(n) \cos(2\pi n f_U t)$

$[I(n) \cdot \cos(2\pi n f_U t) - Q(n) \cdot \sin(2\pi n f_U t)]$ — синфазная составляющая сигнала, $SBQ(t) = \sum_{n=1}^{2N+1} [I(n) \cdot \sin(2\pi n f_U t) + Q(n) \cdot \cos(2\pi n f_U t)]$ — квадратурная составляющая сигнала (преобразование синфазной составляющей по Гильберту).

Связь между частотой f_U и количеством несущих частот $2N + 1$ сигнала в используемой полосе частот можно представить графиком, приведенным на рис.1.9, на котором изображены амплитуды отсчетов составляющих частот $n f_0$ ($n = 1, 2, \dots, 2N + 1$).

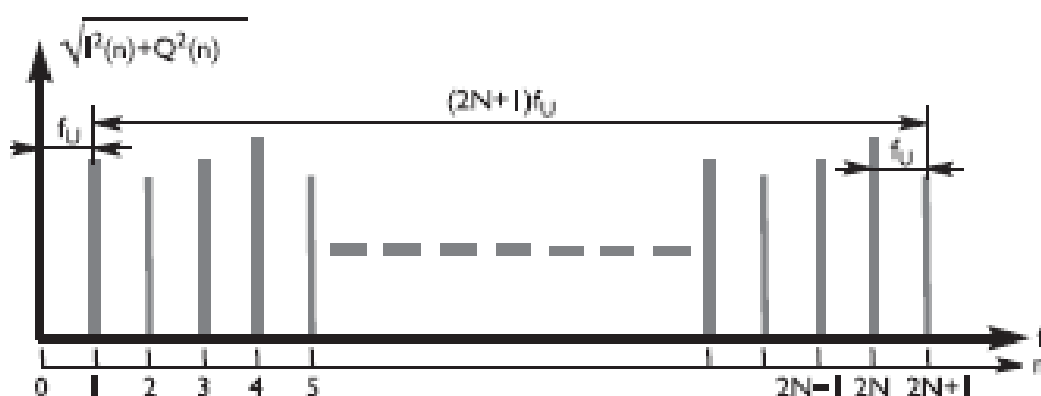


Рисунок 4.34 - График спектральных составляющих ОБПФ.

В соответствии со схемой квадратурного модулятора эти два сигнала $SBI(t)$ и $SBQ(t)$ перемножаются соответственно на сдвинутые по фазе на 90° синусоидальные сигналы. На выходах перемножителей формируются две составляющие:

$$S_I(t) = S_{BI}(t) \cdot \cos(2\pi f_s t) = \sum_{n=1}^{2N+1} \left\{ \frac{1}{2} I(n) \cdot \left[\cos 2\pi(n f_U + f_s) + \cos 2\pi(n f_U - f_s) \right] - \frac{1}{2} Q(n) \cdot \left[\sin 2\pi(n f_U + f_s) + \sin 2\pi(n f_U - f_s) \right] \right\},$$

$$S_Q(t) = -S_{BQ}(t) \cdot \sin(2\pi f_s t) = \sum_{n=1}^{2N+1} \left\{ \frac{1}{2} I(n) \cdot \left[\cos 2\pi(n f_U + f_s) - \cos 2\pi(n f_U - f_s) \right] - \frac{1}{2} Q(n) \cdot \left[\sin 2\pi(n f_U + f_s) - \sin 2\pi(n f_U - f_s) \right] \right\}.$$

В результате суммирования этих двух составляющих окончательно формируется выходной сигнал OFDM-модулятора. При этом разностные частоты ($n f_0 - f_s$) взаимно исключаются.

Структурная схема демодулятора

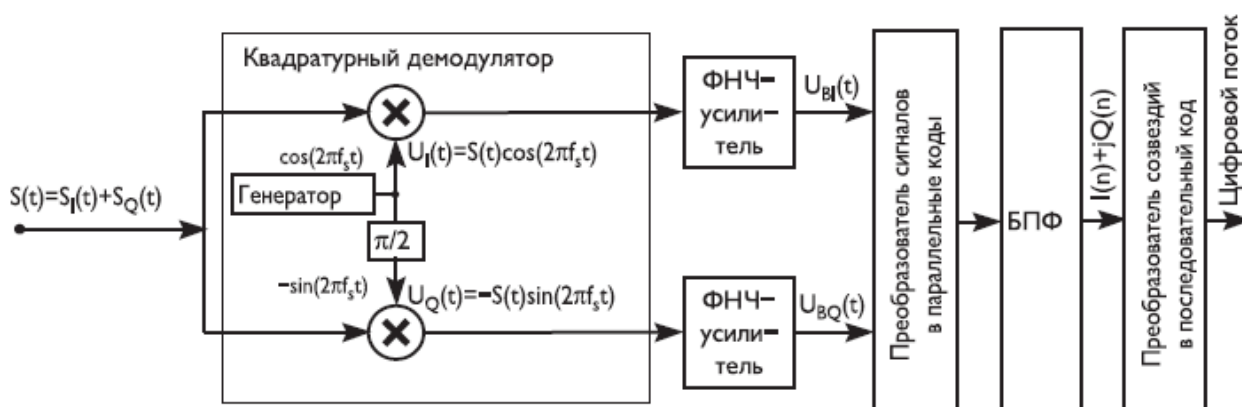


Рисунок 4.35 - Структурная схема идеального демодулятора.

Если частота генератора f_s выбрана равной $f_s = f_c + (N + 1) \cdot f_U$, где f_c - центральная частота радиоканала, то выходной сигнал определяется следующим соотношением:

$$S(t) = \sum_{n=1}^{2N+1} \{I(n) \cdot \cos[2\pi(f_s + nf_U)t] - Q(n) \cdot \sin[2\pi(f_s + nf_U)t]\} =$$

$$= \sum_{n=-N}^N \{I(n) \cdot \cos[2\pi(f_c + nf_U)t] - Q(n) \cdot \sin[2\pi(f_c + nf_U)t]\}.$$

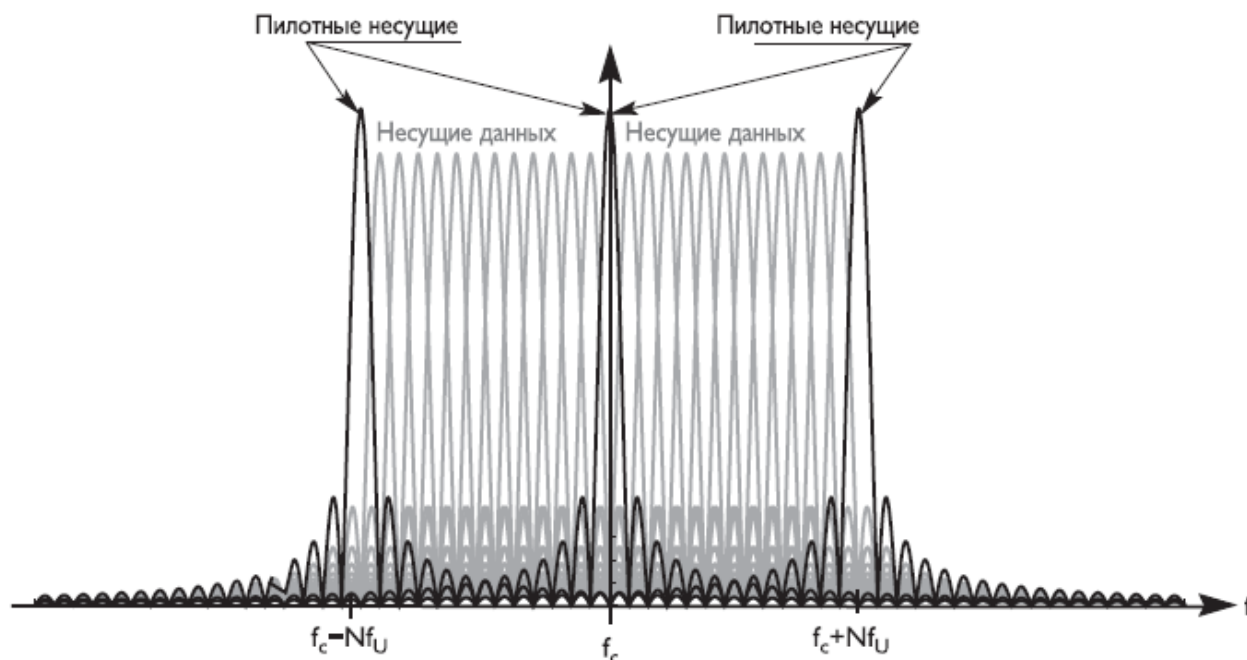


Рисунок 4.36 - Групповой спектр несущих частот.

На рис. 4.36 изображена структура группового спектра несущих частот OFDM-сигнала. Здесь условно показаны составляющие синусоидальные сигналы, промодулированные соответствующими данными созвездий дискретной QAM-модуляции. В

составе сигнала имеются также специальные пилотные несущие, на которых передается информация о параметрах системы. Эти пилотные несущие используются также для обеспечения устойчивой синхронизации и коррекции характеристик в демодуляторе.

Таким образом, в системах с OFDM модуляцией передаваемая цифровая информация разделена на большое число низкоскоростных подканалов, длительность тактового интервала передачи каждой несущей весьма велика. Такое построение системы преобразует широкополосный канал с одной несущей в большое число независимых узкополосных каналов с частотным разделением, что упрощает коррекцию параметров затухающего сигнала. Более того, ряд групп несущих может быть полностью подавлен при приеме, если дополнительно вводится корректирующее кодирование данных в сочетании с временным и частотным перемежением. При таком построении системы модуляцию часто называют COFDM (Coded Orthogonal Frequency Division Multiplexing — кодированное ортогональное частотное мультиплексирование, разновидность технологии OFDM, сочетающая канальное кодирование и OFDM).

На рис. 4.36 приведена структурная схема идеального демодулятора, содержащего квадратурный демодулятор и ряд преобразователей, формирующих передаваемую цифровую информацию.

В квадратурном демодуляторе входной сигнал $S(t)$ перемножается с двум сдвинутыми на 90° синусоидальными сигналами опорной частоты f_s . На двух выходах выделяются сигналы:

$$U_I(t) = S(t) \cdot \cos(2\pi f_s t) = \sum_{n=1}^{2N+1} \left\{ \frac{1}{2} I(n) \cdot [\cos 2\pi n f_U t + \cos 2\pi(2f_s + n f_U)t] - \frac{1}{2} Q(n) \cdot [\sin 2\pi n f_U t + \sin 2\pi(2f_s + n f_U)t] \right\},$$

$$U_Q(t) = S(t) \cdot [-\sin(2\pi f_s t)] = \sum_{n=1}^{2N+1} \left\{ \left(\frac{1}{2} I(n) \cdot [\sin 2\pi n f_U t - \sin 2\pi(2f_s + n f_U)t] + \frac{1}{2} Q(n) \cdot [\cos 2\pi n f_U t - \cos 2\pi(2f_s + n f_U)t] \right) \right\}.$$

Эти два сигнала поступают соответственно на два фильтра нижних частот и усиливаются в два раза. На входы преобразователя в параллельные коды при этом подаются два сигнала:

$$U_{BI}(t) = \sum_{n=1}^{2N+1} [I(n) \cdot \cos(2\pi n f_U t) - Q(n) \cdot \sin(2\pi n f_U t)],$$

$$U_{BQ}(t) = \sum_{n=1}^{2N+1} [I(n) \cdot \sin(2\pi n f_U t) + Q(n) \cdot \cos(2\pi n f_U t)].$$

Таким образом, на вход схемы быстрого преобразования Фурье подается сигнал, который может быть представлен в следующем виде

$$U_B(t) = U_{BI}(t) + jU_{BQ}(t) = \sum_{n=1}^{2N+1} [I(n) + jQ(n)] \exp(2\pi j n f_U t).$$

Очевидно: после преобразования Фурье формируется последовательность, определяющая векторы звездной диаграммы:

$$S(n) = I(n) + jQ(n), \quad 1 \leq n \leq 2N + 1,$$

которые выходным преобразователем преобразуются в передаваемую цифровую последовательность.

Достоинства и недостатки OFDM

Плюсы

- Высокая эффективность использования радиочастотного спектра, объясняемая почти прямоугольной формой огибающей спектра при большом количестве поднесущих.
- Простая аппаратная реализация: базовые операции реализуются методами цифровой обработки.
- Хорошее противостояние межсимвольным помехам (ISI – intersymbol interference) и интерференции между поднесущими (ICI – intercarrier interference).
- Возможность применения различных схем модуляции для каждой поднесущей, что позволяет адаптивно варьировать помехоустойчивость и скорость передачи информации.

Минусы

- Необходима высокая синхронизация частоты и времени.
- Чувствительность к эффекту Доплера, ограничивающая применение OFDM в мобильных системах.
- Не идеальность современных приёмников и передатчиков вызывает фазовый шум, что ограничивает производительность системы.
- Защитный интервал, используемый в OFDM для борьбы с многолучевым распространением, снижает спектральную эффективность сигнала.

Технология ММО

ММО (Multiple Input Multiple Output – множественный вход множественный выход) – это технология, используемая в беспроводных системах связи (WIFI, WI-MAX, сотовые сети связи), позволяющая значительно улучшить спектральную эффективность системы, максимальную скорость передачи данных и емкость сети. Главным способом достижения указанных выше преимуществ является передача данных от источника к получателю через несколько радио соединений, откуда данная технология и получила свое название. Рассмотрим предысторию данного вопроса, и определим основные причины, послужившие широкому распространению технологии ММО.

Необходимость в высокоскоростных соединениях, предоставляющих высокие показатели качества обслуживания (QoS) с высокой отказоустойчивостью растет от года в год. Этому в значительной мере способствует появление таких сервисов как VoIP (Voice over Internet Protocol), видеоконференции, VoD (Video on Demand) и др. Однако большинство беспроводных технологий не позволяют предоставить абонентам высокое качество обслуживания на краю зоны покрытия. В сотовых и других беспроводных системах связи качество соединения, также как и доступная скорость передачи данных стремительно падает с удалением от базовой станции (BTS). Вместе с этим падает и качество услуг, что в итоге приводит к невозможности предоставления услуг реального времени с высоким качеством на всей территории радио покрытия сети. Для решения данной проблемы можно попробовать максимально плотно установить базовые станции и организовать внутреннее покрытие во всех местах с низким уровнем сигнала. Однако это потребует значительных финансовых затрат что в конечном счете приведет к росту стоимости услуги и снижению конкурентоспособности. Таким образом, для решения данной проблемы требуется оригинальное нововведение, использующее, по возможности, текущий частотный диапазон и не требующее строительства новых объектов сети.

Особенности распространения радиоволн

Для того чтобы понять принципы действия технологии ММО необходимо рассмотреть общие принципы распространения радио волн в пространстве. Волны, излучаемые различными системами беспроводной радиосвязи в диапазоне свыше 100 МГц, во многом ведут себя как световые лучи. Когда радиоволны при распространении встречаются какую-либо поверхность, то в зависимости от материала и размера препятствия часть энергии поглощается, часть проходит насквозь, а оставшаяся – отражается. На соотношение долей поглощенной, отраженной и прошедшей насквозь частей энергий влияет множество внешних факторов, в том числе и частота сигнала. Причем отраженная и прошедшая

насквозь энергии сигнала могут изменить направление своего дальнейшего распространения, а сам сигнал разбивается на несколько волн.

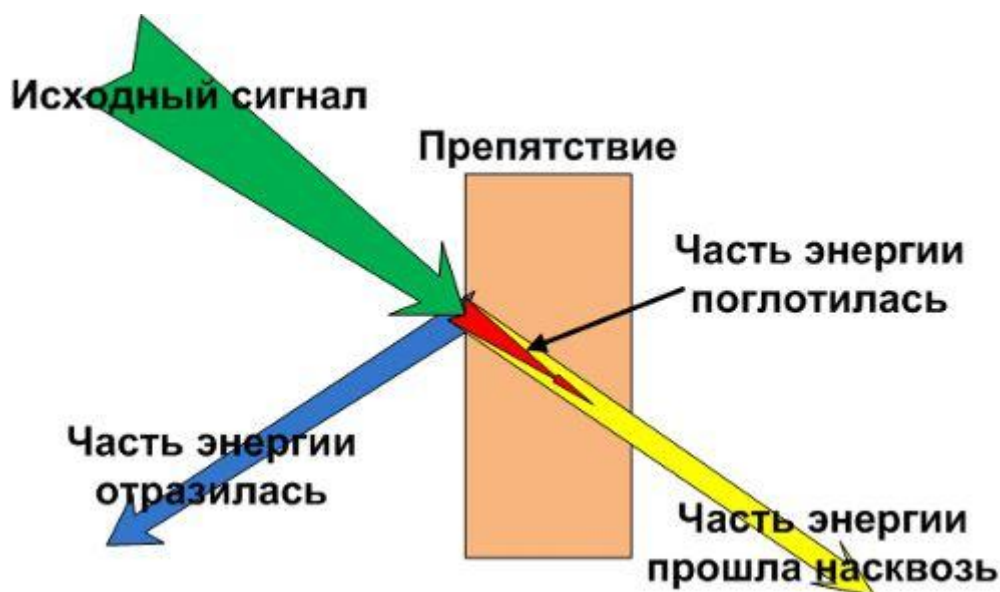


Рисунок 4.37 - Распределение энергии сигнала при взаимодействии с препятствием.

Распространяющийся по вышеуказанным законам сигнал от источника к получателю после встречи с многочисленными препятствиями разбивается на множество волн, лишь часть из которых достигнет приемник. Каждая из дошедших до приемника волн образует так называемый путь распространения сигнала. Причем из-за того, что разные волны отражаются от разного числа препятствий и проходят разное расстояние, различные пути имеют разные временные задержки.

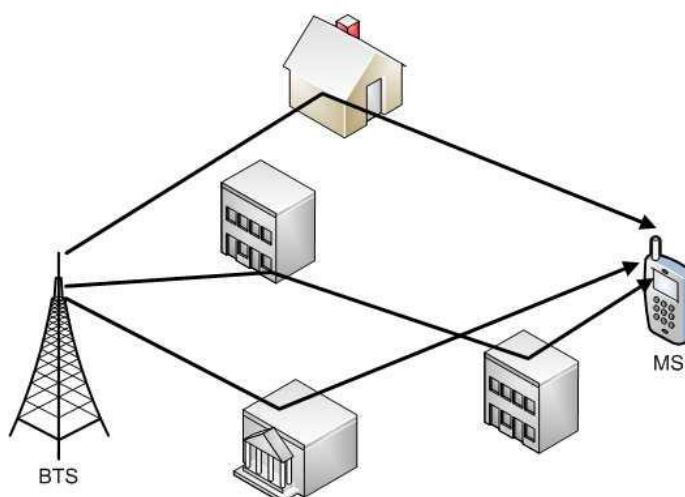


Рисунок 4.38 - Пример многолучевого распространения сигнала.

В условиях плотной городской постройки, из-за большого числа препятствий, таких как здания, деревья, автомобили и др., очень часто возникает ситуация, когда между абонентским оборудованием (MS) и антеннами базовой станции (BTS) отсутствует прямая видимость. В этом случае, единственным вариантом достижения сигнала приемника являются отраженные волны. Однако, как отмечалось выше, многократно отраженный

сигнал уже не обладает исходной энергией и может прийти с запозданием. Особую сложность также создает тот факт, что объекты не всегда остаются неподвижными и обстановка может значительно измениться с течением времени. В связи с этим возникает проблема многолучевого распространения сигнала – одна из наиболее существенных проблем в беспроводных системах связи.

Для борьбы с многолучевым распространением сигналов применяется несколько различных решений. Одной из наиболее распространенных технологий является Receive Diversity – разнесенный прием. Суть его заключается в том, что для приема сигнала используется не одна, а сразу несколько антенн (обычно две, реже четыре), расположенные на расстоянии друг от друга. Таким образом, получатель имеет не одну, а сразу две копии переданного сигнала, пришедшего различными путями. Это дает возможность собрать больше энергии исходного сигнала, т.к. волны, принятые одной антенной, могут не быть принятыми другой и наоборот. Также сигналы, приходящие в противофазе к одной антенне, могут приходить к другой синфазно. Эту схему организации радио интерфейса можно назвать Single Input Multiple Output (SIMO), в противовес стандартной схеме Single Input Single Output (SISO). Также может быть применен обратный подход: когда используется несколько антенн на передачу и одна на прием. Благодаря этому также увеличивается общая энергия исходного сигнала, полученная приемником. Эта схема называется Multiple Input Single Output (MISO). В обеих схемах (SIMO и MISO) несколько антенн устанавливаются на стороне базовой станции, т.к. реализовать разнесение антенн в мобильном устройстве на достаточно большое расстояние сложно без увеличения габаритов самого оконечного оборудования.

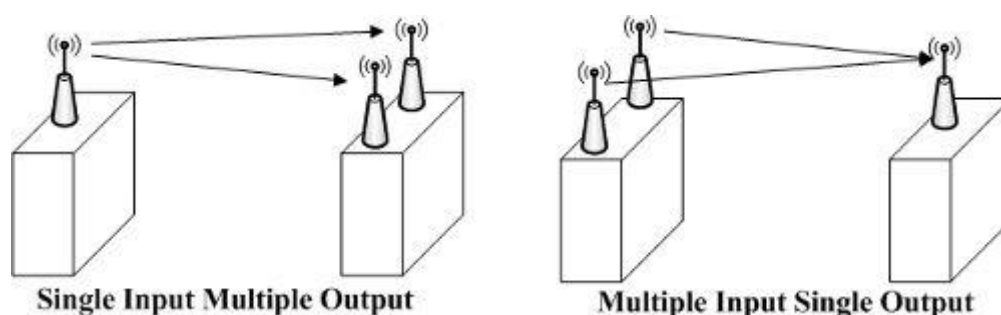


Рисунок 4.39 – Применение нескольких антенн

В результате дальнейших рассуждений мы приходим к схеме Multiple Input Multiple Output (MIMO). В этом случае устанавливаются несколько антенн на передачу и прием. Однако в отличие от указанных выше схем эта схема разнесения позволяет не только бороться с многолучевым распространением сигнала, но и получить некоторые дополнительные преимущества. За счет использования нескольких антенн на передаче и приеме каждой паре, передающей/приемной антенне можно сопоставить отдельный тракт

для передачи информации. При этом разнесенный прием будет выполняться оставшимися антеннами, а данная антенна также будет выполнять функции дополнительной антенны для других трактов передачи. В результате, теоретически, можно увеличить скорость передачи данных во столько раз, сколько дополнительных антенн будет использоваться. Однако существенное ограничение накладывается качеством каждого радио тракта.

Принцип работы и структурная схема ММО

Как уже отмечалось выше, для организации технологии ММО необходима установка нескольких антенн на передающей и на приемной стороне. Обычно устанавливается равное число антенн на входе и выходе системы, т.к. в этом случае достигается максимальная скорость передачи данных. Чтобы показать число антенн на приеме и передаче вместе с названием технологии «ММО» обычно упоминается обозначение « $A \times B$ », где A – число антенн на входе системы, а B – на выходе. Под системой в данном случае понимается радио соединение.

Для работы технологии ММО необходимы некоторые изменения в структуре передатчика по сравнению с обычными системами. Рассмотрим лишь один из возможных, наиболее простых, способов организации технологии ММО. В первую очередь, на передающей стороне необходим делитель потоков, который будет разделять данные, предназначенные для передачи на несколько низкоскоростных подпотоков, число которых зависит от числа антенн. Например, для ММО 4×4 и скорости поступления входных данных 200 Мбит/сек делитель будет создавать 4 потока по 50 Мбит/сек каждый. Далее каждый из данных потоков должен быть передан через свою антенну. Обычно, антенны на передаче устанавливаются с некоторым пространственным разнесением, чтобы обеспечить как можно большее число побочных сигналов, которые возникают в результате переотражений. В одном из возможных способов организации технологии ММО сигнал передается от каждой антенны с различной поляризацией, что позволяет идентифицировать его при приеме. Однако в простейшем случае каждый из передаваемых сигналов оказывается промаркированным самой средой передачи (задержкой во времени, затуханием и другими искажениями).

На приемной стороне несколько антенн принимают сигнал из радиоэфира. Причем антенны на приемной стороне также устанавливаются с некоторым пространственным разнесением, за счет чего обеспечивается разнесенный прием, обсуждавшийся ранее. Принятые сигналы поступают на приемники, число которых соответствует числу антенн и трактов передачи. Причем на каждый из приемников поступают сигналы от всех антенн системы. Каждый из таких сумматоров выделяет из общего потока энергию сигнала только того тракта, за который он отвечает. Делает он это либо по какому-либо заранее

предусмотренному признаку, которым был снабжен каждый из сигналов, либо благодаря анализу задержки, затухания, сдвига фазы, т.е. набору искажений или «отпечатку» среды распространения. В зависимости от принципа работы системы (Bell Laboratories Layered Space-Time - BLAST, Selective Per Antenna Rate Control (SPARC) и т.д.), передаваемый сигнал может повторяться через определенное время, либо передаваться с небольшой задержкой через другие антенны.

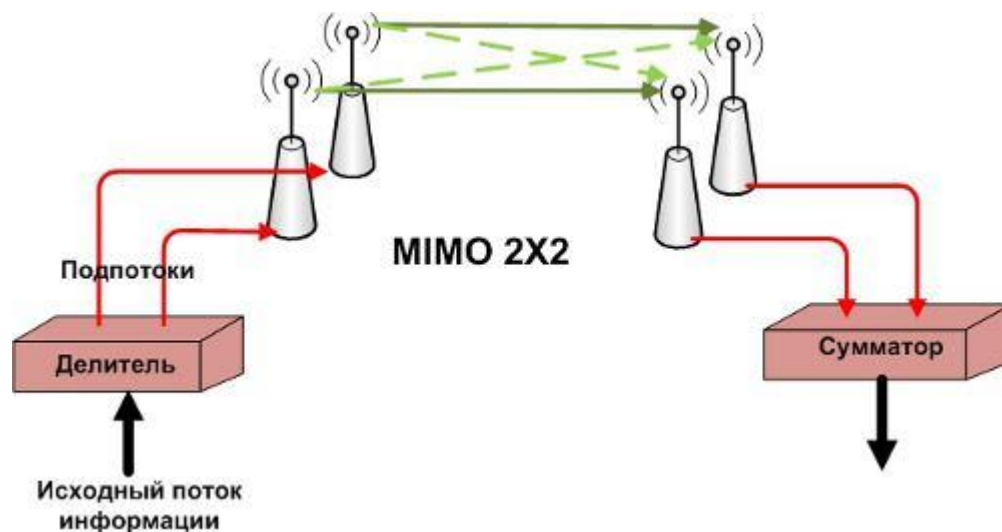


Рисунок 4.40 – Принцип организации технологии MIMO

В системе с технологией MIMO может возникнуть необычное явление, которое заключается в том, что скорость передачи данных в системе MIMO может снизиться в случае появления прямой видимости между источником и приемником сигнала. Это обусловлено в первую очередь уменьшением выраженности искажений окружающего пространства, который маркирует каждый из сигналов. В результате на приемной стороне становится проблематичным разделить сигналы, и они начинают оказывать влияние друг на друга. Таким образом, чем выше качество радио соединения, тем меньше преимуществ можно получить от MIMO.

Применение MIMO

Технология MIMO в последнее десятилетие является одним из самых актуальных способов увеличения пропускной способности и емкости беспроводных систем связи. Рассмотрим некоторые примеры использования MIMO в различных системах связи.

Стандарт WiFi 802.11n – один из наиболее ярких примеров использования технологии MIMO. Согласно ему, он позволяет поддерживать скорость до 300 Мбит/сек. Причем предыдущий стандарт 802.11g позволял предоставлять лишь 50 Мбит/сек. Кроме увеличения скорости передачи данных, новый стандарт благодаря MIMO также позволяет обеспечить лучшие характеристики качества обслуживания в местах с низким уровнем сигнала. 802.11n используется не только в системах точка/многоточка (Point/Multipoint) – наиболее

привычной нише использования технологии WiFi для организации LAN (Local Area Network), но и для организации соединений типа точка/точка которые используются для организации магистральных каналов связи со скоростью несколько сотен Мбит/сек и позволяющих передавать данные на десятки километров (до 50 км).

Стандарт WiMAX также имеет два релиза, которые раскрывают новые возможности перед пользователями с помощью технологии MIMO. Первый – 802.16e – предоставляет услуги мобильного широкополосного доступа. Он позволяет передавать информацию со скоростью до 40 Мбит/сек в направлении от базовой станции к абонентскому оборудованию. Однако MIMO в 802.16e рассматривается как опция и используется в простейшей конфигурации – 2x2. В следующем релизе 802.16m MIMO рассматривается как обязательная технология, с возможной конфигурацией 4x4. В данном случае WiMAX уже можно отнести к сотовым системам связи, а именно четвертому их поколению (за счет высокой скорости передачи данных), т.к. обладает рядом присущих сотовым сетям признаков: роуминг, хэндовер, голосовые соединения. В случае мобильного использования, теоретически, может быть достигнута скорость 100 Мбит/сек. В фиксированном исполнении скорость может достигать 1 Гбит/сек.

Наибольший интерес представляет использование технологии MIMO в системах сотовой связи. Данная технология находит свое применение, начиная с третьего поколения систем сотовой связи. Например, в стандарте UMTS, в Rel. 6 она используется совместно с технологией HSPA с поддержкой скоростей до 20 Мбит/сек, а в Rel. 7 – с HSPA+, где скорости передачи данных достигают 40 Мбит/сек. Однако в системах 3G MIMO так и не нашла широкого применения.

Системы 4G, а именно LTE, также предусматривают использование MIMO в конфигурации до 8x8. Это в теории может дать возможность передавать данные от базовой станции к абоненту свыше 300 Мбит/сек. Также важным положительным моментом является устойчивое качество соединения даже на краю соты. При этом даже на значительном удалении от базовой станции, или при нахождении в глухом помещении будет наблюдаться лишь незначительное снижение скорости передачи данных.

Таким образом, технология MIMO находит применение практически во всех системах беспроводной передачи данных. Причем потенциал ее не исчерпан. Уже сейчас разрабатываются новые варианты конфигурации антенн, вплоть до 64x64 MIMO. Это в будущем позволит добиться еще больших скоростей передачи данных, емкости сети и спектральной эффективности.

1. Задание на лабораторную работу

Для исследования OFDM:

1. Изменяя вероятность нуля (probability of a zero) от 0 до 1 с шагом 0,1 снять зависимость частоты ошибок от вероятности нуля. Построить график.
2. Изменяя отношение сигнал/шум от 0 дБ до 20 дБ с шагом 4 снять зависимость частоты ошибок от отношения сигнал/шум.
3. Снять созвездие «IFFT + white noise on amplitude Constellation» для значений отношения сигнал/шум 5 дБ, 10 дБ, 20 дБ.

Схема исследования OFDM представлена на рисунке 3.1:

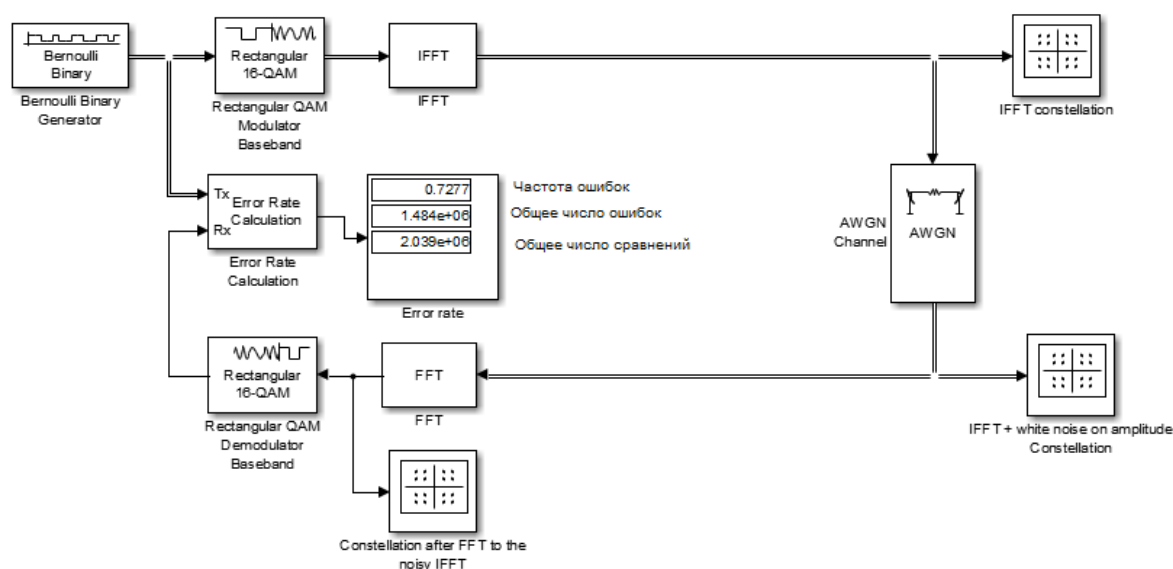


Рисунок 4.41 – Схема исследования системы OFDM.

В работе использованы следующие блоки:

- 1) Bernoulli Binary Generator – генератор ПСП.
- 2) QAM – модулятор.
- 3) Блок обратного быстрого преобразования Фурье.
- 4) Блок AWGN (канал с шумом).
- 5) Блок QAM-демодулятора.
- 6) Блок подсчёта ошибок.

Настройки блоков представлены на рисунках 3.2 – 3.7.

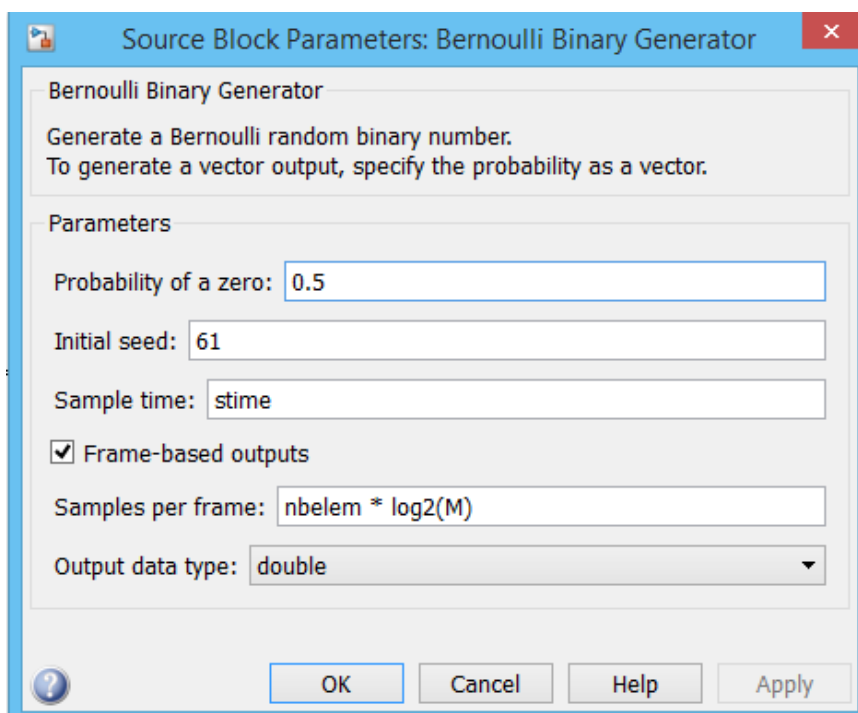


Рисунок 4.42 – Настройки блока генератора ПСП.

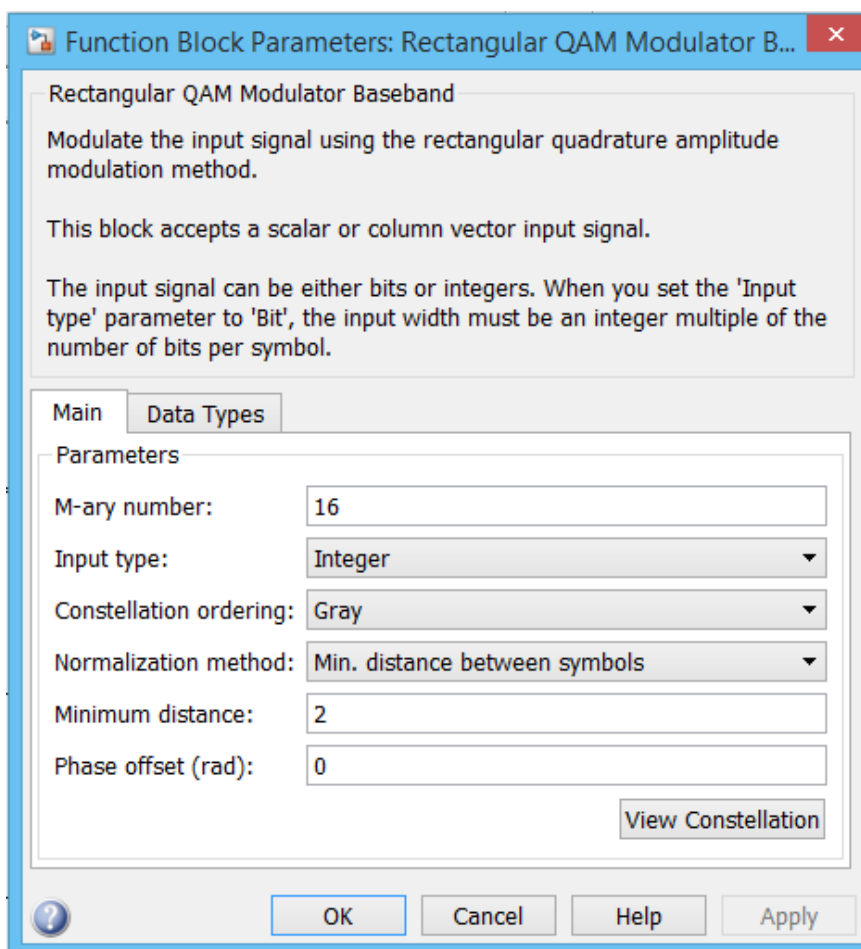


Рисунок 4.43 – Настройки QAM – модулятора.

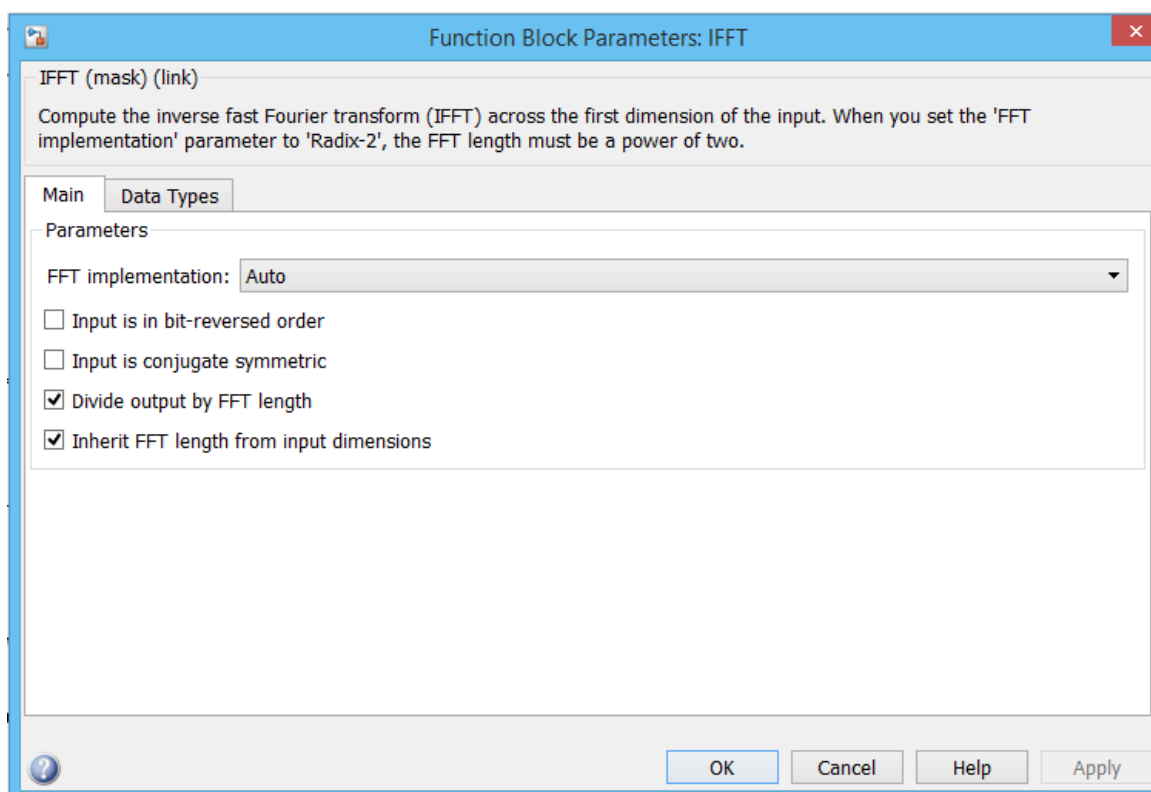


Рисунок 4.44 – Настройки блока преобразования Фурье.

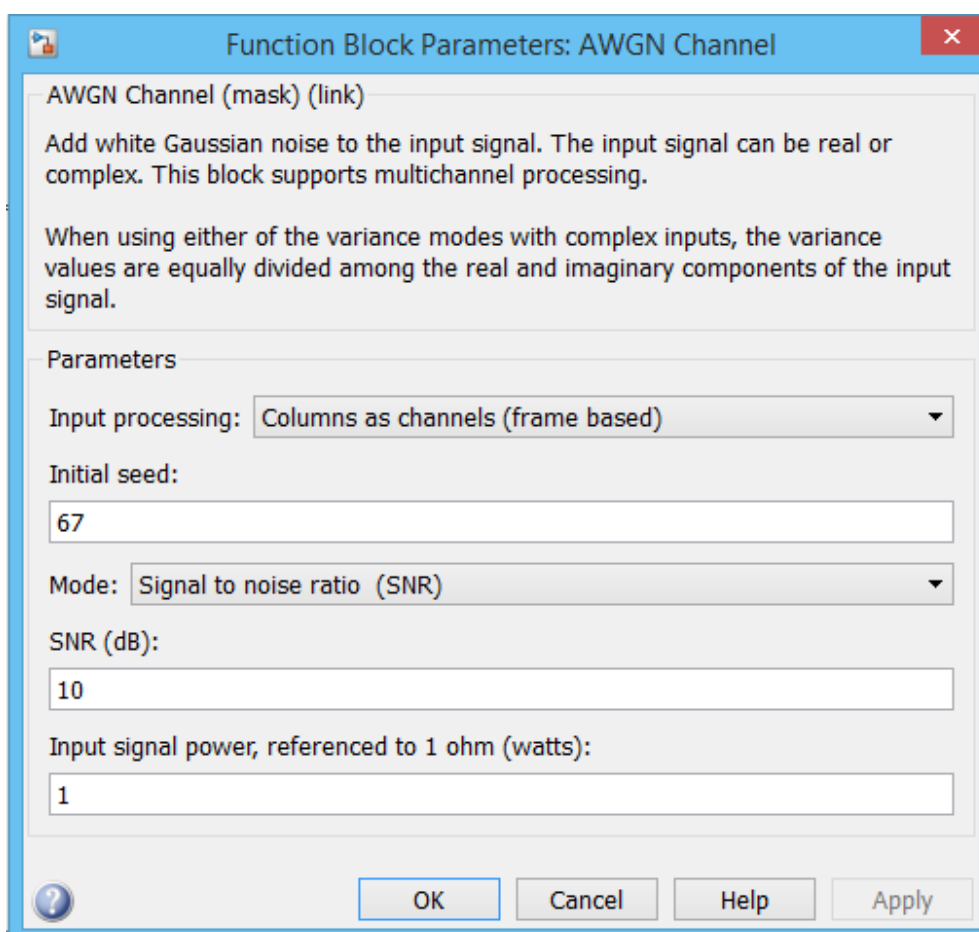


Рисунок 4.45 – Настройки блока канала с шумами.

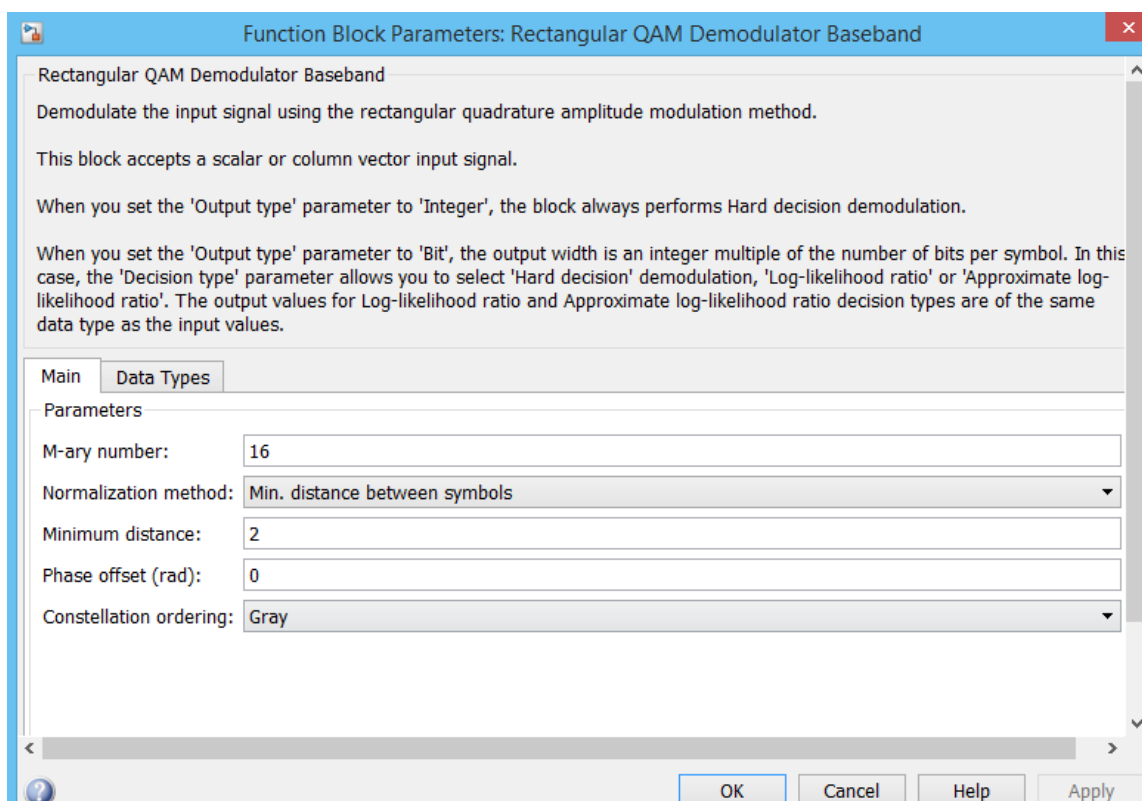


Рисунок 4.46 – Настройки QAM-демодулятора.

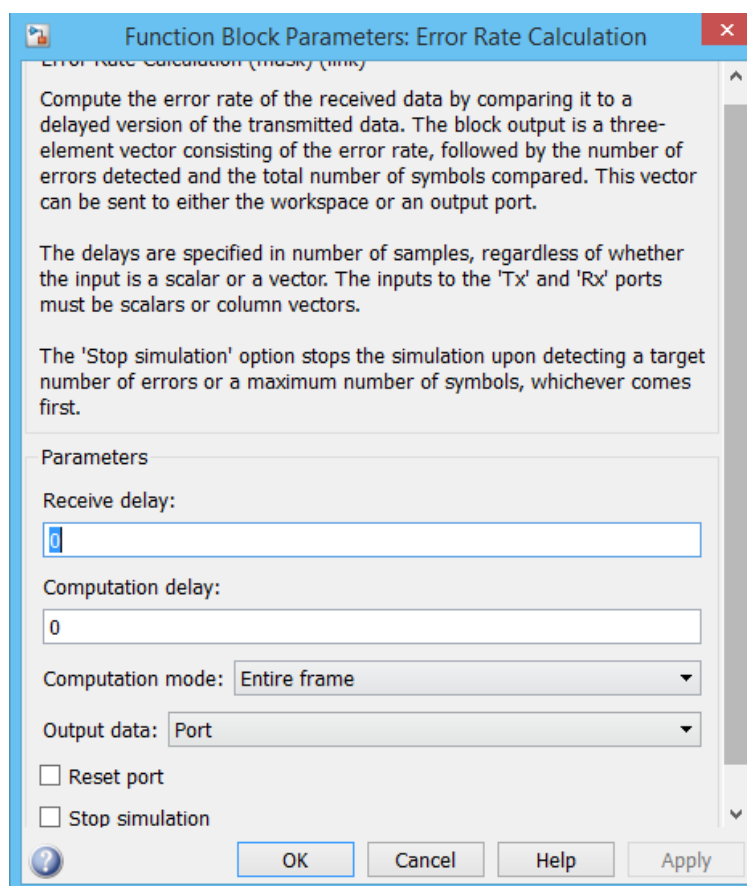


Рисунок 4.47 – Настройки блока подсчёта ошибок.

Для построения зависимости частоты ошибок от вероятности нуля нужно в генераторе ПСП изменять значение «probability of a zero» от 0 до 1.

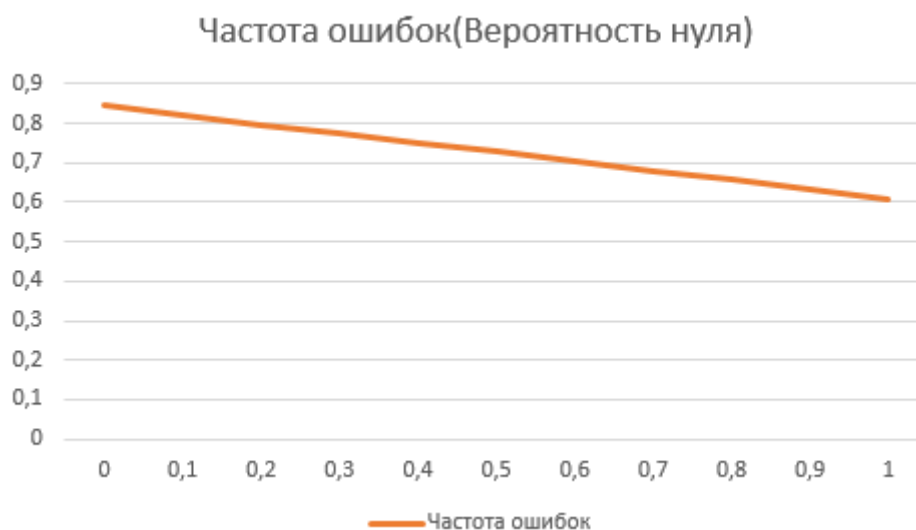


Рисунок 4.48 – Зависимость частоты ошибок от вероятности нуля.

Для построения зависимости частоты ошибок от отношения сигнал/шум нужно в блока канала с шумами изменять значение «probability of a zero» от 0 дБ до 20 дБ.

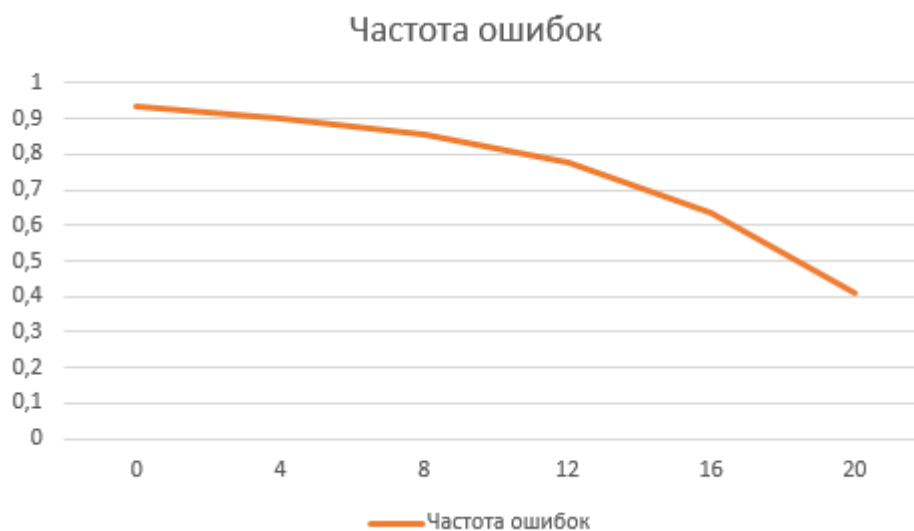


Рисунок 4.49 – Зависимость частоты ошибок от отношения сигнал/шум.

На рисунках ниже представлены созвездия для различных значений отношения сигнал/шум:

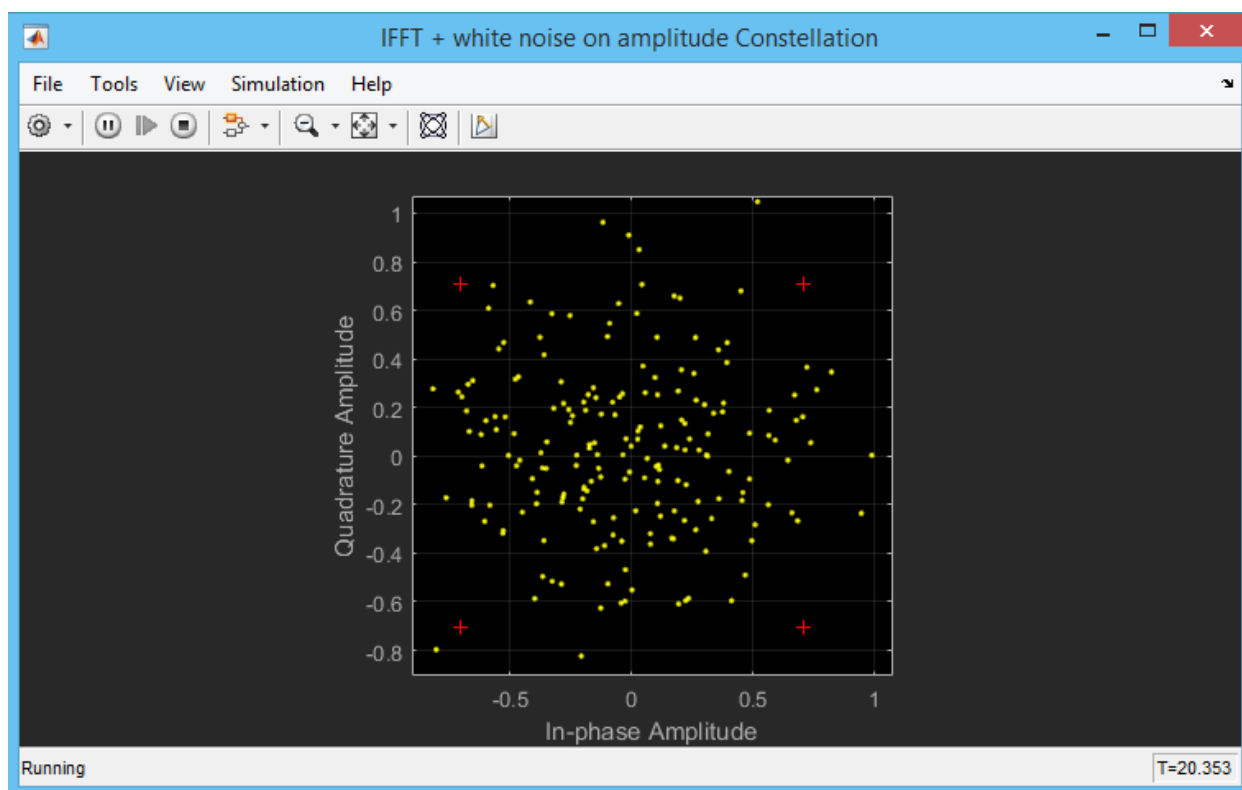


Рисунок 4.50 – Созвездие для значения отношения сигнал/шум 5 дБ.

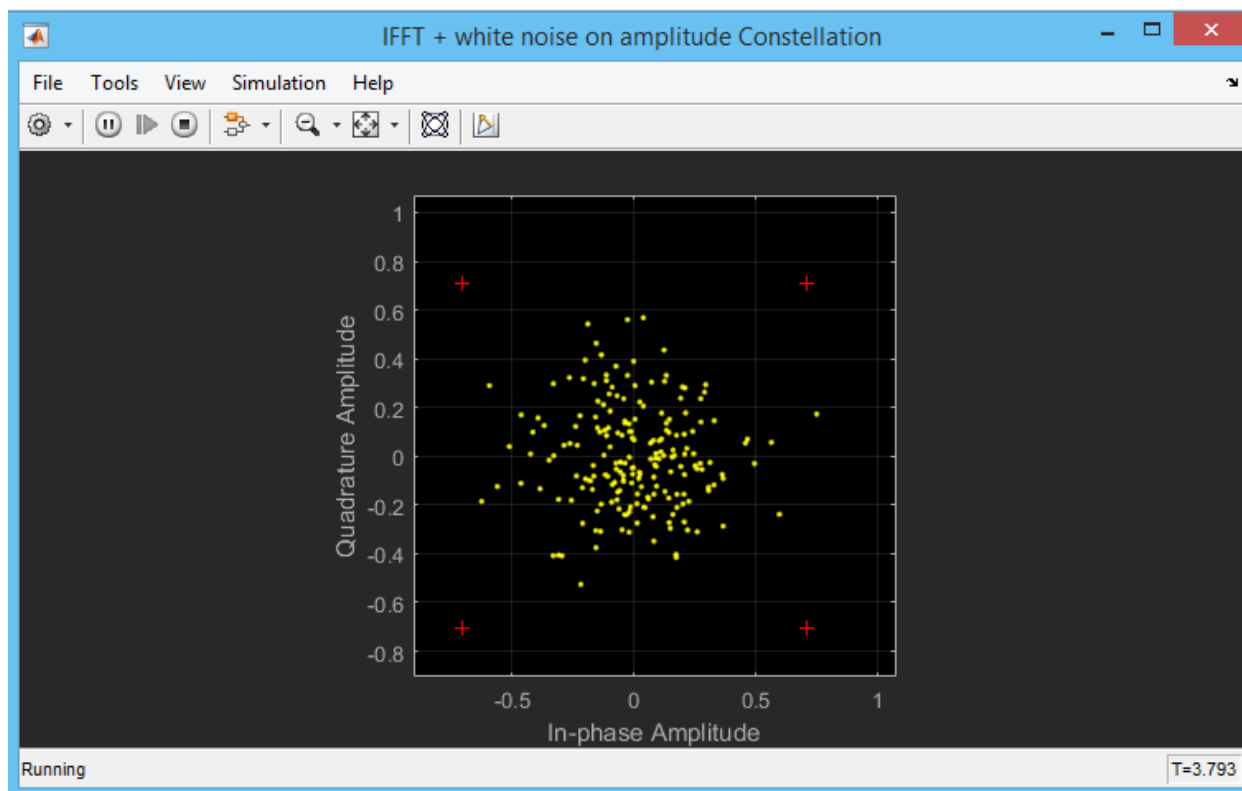


Рисунок 4.51 - Созвездие для значения отношения сигнал/шум 10 дБ.

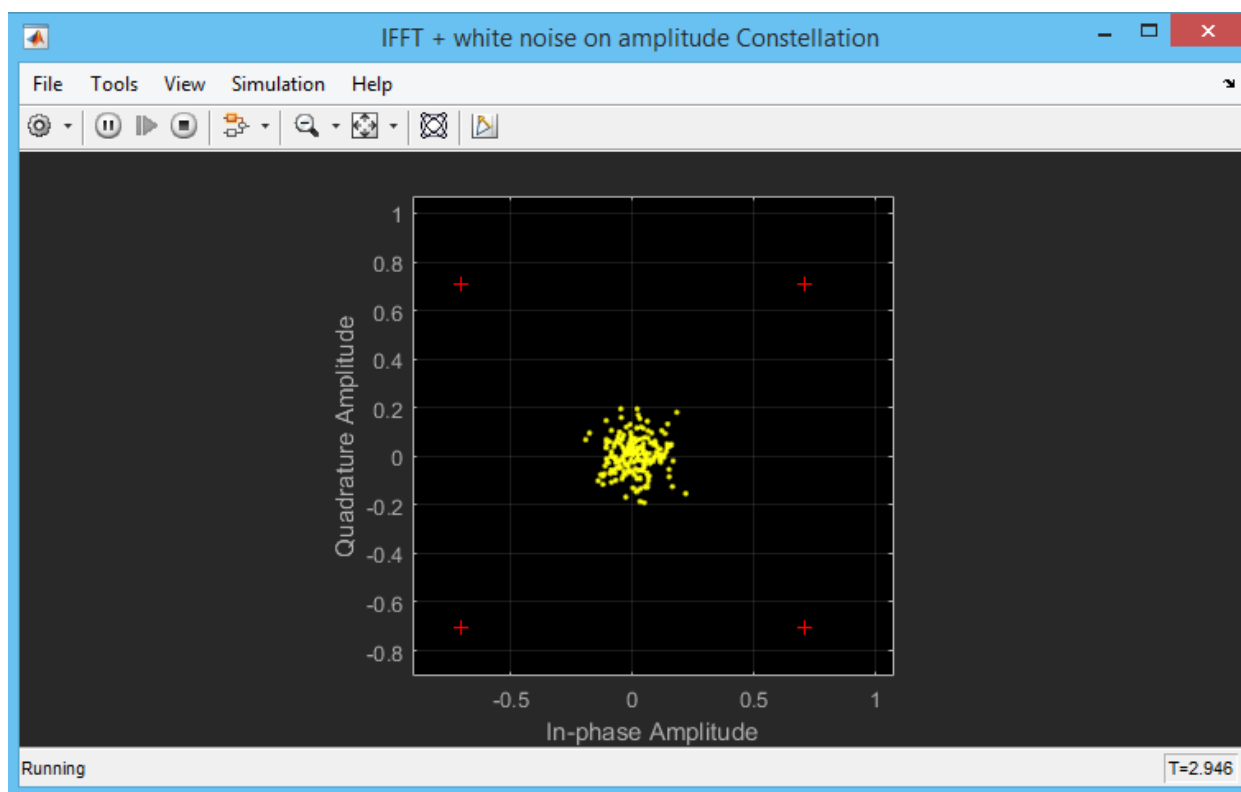


Рисунок 4.52 - Созвездие для значения отношения сигнал/шум 20 дБ.

Для исследования MIMO:

1. Изменяя SNR, dB (signal noise rate) в блоке Model Parameters от 0 до 30 с шагом 5, снять зависимость FER (частоты ошибок в кадре) от SNR. Построить график для адаптивного случая MIMO. Зафиксировать характеристики сигнала из блока Signal Visualization (двойной щелчок, рисунок 3.16).

Справка. В адаптивном случае система автоматически подстраивается под условия в канале под заданное в блоке Model Parameters значение FER. Для включения адаптивного случая в блоке Adaptive Control Panel нужно выставить переключатели в соответствующие положения.

2. Перевести переключатели в блоке Adaptive Control Panel в положения, соответствующие ручному вводу количества антенн (рисунок 3.15).

Изменяя SNR, dB (signal noise rate) в блоке Model Parameters от 0 до 30 с шагом 5, снять зависимость FER (частоты ошибок в кадре) от SNR. Построить график. Зафиксировать характеристики сигнала из блока Signal Visualization.

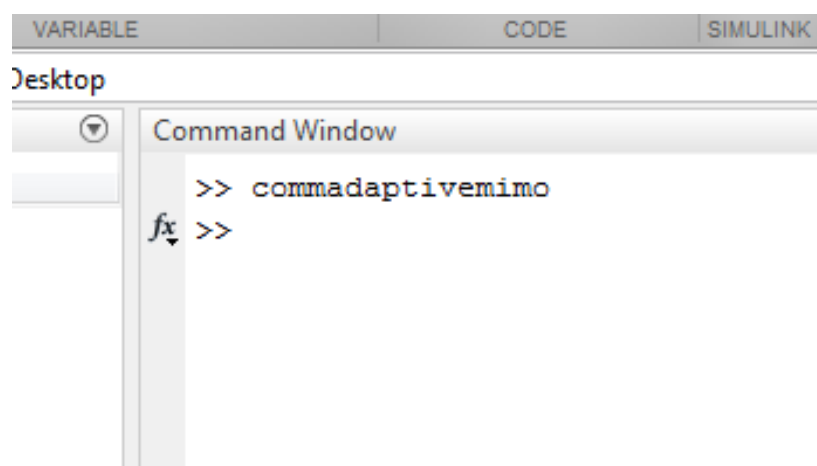


Рисунок 4.53 – Команда для создания схемы исследования MIMO.

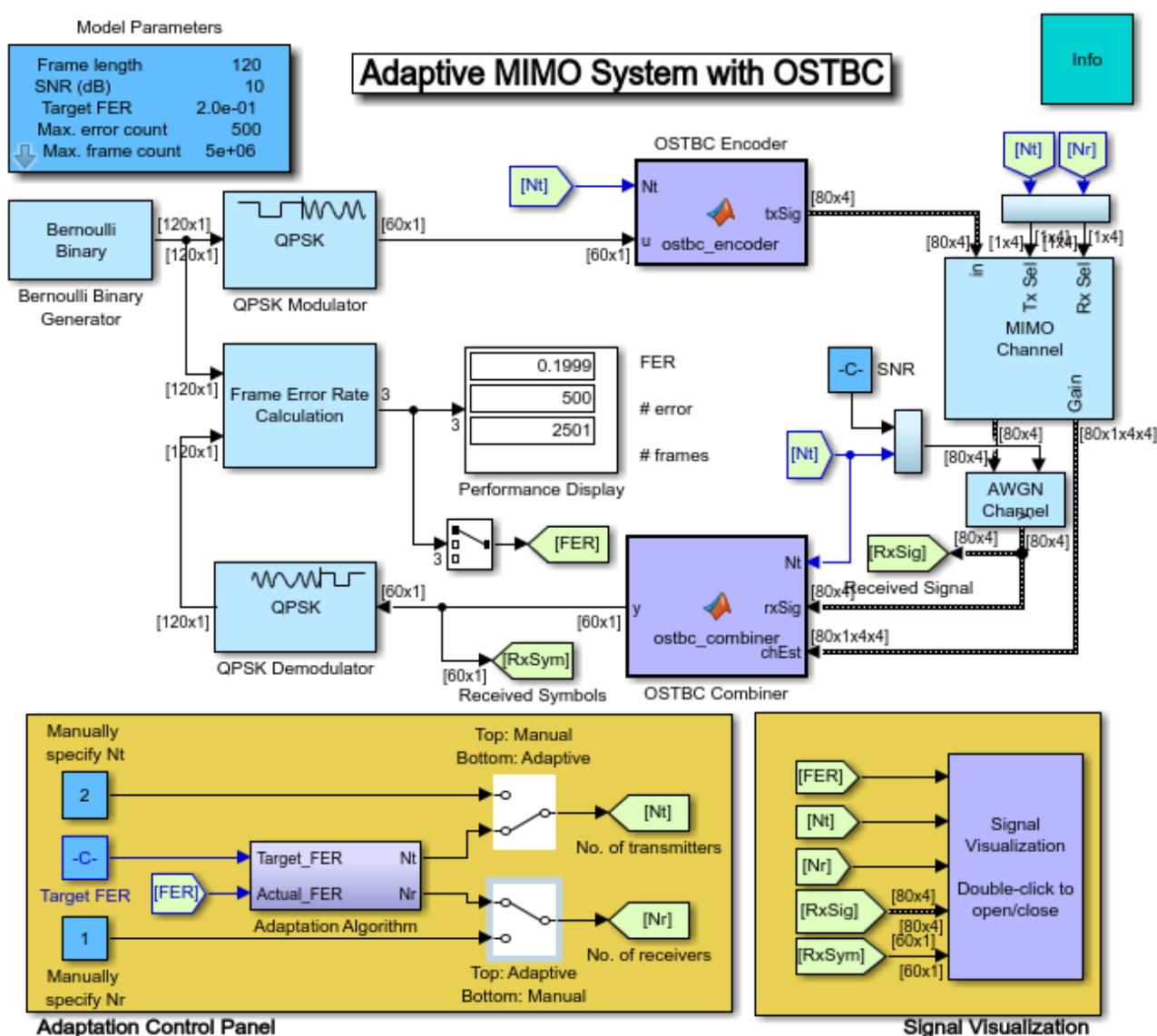


Рисунок 4.54 – Схема для исследования MIMO.

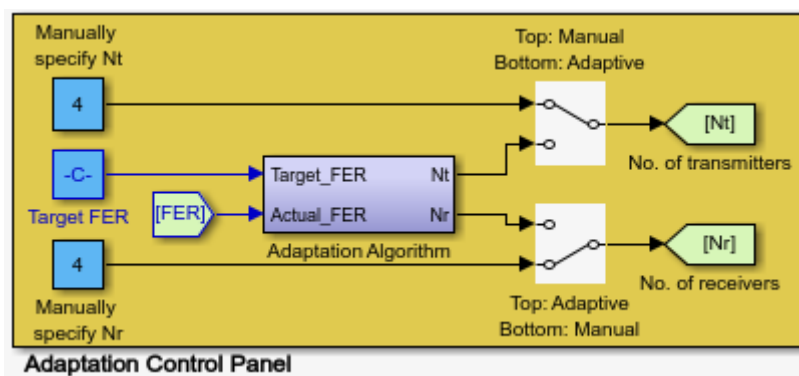


Рисунок 4.55 – Настройка блока Adaptation Control Panel для исследования случая ММО 4x4.

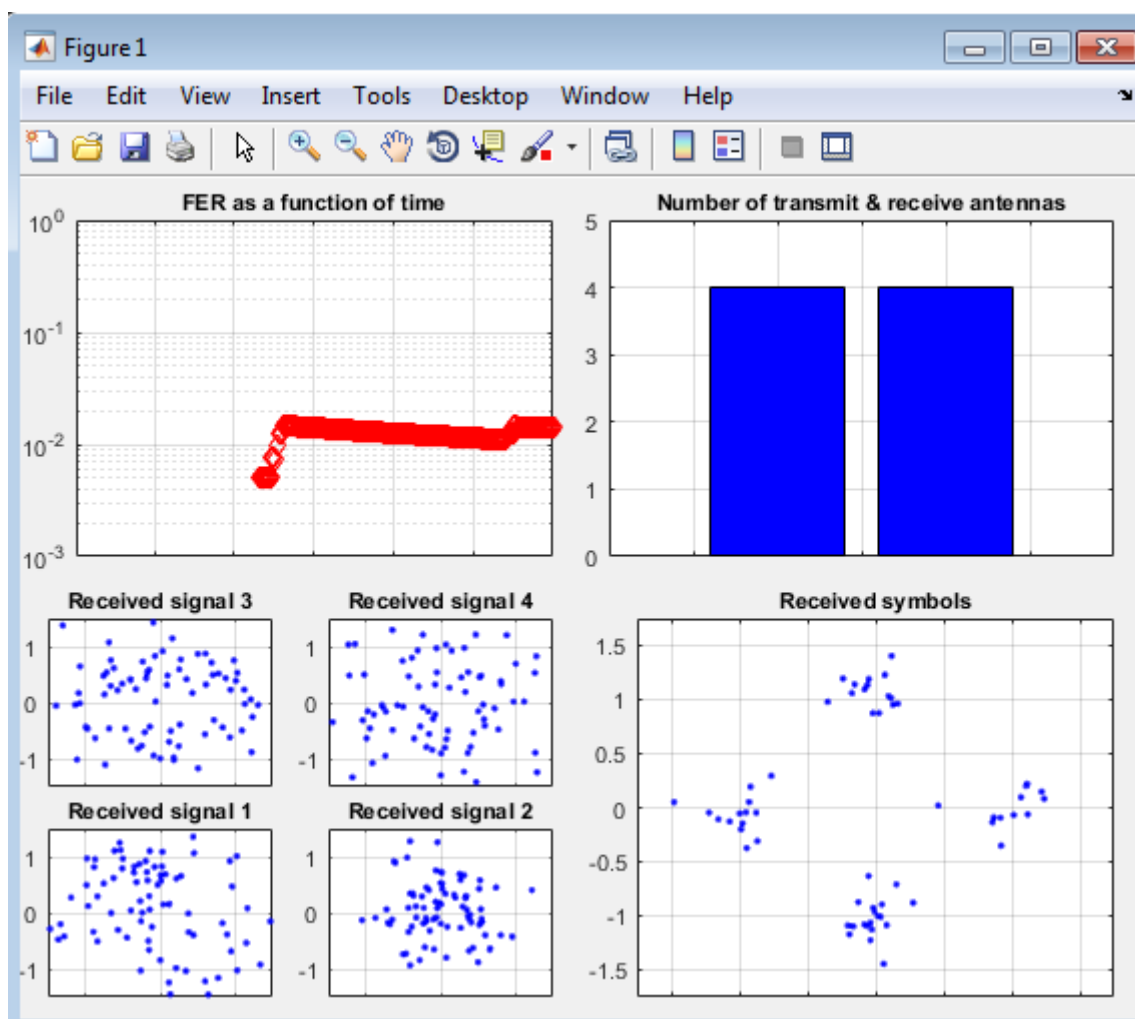


Рисунок 4.56 – Пример работы блока Signal Visualization.

5. ОПТИМИЗАЦИЯ МЕТОДОВ ПОМЕХОУСТОЙЧИВОГО КОДИРОВАНИЯ ДЛЯ ТЕЛЕКОММУНИКАЦИОННЫХ СИСТЕМ (ЗАДАНИЕ НА САМОСТОЯТЕЛЬНУЮ РАБОТУ)

Помехоустойчивое кодирование является эффективным способом оптимизации ТКС. На практике инженеру проектировщику ТКС приходится решать задачи оптимизации на основе численных расчетов и соответствующего сравнения методов помехоустойчивого кодирования и выбора конкретных методов и соответствующим им кодов. Решение именно такой задачи положено в основу СР [1].

Исходные данные заданы в таблице вариантов 5.2:

1. Цифровая информация передается двоичным кодом. Виды передаваемой цифровой информации:

ДК - данные компьютерного обмена;

ЦТЛФ - цифровая телефония;

ЦТВ - сообщения цифрового ТВ;

ЦЗВ - сообщения цифрового звукового вещания.

2. Канал связи - канал с постоянными параметрами и аддитивным белым гауссовым шумом.

3. Отношение с/ш на входе демодулятора $h_0 = E_0 / N_0$.

4. Методы модуляции: ФМ-2, ФМ-4.

5. Прием - когерентный.

6. Производительность источника $R_{ист}$ (бит/с).

7. Полоса пропускания канала F_K (кГц).

8. Вероятность ошибки бита в сообщениях, отдаваемых получателю, не более p .

9. Допустимая сложность декодера СК (показатель сложности решетки кода) - не более W .

Необходимо:

1. Выбрать и обосновать выбор корректирующего кода для проектируемой ТКС, обеспечивающего требуемую вероятность ошибки бита p в сообщениях, отдаваемых получателю, при условии выполнения следующих **ограничений**:

1.1. Полоса частот кодированного сигнала не должна превышать полосу пропускания канала F_K .

1.2. При использовании сверточных кодов показатель сложности решетки кода должен быть не более величины W .

2. Разработать и дать подробное описание *структурной и функциональных схем кодера и декодера* выбранного кода и обосновать их параметры.

3. Проанализировать показатели энергетической и частотной эффективности телекоммуникационной системы и сравнить их с предельными значениями эффективности.

4. Сделать *заключение* по выполненной работе.

Содержание пояснительной записки работы:

1. Задание и исходные данные.

2. Описание структурной схемы проектируемой телекоммуникационной системы с указанием мест включения кодера помехоустойчивого кода, модулятора, демодулятора и декодера с подробными пояснениями выполняемых ими функций.

3. Классификация корректирующих кодов по структуре. Сравнительный анализ преимуществ и недостатков помехоустойчивых блочных и сверточных кодов. Обоснование применения в проекте сверточных кодов.

4. Классификация и сравнительный анализ алгоритмов декодирования сверточных кодов. Обоснование выбора алгоритма Витерби для декодирования СК.

5. Расчет ширины спектра цифрового сигнала с заданным видом модуляции.

6. Расчет ширины спектра кодированного цифрового сигнала с заданным видом модуляции в зависимости от скорости кода.

7. Определение допустимой скорости кода $R_{КОД}^*$ из условия *непревышения* полосы частот кодированного сигнала полосы пропускания канала (ограничение 1.1).

8. Определение перечня кодов со скоростями, превышающими допустимую скорость $R_{КОД}^*$, которые могут быть использованы для решения поставленной задачи.

9. Выбор СК из этого перечня, обеспечивающего заданную вероятность ошибки бита (условие 1) и удовлетворяющего требованию ограничения по сложности декодера (ограничение 1.2).

10. Проверочный расчет зависимости вероятности ошибки на выходе декодера выбранного СК.

11. Разработка и описание структурных и функциональных схем кодера и декодера выбранного СК.

12. Заключение с подведением итогов выполненной работы.

13. Список использованных источников.

Методические указания к выполнению КР

Расчет ширины спектра сигнала ФМ-2 (ФМ-4) следует производить по рекомендациям материалов главы 1. Применение корректирующих кодов со скоростью $R_{КОД}^*$ приводит к расширению спектра кодированного сигнала в $(K_F = 1/R_{КОД})$ раз. С другой стороны, корректирующая способность кода возрастает с уменьшением скорости кода (т.е. с увеличением избыточности). Поэтому *задача оптимизации* параметров корректирующего кода состоит в выборе кода со скоростью, при которой ширина спектра кодированного сигнала *не превышает заданную полосу пропускания канала*. Если требуемая полоса пропускания канала для передачи ФМ сигнала с информационной скоростью $R_{ИСТ}$ равна $F_{(ФМ)}$, а скорость кода выбрана равной $R_{КОД}$, то полоса пропускания канала, необходимая для передачи кодированного ФМ сигнала, будет равна

$$F_{K(ФМ-СК)} = \frac{F_{(ФМ)}}{R_{КОД}}.$$

Тогда из условия неперевышения этой полосой частот сигнала полосы пропускания канала ($F_{K(ФМ-СК)} < F_K$) получаем простое *условие для выбора скорости кода*

$$R_{КОД}^* > R_{КОД} = \frac{F_{(ФМ)}}{F_K}. \quad (5.1)$$

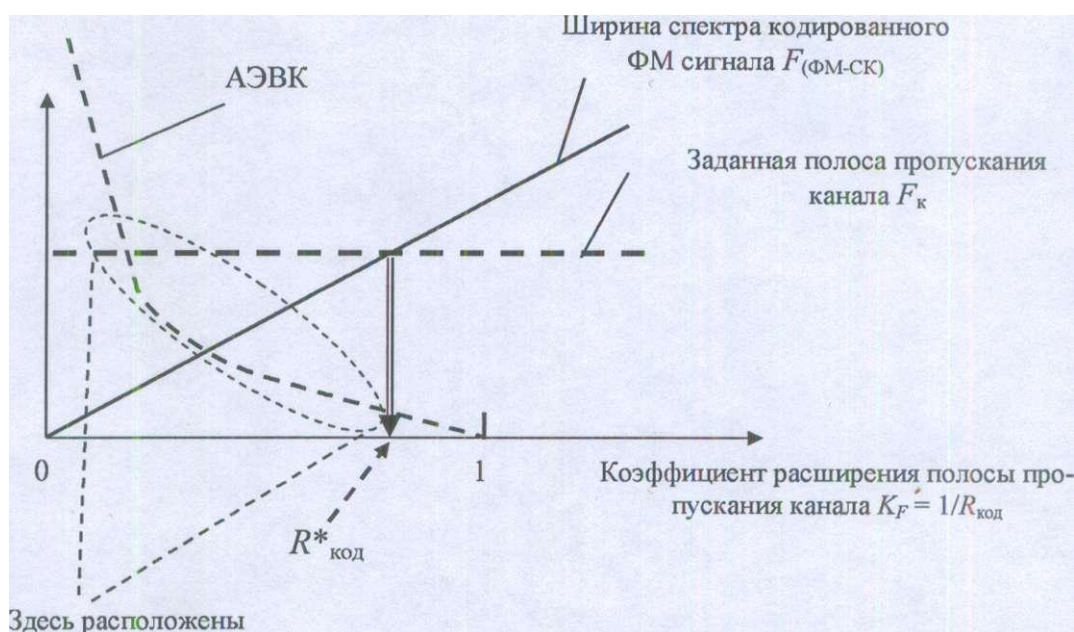
Сказанное иллюстрируется рисунком 5.1. Ширина спектра кодированного ФМ сигнала пропорциональна коэффициенту расширения полосы. По мере снижения скорости кода (возрастания K_F) полоса расширяется и достигает значения полосы пропускания канала. На этом же рисунке показана зависимость АЭВК от K_F (что равноценно скорости кода). Пересечение кривой полосы с граничным заданным значением F_K^* определяет допустимое значение коэффициента расширения полосы пропускания канала $K_p = 1/R_{КОД}$ и, соответственно, скорость кода $R_{КОД}^*$. Первым этапом выбора корректирующего кода является выбор класса кодов (класс блоковых либо непрерывных (сверточных) кодов). Используя материалы разделов 8 и 11, рекомендуется *аргументированно обосновать выбор класса сверточных кодов* для применения в своей работе. Среди алгоритмов декодирования СК по широте практического применения *лидирующее место* занимает алгоритм Витерби. Рекомендуется в работе применить именно алгоритм Витерби. В разделе проекта с обоснованием применения этого алгоритма следует привести сведения о сложности реализации алгоритма. Среди кодов, отобранных по критерию скорости в соответствии с формулой (5.1), могут оказаться коды с различной длиной кодового ограничения (и,

соответственно, с различной сложностью декодера). Помехоустойчивость декодирования СК характеризуется величиной ЭВК. В таблицах кодов не приводятся значения ЭВК при определенном уровне вероятности ошибки декодирования. В то же время, величина асимптотического энергетического выигрыша (АЭВК) является верхней оценкой ЭВК. Поэтому при отборе кодов рекомендуется использовать величины АЭВК, значения которых имеются в таблицах приложения А. Среди отобранных кодов-кандидатов следует применить код, обеспечивающий максимальный АЭВК и удовлетворяющий требованиям по скорости и сложности декодера. Окончательные данные о вероятности ошибки на выходе декодера следует получить на основе расчетов зависимости вероятности ошибки декодирования от отношения сигнал/шум для выбранного кода. В случае невыполнения требований задания рекомендуется применить код с большей величиной АЭВК.

Пример расчетов и процедуры оптимизации кода

Исходные данные:

1. Вид передаваемой цифровой информации - ЦТЛФ.
3. Отношение с/ш $h_0 = 4$ дБ.
4. Метод модуляции: ФМ-4.
5. Прием-когерентный.
6. Производительность источника $R_{ист} = 64$ кбит/с
7. Ширина полосы частот канала $F_K = 100$ кГц.
8. Допустимая вероятность ошибки бита $p = 10^{-5}$.
9. Допустимая сложность решетки кода $W = 150$.



Здесь расположены коды-кандидаты на выбор

Рис. 5.1. К процедуре оптимизации кода

1. Расчет полосы пропускания канала связи, необходимой для передачи цифровой информации с заданной скоростью методом ФМ-4, производим по формуле

$F_{(\text{ФМ-4})} = [R_{\text{ИСТ}}(1 - \alpha)]/2$, где α - коэффициент ската спектра. Задавая значение $\alpha = 0,4$, получаем $F_{(\text{ФМ-4})} = [R_{\text{ИСТ}}(1 - \alpha)]/2 = [64(1 + 0,4)]/2 = 44,8$ кГц.

2. В соответствии с формулой (5.1) определяем предельное значение скорости $R_{\text{К}}$

$$R_{\text{КОД}}^* > \frac{F_{\text{ФМ-СК}}}{F_{\text{К}}} = \frac{44,8}{100} = 0,448.$$

3. По таблицам СК отбираем коды, удовлетворяющие требованию по скорости. Данные об этих кодах сведены в таб. 5.1.

Таблица 5.1-Характеристики СК для выбора кода

Скорость кода $R_{\text{КОД}}$	Порождающие многочлены	ДКО v	Сложность решетки W	АЭВК дБ
1/4	463,535,733,745	8	512	8,29
1/3	557,663,711	8	512	7,78
1/2	53,75	5	64	6,02
1/2	61,73	5	64	6,02
1/2	71,73	5	64	6,02
1/2	133,171	6	128	6,99
1/2	247,371	7	256	6,99

Из таблицы видно, что для выполнения поставленной задачи могут быть использованы СК со скоростями $R_{\text{КОД}} = 1/2$, которые обеспечивают достаточно большой АЭВК. На основе данных таблицы выбираем для проекта код с порождающими многочленами (133, 171),

который при скорости $R_{код} = 0,5$ обеспечивает АЭВК = 6,99 дБ. Данные расчета вероятности ошибки приведены в главе 1.

Видно, что применение выбраного кода обеспечивает выполнение задания: при отношении сигнал/шум $h_o^2 = 4$ дБ вероятность ошибки декодирования менее $3 \cdot 10^{-5}$. Сравнение с кривыми помехоустойчивости некодированной ФМ (рис. 11.1) показывает, что при вероятности ошибки $p = 10^{-5}$ этот код обеспечивает ЭВК 5,3 дБ.

Таблица 5.2. Исходные данные для выполнения СР

Номер варианта для выполнения СР должен соответствовать номеру фамилии студента в журнале академической группы							
Номер варианта	Вид перед, информ.	Отношение С/Ш на входе h_o^2 , дБ	Метод на модуль.	Производит. источника $R_{ист}$, кбит/с	Полоса пропуск, канала F_k , кГц	Вер. ошибки бита p	Сложн. декодера W
1	ДК	4,0	ФМ-4	64	80	10^{-6}	150
2	ЦТЛФ	5,0	ФМ-4	16	25	10^{-4}	160
3	ЦЗВ	6,0	ФМ-2	256	800	10^{-5}	170
4	ДК	6,5	ФМ-2	64	200	10^{-6}	180
5	ЦТЛФ	4,0	ФМ-4	16	25	10^{-4}	250
6	ЦЗВ	7,0	ФМ-4	128	200	10^{-5}	350
7	НТВ	5,0	ФМ-2	2400	7000	10^{-8}	560
8	ДК	6,0	ФМ-4	32	50	10^{-6}	200
9	ЦТЛФ	5,0	ФМ-2	24	70	10^{-4}	300
10	ЦЗВ	4,5	ФМ-4	256	400	10^{-5}	250
11	ЦТВ	5,5	ФМ-2	3000	1200	10^{-8}	550
12	ДК	4,0	ФМ-4	48	70	10^{-6}	150
13	ЦТЛФ	5,0	ФМ-4	32	50	10^{-4}	250
14	ЦЗВ	7,0	ФМ-2	256	800	10^{-5}	300
15	ЦТВ	4,0	ФМ-4	4500	1300	10^{-9}	550
16	ДК	7,0	ФМ-4	56	90	10^{-6}	150

17	ЦТЛФ	5,0	ФМ-2	24	70	10^{-4}	160
18	ЦЗВ	4,5	ФМ-4	256	400	10^{-5}	200
19	ЦТВ	5,5	ФМ-4	5000	1400	10^{-9}	550
20	ДК	6,0	ФМ-2	64	200	10^{-6}	150
21	ЦТЛФ	7,5	ФМ-4	256	400	10^{-4}	250
23	ЦЗВ	6,5	ФМ-4	16	50	10^{-5}	150
24	ДК	6,0	ФМ-4	64	150	10^{-6}	150
25	ЦГЛФ	4,5	ФМ-2	16	25	10^{-6}	200
26	ЦТВ	5,0	ФМ-2	6000	16000	10^{-9}	550
27	ЦЗВ	6,0	ФМ-4	384	600	10^{-5}	250
28	ДК	4,5	ФМ-4	64	100	10^{-6}	150
29	ЦГЛФ	5,0	ФМ-2	16	50	10^{-4}	250
30	ЦТВ	5,5	ФМ-2	5500	32000	10^{-9}	560
31	ЦГЛФ	4,5	ФМ-4	64	200	10^{-5}	150
32	ДК	5,0	ФМ-4	64	300	10^{-5}	250

Примеры расчетов для разных вариантов

Вариант №7

Таблица 5.3. Параметры проектируемой ТКС

Номер варианта для выполнения индивидуальной работы должен соответствовать номеру фамилии студента в журнале академической группы							
Номер варианта	Вид информации	Отношение С/Ш h_b^2 , дБ	Метод модуляции	Пром. источник $R_{ист}$, кбит/с	Пропускная способность канала F_k , кГц	Вер. Ошибки и бита	Сло жн. декодера
7	ЦТВ	5.0	ФМ-2	2400	7000	10^{-8}	560

Структурная схема проектируемой телекоммуникационной системы

В общем виде обобщенная структурная схема проектируемой ТКС может быть сформирована в виде, представленном на рисунке 1.1.

В передатчике кодер вносит в информационное сообщение избыточность в виде проверочных символов. Закодированные символы поступают на модулятор, который преобразует их в аналоговый сигнал.

В приемнике демодулятор преобразует принятый сигнал в последовательность чисел, представляющих оценку переданных данных – метрики. Метрики поступают в декодер, который исправляет возникающие при передаче ошибки, используя внесенную кодером избыточность [24].

Классификация корректирующих кодов

Обнаружение ошибок в технике связи — действие, направленное на контроль целостности данных при записи/воспроизведении информации или при её передаче по линиям связи. Исправление ошибок (коррекция ошибок) — процедура восстановления информации после чтения её из устройства хранения или канала связи.

Для обнаружения ошибок используют коды обнаружения ошибок, для исправления — корректирующие коды (коды, исправляющие ошибки, коды с коррекцией ошибок, помехоустойчивые коды).

В общем виде классификация корректирующих кодов может быть представлена в следующем виде:

1. Блочные коды:

1.1 Линейные коды общего вида;

1.1.2 Коды Хемминга;

1.2 Линейные циклические коды:

1.2.1 Коды CRC;

1.2.2 Коды BCH;

1.2.3 Коды коррекции ошибок Рида — Соломона;

2. Сверточные коды;

3. Каскадные коды.

Стоит отметить, что блочные коды, как правило, хорошо справляются с редкими, но большими пачками ошибок, их эффективность при частых, но небольших ошибках (например, в канале с АБГШ), менее высока.

Вместе с этим, сверточные коды эффективно работают в канале с белым шумом, но плохо справляются с пакетами ошибок. Более того, если декодер ошибается, на его выходе всегда возникает пакет ошибок.

Так как в начальных условиях поставленной задачи не были сформулированы требования к методам кодирования, выбор остановился на сверточных кодах. Однако, при

проектировании телекоммуникационных систем необходимо четко формировать критерии оптимальности разрабатываемой системы.

Классификация методов декодирования сверточных кодов

Классификация методов декодирования сверточных кодов имеет следующий вид:

1. Алгебраические методы декодирования;
2. Вероятностные методы декодирования:
 - 2.1 Алгоритм последовательного декодирования;
 - 2.2 Алгоритм Витерби.

Алгоритм Витерби характеризуется постоянством вычислительной работы, однако сложность декодера Витерби растет, как при переборных алгоритмов, по экспоненциальному закону от длины кодового ограничения сверточного кода.

Так как в данной работе в целях оптимизации проектируемой системы будут использоваться короткие сверточные коды, сложность декодера будет мала, что позволяет использовать алгоритм декодирования Витерби.

Расчет и оптимизация параметров телекоммуникационной системы

Расчет ширины спектра цифрового сигнала с заданным видом модуляции:

$$F_{\Phi M-2} = \frac{R_{\text{уст}} \cdot (1 + \alpha)}{2} = \frac{2400 \cdot 10^3 \cdot (1 + 0.4)}{2} = 1.68 \text{ МГц} .$$

Расчет ширины спектра кодированного цифрового сигнала с заданным видом модуляции в зависимости от скорости кода:

$$R_{\text{код}^*} = \frac{F_{\Phi M-2}}{F_{\kappa}} = \frac{1680 \cdot 10^3}{7000 \cdot 10^3} = 0.24 .$$

Следовательно скорость кода должна быть не менее 0.24. Полученный результат позволяет сформировать список подходящих сверточных кодов в виде представленном в таблице 5.3.

Таблица 5.4. Перечень подходящих сверточных кодов

Скорость кода $R_{\text{код}}$	Порождающие многочлены	ДКО ν	Сложность решетки W	АЭВК, дБ
1/4	463,535,733,745	8	512	8,29
1/3	557,663,711	8	512	7,78
1/2	53,75	5	64	6,02
1/2	61,73	5	64	6,02
1/2	71,73	5	64	6,02
1/2	133,171	6	128	6,99
1/2	247,371	7	256	6,99

В силу того, критерием оптимальности проектируемой ТКС является простота используемого кодера/декодера, был выбран код /133,171/ с длиной кодового ограничения 7, который при скорости кода 0.5 обеспечивает АЭВК = 6.99 дБ.

Изложенное позволяет рассчитать ширину спектра кодированного цифрового сигнала:

$$F_{\Phi M-2+CK} = \frac{F_{\Phi M-2}}{R_{код}} = \frac{1680 \cdot 10^3}{0.5} = 3.36 \text{ МГц}$$

Рисунок 6.2 позволяет сделать вывод о том, что применение выбранного кода обеспечивает выполнение поставленной задачи, так как при отношении С/Ш = 5 дБ вероятность ошибки декодирования меньше 10^{-5} .

Сравнение с кривыми помехоустойчивости некодированной ФМ показывает, что при вероятности ошибки 10^{-8} этот код обеспечивает значение ЭВК более 10 дБ.

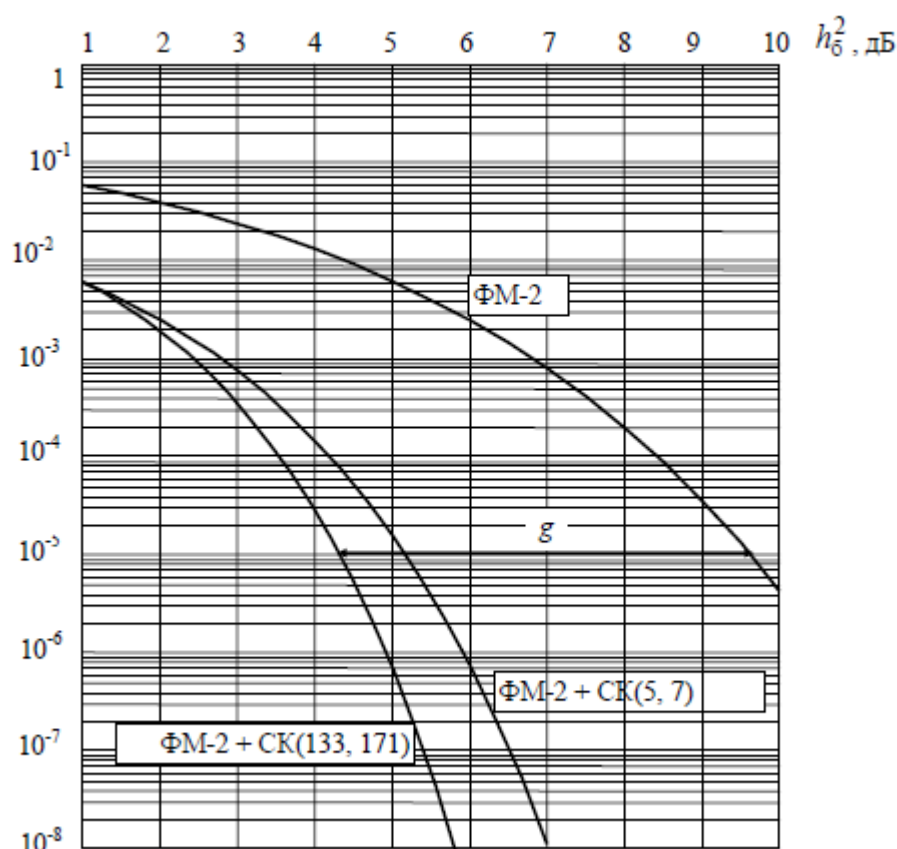


Рис. 5.2. Помехоустойчивость декодирования сверточных кодов

Проверочный расчет вероятности ошибки на выходе декодера:

$$Q = 0.65 \cdot \exp(-0.44 \cdot (z + 0.75)^2) = 0.65 \cdot \exp(-0.44 \cdot (5.01 + 0.75)^2) = 2.972 \cdot 10^{-7}$$

$$p_o = w_{df} \cdot Q \cdot (\sqrt{2 \cdot d_f \cdot R_{код} \cdot h_o^2}) = 36 \cdot 2.972 \cdot 10^{-7} \cdot (\sqrt{2 \cdot 10 \cdot 0.5 \cdot 5}) = 7.565 \cdot 10^{-5}$$

Расчет показал, что реальное значение вероятности ошибки кодера меньше теоретического значения, следовательно, условия задачи были выполнены.

Разработка кодера и декодера сверточного кода 133,171

В предыдущем разделе был описан выбор сверточного кодера /133,171/. Функциональная и структурная схема кодера/декодера может быть представлена в следующем виде:

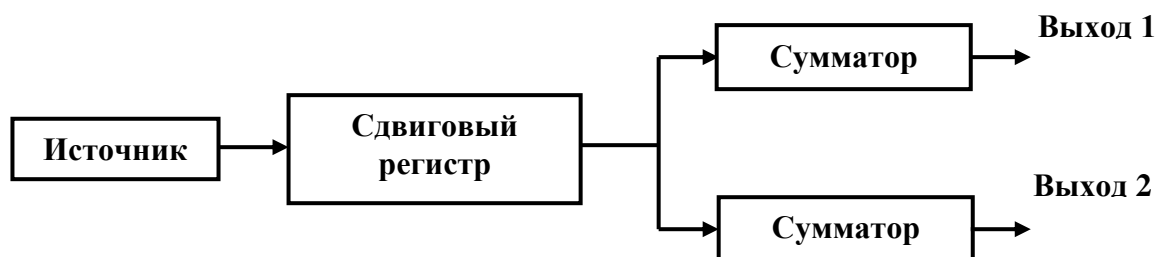


Рис. 5.3. Структурная схема сверточного кодера

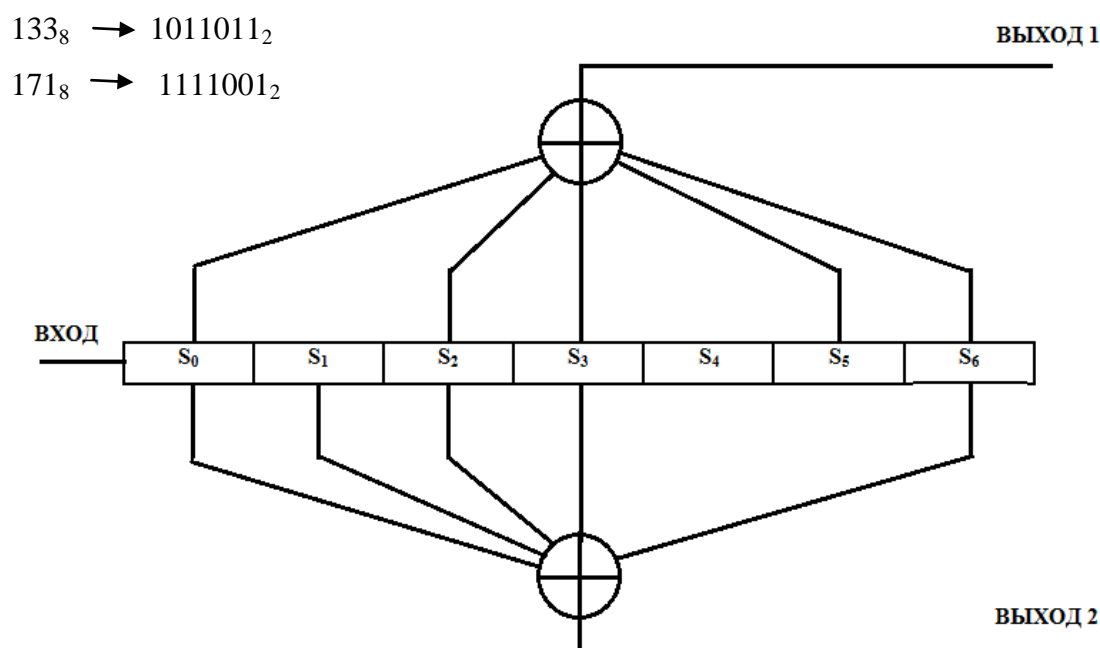


Рис. 5.4. Функциональная схема сверточного кодера 133,171

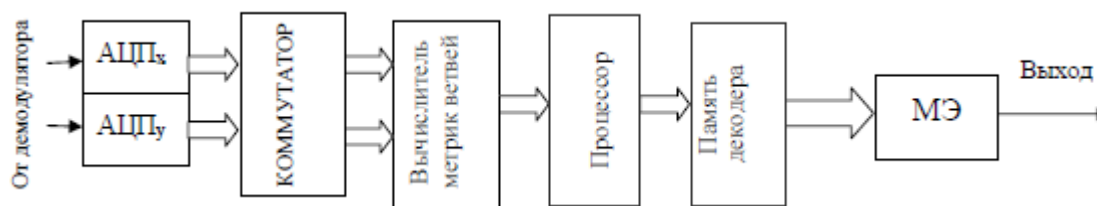


Рис. 5.5. Структурная схема декодера Витерби

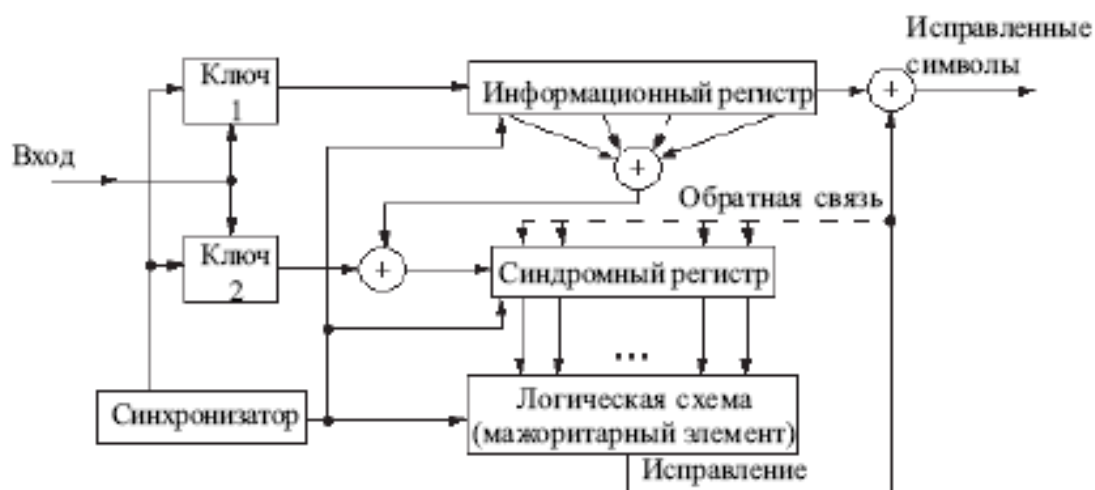


Рис. 5.6. Функциональная схема декодера Витерби

В результате выполнения данной индивидуальной работы было сделано следующее:

1. Спроектирована телекоммуникационная система с использованием сверточного кодера;
2. Рассчитаны и оптимизированы параметры сверточного кода используемого в ТКС в целях повышения ее эффективности и помехоустойчивости;
3. Предложены структурные и функциональные схемы кодера и декодера, используемых в разработанной ТКС.

Варианты № 16, 3, 8

Для решения поставленной задачи предложены общие параметры проектируемой ТКС, которые представлены в таблице 5.5.

Таблица 5.5. Параметры проектируемой ТКС

Ном ер варианта	Ви д перед. инф-ии	Отн ошение С/Ш h_0^2 , дБ	Метод модуляции	Произв . источника $R_{ист}$, кбит/с	Пропуск ная способность канала F_k , кГц	Вер. ошибки бита p	Слож н. декодера W
16	ДК	7,0	ФМ-4	56	90	10^{-6}	150
3	ЦЗ В	6,0	ФМ-2	256	800	10^{-5}	170
8	ДК	6,0	ФМ-4	32	50	10^{-6}	200

Структурная схема проектируемой телекоммуникационной системы

Структурная схема проектируемой телекоммуникационной системы представлена на рисунке 1.2.

Источник сообщения генерирует бинарную последовательность с определенной скоростью $R_{\text{ист}}$. Курсивом отмечены блоки, которые кодируют и декодируют информацию с применением помехоустойчивых кодов (вводится избыточность при кодировании, например код Хемминга, БЧХ, сверточный код). Что касается источника, то он кодируется и декодируется с помощью таких алгоритмов как, Хаффмана, Шеннона-Фано или Лемпел-Зива. В данных алгоритмах не вводится избыточность. Помимо кодирования система связи содержит в себе квадратурную модуляцию/демодуляцию. Где на выходе модулятора мы получаем сначала комплексные числа (квадратурные и синфазные составляющие), которые в свою очередь садятся на несущие, сдвинутые на 90 градусов и в конечном итоге суммируются. Демодуляция представляет собой обратный процесс. Варианты работы содержит в себе модуляцию ФМ-2 или BPSK, которая имеет только два синфазных значения постоянной амплитуды и фазы 0 и 180 градусов и ФМ-4 или QPSK, которая имеет четыре значения постоянной амплитуды и фазы. И, конечно же, любая система передачи не обходится без воздействия на нее шумов, в канале беспроводной сети (канал связи).

4 Классификация корректирующих кодов

Обнаружение ошибок в технике связи — действие, направленное на контроль целостности данных при записи/воспроизведении информации или при её передаче по линиям связи. Исправление ошибок (коррекция ошибок) — процедура восстановления информации после чтения её из устройства хранения или канала связи.

Для обнаружения ошибок используют коды обнаружения ошибок, для исправления — корректирующие коды (коды, исправляющие ошибки, коды с коррекцией ошибок, помехоустойчивые коды).

Преимущества и недостатки блоковых кодов:

Блоковые коды, как правило, хорошо справляются с редкими, но большими пачками ошибок, их эффективность при частых, но небольших ошибках (например, в канале с АБГШ), менее высока.

Преимущества и недостатки сверточных кодов:

Сверточные коды эффективно работают в канале с белым шумом, но плохо справляются с пакетами ошибок. Более того, если декодер ошибается, на его выходе всегда возникает пакет ошибок. Выбор в индивидуальной работе сверточных кодов обосновывается тем, что

свёрточное кодирование - очень простая операция. Кодирование свёрточным кодом производится с помощью регистра сдвига, отводы от которого суммируются по модулю два. Таких сумм может быть две (чаще всего) или больше.

Классификация корректирующих кодов по структуре представлена на рисунке в.

Классификация методов декодирования сверточных кодов

Классификация методов декодирования сверточных кодов имеет следующий вид:

3. Алгебраические методы декодирования;
4. Вероятностные методы декодирования:
 - 4.1 Алгоритм последовательного декодирования;
 - 4.2 Алгоритм Витерби.

Задача декодирования сверточного кода заключается в выборе пути (в этом и состоит отличие декодирования сверточных кодов) вдоль решетки наиболее похожего на принятую последовательность. Каждый путь вдоль решетчатой диаграммы складывается из ветвей соединяющих узлы. Каждой ветви решетки соответствует кодовое слово из двух бит. Каждую ветвь на каждом периоде можно пометить расстоянием Хемминга между полученным кодовым словом и кодовым словом, соответствующим ветви. Складывая расстояния Хемминга ветвей, составляющих путь, получим метрику соответствующего пути. Данная метрика будет характеризовать степень подобия каждого пути принятой последовательности. Чем меньше метрика, тем более похожи путь и принятая последовательность. Таким образом, результатом декодирования будет информационная последовательность, соответствующая пути с минимальной метрикой. Если в одно и тоже состояние входят два пути выбирается тот, который имеет лучшую метрику. Такой путь называется выжившим. Отбор выживших путей проводится для каждого состояния. Это не иначе как алгоритм декодирования Витерби и он наиболее эффективный.

Расчет ширины спектра цифрового сигнала с заданным видом модуляции

Вариант	Расчеты
16	$F_{ФМ4} = \frac{R_{ист} * (1 + \alpha)}{2} = \frac{56 * (1 + 0,4)}{2} = 39,2 \text{ кГц}$
3	$F_{ФМ4} = \frac{R_{ист} * (1 + \alpha)}{2} = \frac{256 * (1 + 0,4)}{2} = 179,2 \text{ кГц}$
8	$F_{ФМ4} = \frac{R_{ист} * (1 + \alpha)}{2} = \frac{32 * (1 + 0,4)}{2} = 22,4 \text{ кГц}$

Определение допустимой скорости кода из условия непревышения полосой частот кодированного сигнала полосы пропускания канала

Вариант	Расчеты

16	$R_{\text{код}^*} = \frac{F_{\text{ФМ4}}}{F_K} = 0,436$
3	$R_{\text{код}^*} = \frac{F_{\text{ФМ4}}}{F_K} = 0,224$
8	$R_{\text{код}^*} = \frac{F_{\text{ФМ4}}}{F_K} = 0,448$

Определение кода

Полученный результат позволяет сформировать список подходящих сверточных кодов в виде, представленном в таблице 5.6.

Таблица 5.6. Характеристики СК для выбора кода

Скорость кода $R_{\text{код}}$	Порождающие многочлены	ДКО ν	Сложность решетки W	АЭВК, дБ
1/4	463,535,733,745	8	512	8,29
1/3	557,663,711	8	512	7,78
1/2	53,75	5	64	6,02
1/2	61,73	5	64	6,02
1/2	71,73	5	64	6,02
1/2	133,171	6	128	6,99
1/2	247,371	7	256	6,99

Вариант	Условия
16	СК со скоростями 1/2 и сложностью решетки W не более 150
3	Все СК со сложностью решетки W не более 170
8	СК со скоростями 1/2 и сложностью решетки W не более 200

Произведен выбор СК из перечня, обеспечивающего заданную вероятность ошибки бита и удовлетворяющего требованию ограничения по сложности декодера.

Вариант	Выбранный СК
16	Код с порождающими многочленами (133, 171), который при скорости 1/2 обеспечивает АЭВК = 6,99 дБ
3	Код с порождающими многочленами (133, 171), который при скорости 1/2 обеспечивает АЭВК = 6,99 дБ
8	Код с порождающими многочленами (133, 171), который при скорости 1/2 обеспечивает АЭВК = 6,99 дБ

Расчет ширины спектра кодированного цифрового сигнала с заданным видом модуляции в зависимости от скорости кода

Вариант	Расчеты
---------	---------

иант	
16	$F_{ФМ4+СК} = \frac{F_{ФМ4}}{R_{код}} = \frac{39,2}{0,5} = 78,4 \text{ кГц}$
3	$F_{ФМ2+СК} = \frac{F_{ФМ2}}{R_{код}} = \frac{179,2}{0,5} = 358,4 \text{ кГц}$
8	$F_{ФМ4+СК} = \frac{F_{ФМ4}}{R_{код}} = \frac{22,4}{0,5} = 44,8 \text{ кГц}$

Рисунок 5.7 позволяет сделать вывод о том, что применение выбранного кода обеспечивает выполнение поставленной задачи, так как

Вариант	Отношение С/Ш h_0^2 , дБ	Вероятность ошибки декодирования меньше
16	7,0	10^{-6}
3	6,0	10^{-5}
8	6,0	10^{-6}

Сравнение с кривыми помехоустойчивости некодированной ФМ показывает, что

Вариант	Вероятность ошибки	АЭВК, дБ
16	10^{-6}	более 10
3	10^{-5}	9,4
8	10^{-6}	более 10

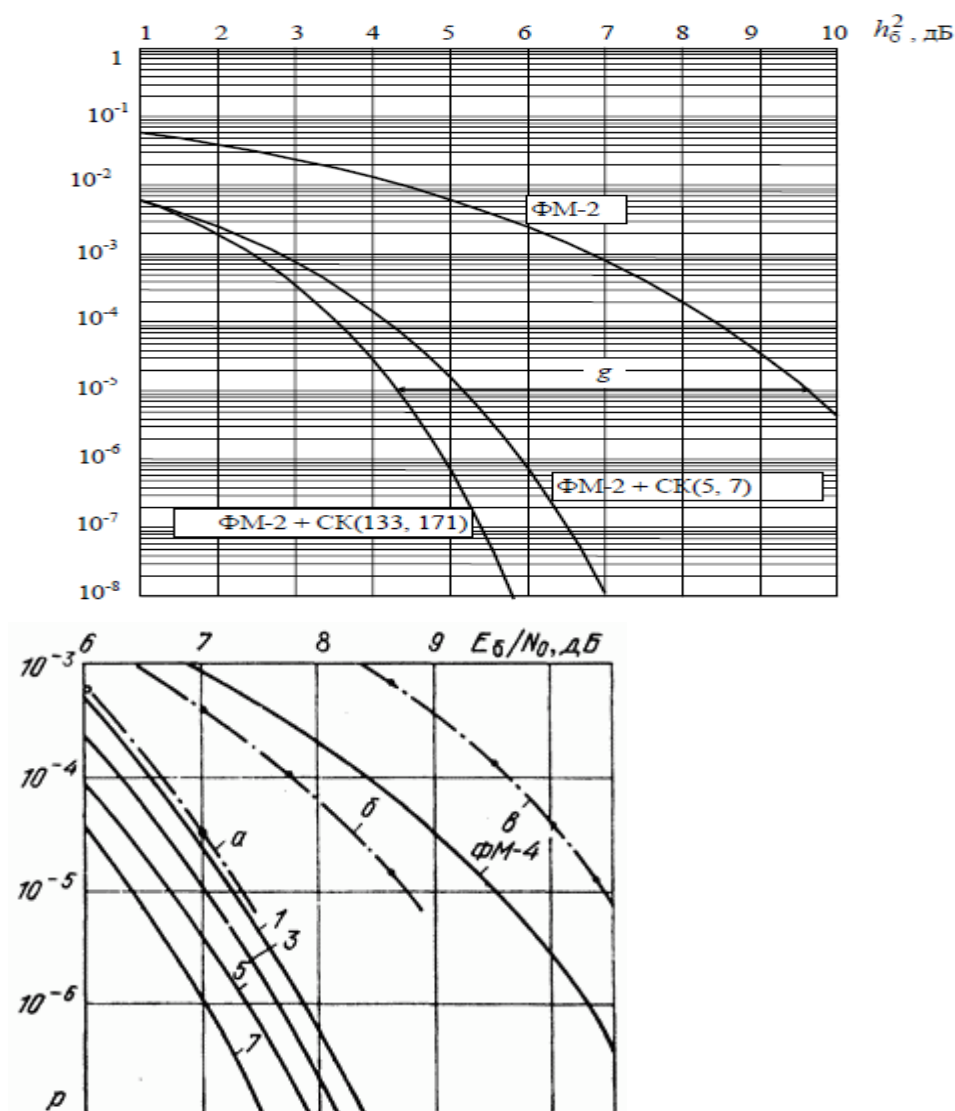


Рис. 5.7. Помехоустойчивость декодирования сверточных кодов

Проверочный расчет зависимости вероятности ошибки на выходе декодера

В результате получим (примерно для заданной вероятности ошибки бита):

Вариант	Расчеты
16	$Q = \frac{1}{x \cdot \sqrt{2\pi}} * \exp\left(-\frac{x^2}{2}\right) = \frac{1}{5,01 * \sqrt{2\pi}} * \exp\left(-\frac{5,01^2}{2}\right) = 4,45 * 10^{-5}$ $p_d = w_{df} \cdot Q \cdot \sqrt{2 \cdot d_f \cdot R_{kod} \cdot h_b} = 36 \cdot 4,45 \cdot 10^{-5} \cdot \sqrt{2 \cdot 10 \cdot 0,5 \cdot 7} = 4,2 \cdot 10^{-3}$
3	$Q = \frac{1}{x \cdot \sqrt{2\pi}} * \exp\left(-\frac{x^2}{2}\right) = \frac{1}{4 * \sqrt{2\pi}} * \exp\left(-\frac{4^2}{2}\right) = 3,3 * 10^{-5}$ $p_d = w_{df} \cdot Q \cdot \sqrt{2 \cdot d_f \cdot R_{kod} \cdot h_b} = 36 \cdot 3,3 \cdot 10^{-5} \cdot \sqrt{2 \cdot 10 \cdot 0,5 \cdot 6} = 9,2 \cdot 10^{-3}$

8	$Q = \frac{1}{x \cdot \sqrt{2\pi}} \cdot \exp\left(-\frac{x^2}{2}\right) = \frac{1}{5,01 \cdot \sqrt{2\pi}} \cdot \exp\left(-\frac{5,01^2}{2}\right) = 4,45 \cdot 10^{-5}$ $p_d = w_{df} \cdot Q \cdot \sqrt{2 \cdot d_f \cdot R_{kod} \cdot h_b} = 36 \cdot 4,45 \cdot 10^{-5} \cdot \sqrt{2 \cdot 10 \cdot 0,5 \cdot 6} = 4,2 \cdot 10^{-3}$
---	--

Расчет показал, что реальное значение вероятности ошибки кодера меньше теоретического значения, следовательно, условия задачи были выполнены.

Разработка кодера и декодера СК 133, 171

В предыдущем разделе был описан выбор сверточного кодера (133,171).

$$133_8 = 1011011_2; 171_8 = 1111001_2$$

Функциональная и структура схема кодера/декодера может быть представлена в следующем виде:

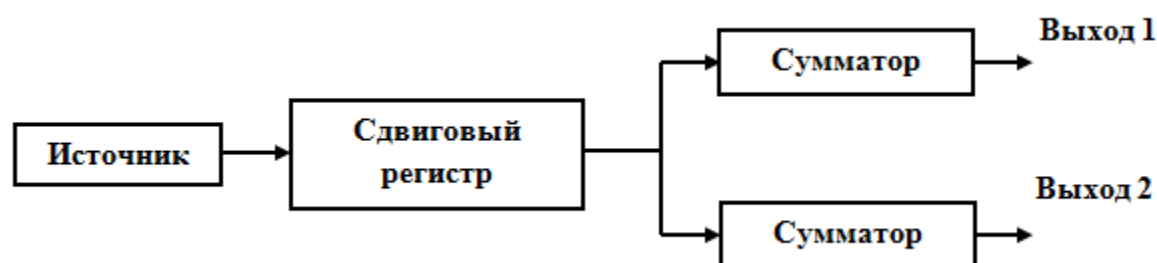


Рис. 5.8. Структурная схема сверточного кодера

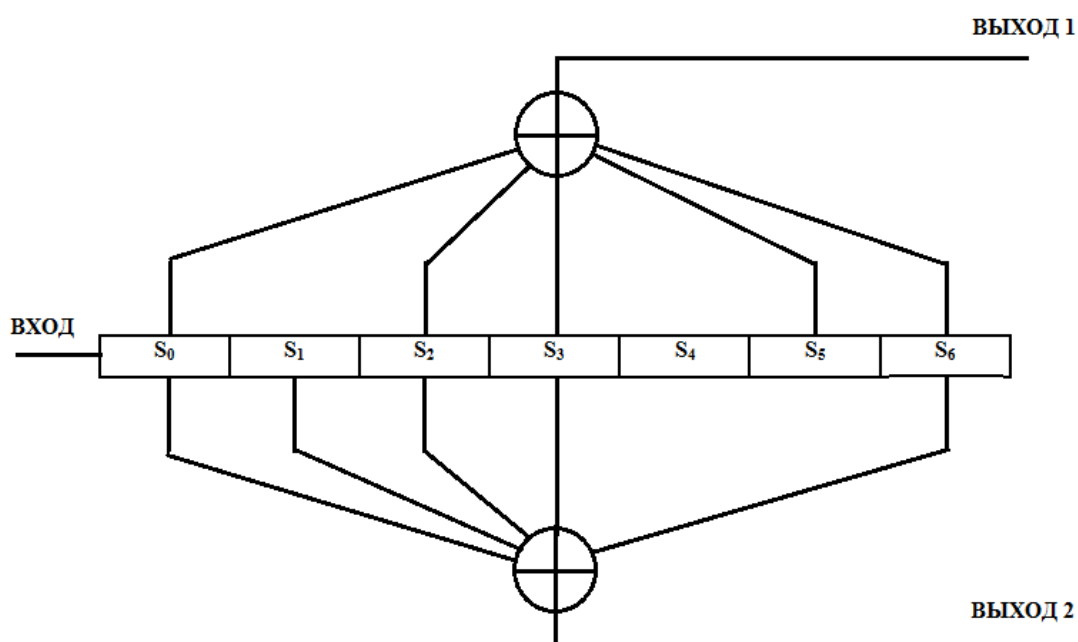


Рис. 5.9. Функциональная схема сверточного кодера 133,171

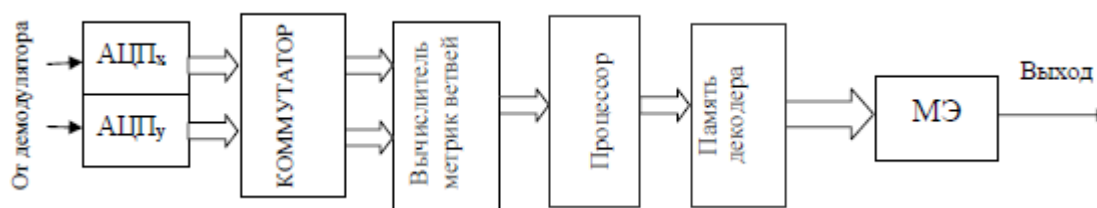


Рис. 5.10. Структурная схема декодера Витерби

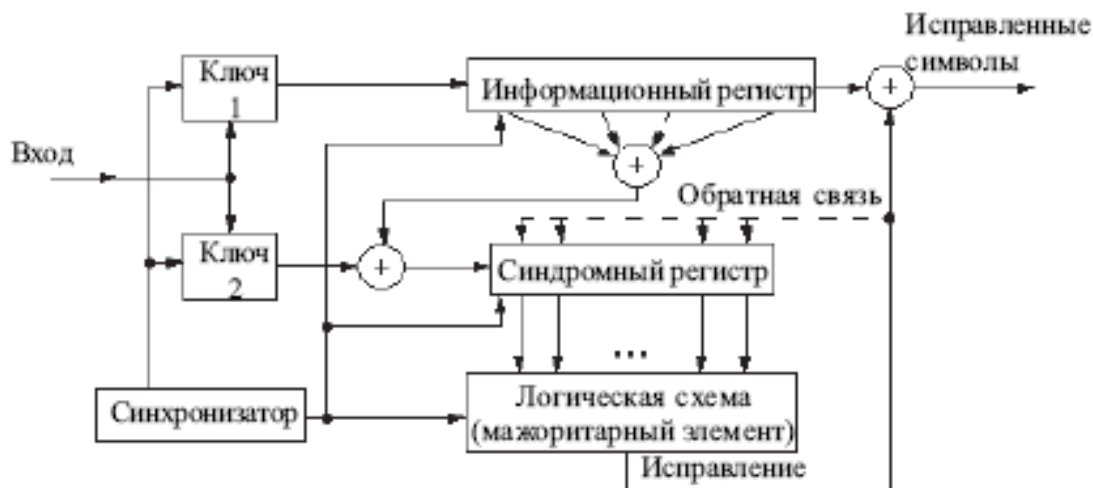


Рис. 5.11. Функциональная схема декодера Витерби

кодера со скоростью 1/2.

В результате выполнения данного индивидуального задания было выполнено следующее:

- Спроектирована телекоммуникационная система с использованием сверточного кодера;
- Рассчитаны и оптимизированы параметры сверточного кода используемого в ТКС в целях повышения ее эффективности и помехоустойчивости при различных начальных заданных условиях (ширина спектра, скорость кода, битовая вероятность ошибки в зависимости от заданного значения отношения сигнал/шум);
- Предложены структурные и функциональные схемы кодера и декодера, используемых в разработанной ТКС.

ЧАСТЬ 2. ШИФРОВАНИЕ В РАДИОЭЛЕКТРОННЫХ СИСТЕМАХ ПЕРЕДАЧИ ИНФОРМАЦИИ

6. КЛАССИЧЕСКИЕ ШИФРЫ

Разные люди понимают под шифрованием разные вещи. Дети играют в игрушечные шифры и секретные языки. Это, однако, не имеет ничего общего с настоящей криптографией. Настоящая криптография (strong cryptography) должна обеспечивать такой

уровень секретности, чтобы вы имели возможность надежно защитить критическую информацию от расшифровки крупными организациями — такими как мафия, транснациональные корпорации и крупные государства. Настоящая криптография в прошлом использовалась лишь в военных целях. Однако сейчас, со становлением информационного общества, она становится центральным инструментом для обеспечения конфиденциальности.

По мере образования информационного общества, крупным государствам становятся доступны Технологические средства тотального надзора за миллионами людей. Поэтому криптография становится одним из основных инструментов обеспечивающих конфиденциальность, доверие, авторизацию, электронные платежи, корпоративную безопасность и бесчисленное множество других важных вещей.

Криптография не является более придумкой военных, с которой не стоит связываться. Настала пора снять с криптографии покровы таинственности и использовать все ее возможности на пользу современному обществу. Широкое распространение криптографии является одним из немногих способов защитить человека от ситуации, когда он вдруг обнаруживает, что живет в тоталитарном государстве, которое может контролировать каждый его шаг.

Основные термины криптографии и понятия

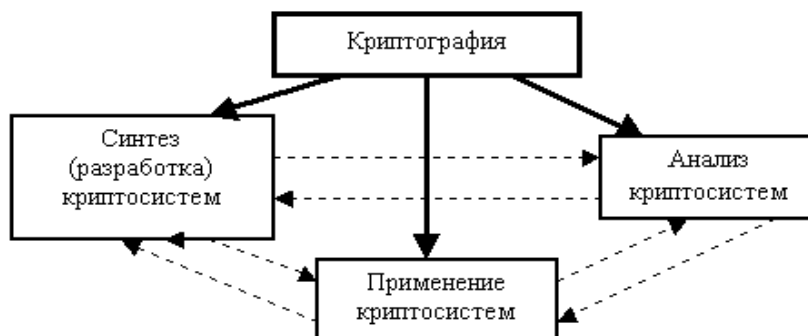
Криптография – до 70-х гг. XX в. – область науки и практической деятельности, связанная с разработкой, применением и анализом шифрсистем; в настоящее время – область науки, техники и практической деятельности, связанная с разработкой, применением и анализом криптографических систем защиты информации. Основными функциями криптографических систем являются обеспечение конфиденциальности и аутентификации различных аспектов информационного взаимодействия. Источником угроз при решении криптографических задач считаются преднамеренные действия противника или недобросовестного участника информационного взаимодействия, а не случайные искажения информации вследствие помех, отказов и т. п.

Конфиденциальность – защищенность информации от ознакомления с ее содержанием со стороны лиц, не имеющих права доступа к ней.

Аутентификация – установление (то есть проверка и подтверждение) подлинности различных аспектов информационного взаимодействия: сеанса связи, сторон (идентификация), содержания (имитозащита) и источника (установление авторства) передаваемых сообщений, времени взаимодействия и т. д. Является важной составной частью проблемы обеспечения достоверности получаемой информации. Особенно остро эта проблема стоит в случае не доверяющих друг другу сторон, когда источником угроз может

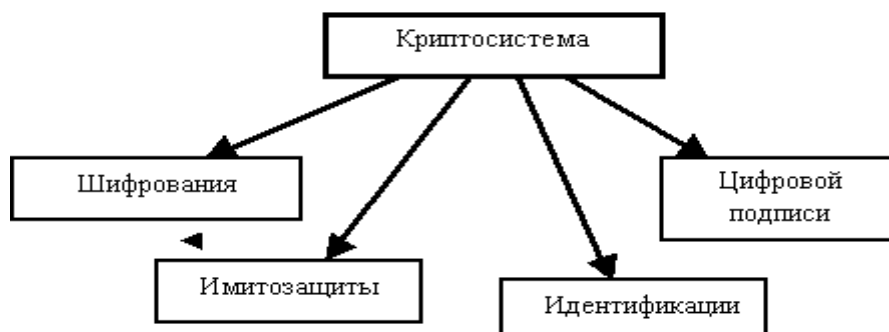
служить не только третья сторона (противник), но и сторона, с которой осуществляется информационное взаимодействие.

Определение криптографии и показывает основные составляющие ее части. Пунктирные стрелки показывают тесные взаимосвязи между этими тремя составляющими.



Виды криптосистем

Система криптографическая (криптосистема) – система обеспечения безопасности защищенной сети, использующая криптографические средства. В качестве подсистем может включать системы шифрования, идентификации, имитозащиты, цифровой подписи и др., а также ключевую систему, обеспечивающую работу остальных систем. В основе выбора и построения криптосистемы лежит условие обеспечения криптографической стойкости. В зависимости от ключевой системы различают симметричные и асимметричные криптосистемы.



Средства криптографические – в широком смысле – методы и средства обеспечения безопасности информации, использующие криптографические преобразования информации; в узком смысле – средства, реализованные в виде документов, механических, электро-механических, электронных технических устройств или программ, предназначенные для выполнения функций криптографической системы.

Криптографическое преобразование информации – преобразование информации с использованием одного из криптографических алгоритмов, определяемое целевым назначением криптографической системы.

Симметричные криптосистемы – криптосистемы с симметричными (секретными) ключами. Симметричность означает здесь, что ключи, задающие пару взаимно обратных криптографических преобразований, могут быть получены один из другого с небольшой трудоемкостью. Стойкость симметричной криптосистемы определяется трудоемкостью, с которой противник может вычислить любой из секретных ключей, и оценивается при общепринятом допущении, что противнику известны все элементы криптосистемы, за исключением секретного ключа.

Асимметричные криптосистемы – криптосистемы с асимметричными (секретными и открытыми) ключами. Асимметричность означает здесь, что из двух ключей, задающих пару взаимно обратных криптографических преобразований, один является секретным, а другой открытым. Открытые ключи известны всем участникам защищенной сети и противнику, но каждый участник сети хранит в тайне собственный секретный ключ. Стойкость асимметричной криптосистемы определяется трудоемкостью, с которой противник может вычислить секретный ключ, исходя из знания открытого ключа и другой дополнительной информации о криптосистеме.

Шифрсистема – криптографическая система обеспечения конфиденциальности, предназначенная для защиты информации от ознакомления с ее содержанием лиц, не имеющих права доступа к ней, путем шифрования информации. Математическая модель шифрсистемы включает способ кодирования исходной и выходной информации, *шифр* и ключевую систему.

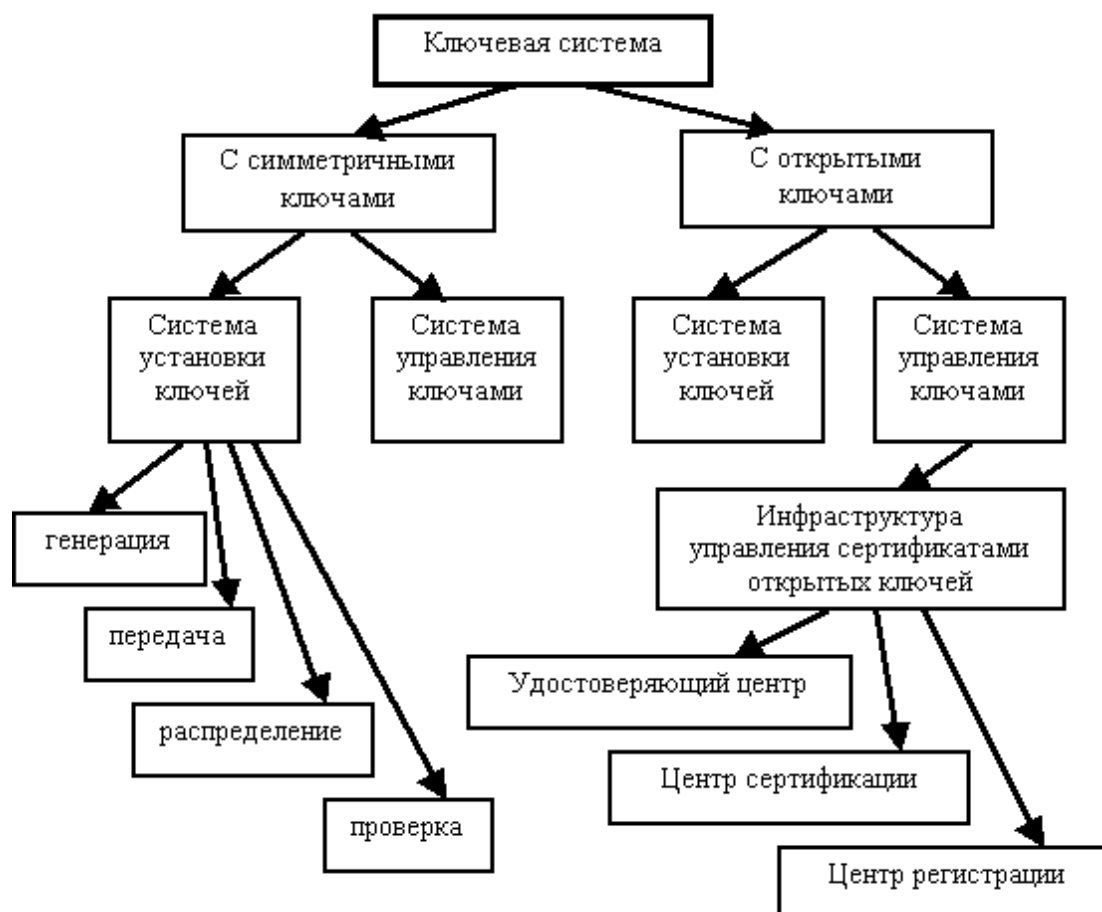
Система имитозащиты (обеспечения целостности) информации – криптографическая система, выполняющая функцию аутентификации содержания сообщения или документа и предназначенная для защиты от несанкционированного изменения информации или навязывания ложной информации. Математическая модель системы имитозащиты включает криптографический алгоритм имитозащищенного кодирования информации (это может быть алгоритм шифрования, код аутентификации, либо другое преобразование) и алгоритм принятия решения об истинности полученной информации, а также ключевую систему.

Система идентификации – криптографическая система, выполняющая функцию аутентификации сторон в процессе информационного взаимодействия. Математическая модель системы идентификации включает протокол идентификации и ключевую систему.

Система цифровой подписи – криптографическая система, выполняющая функцию аутентификации источника сообщения или документа и предназначенная для защиты от отказа субъектов от некоторых из ранее совершенных ими действий. Например, отправитель может отказаться от факта передачи сообщения, утверждая, что его создал сам получатель, а получатель легко может модифицировать, подменить или создать новое сообщение, а затем

утверждать, что оно получено от отправителя. Математическая модель системы цифровой подписи включает схему цифровой подписи и ключевую систему.

Система ключевая – определяет порядок использования криптографической системы и включает системы установки и управления ключами.



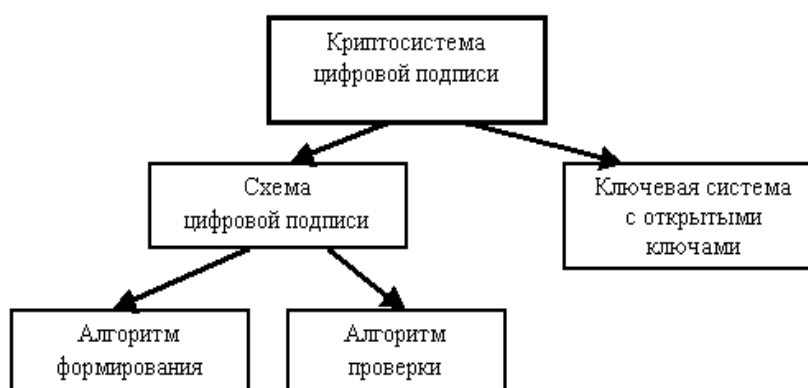
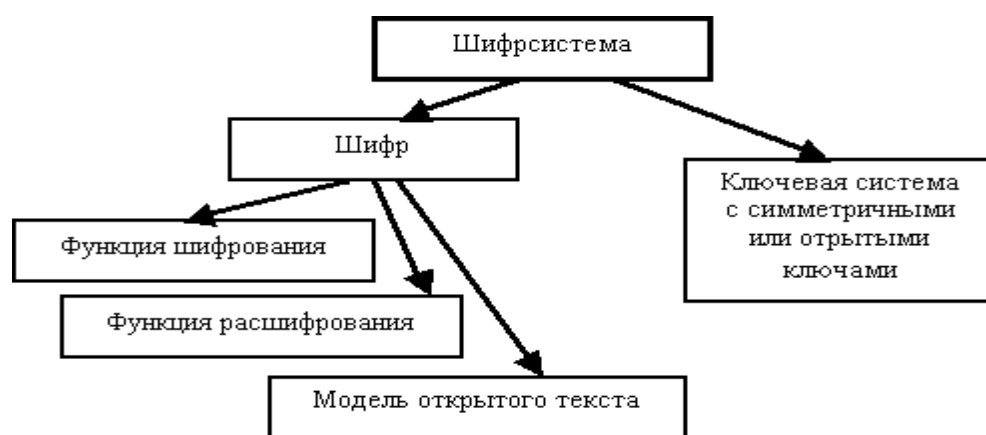
Система установки ключей – определяет алгоритмы и процедуры генерации, распределения, передачи и проверки ключей.

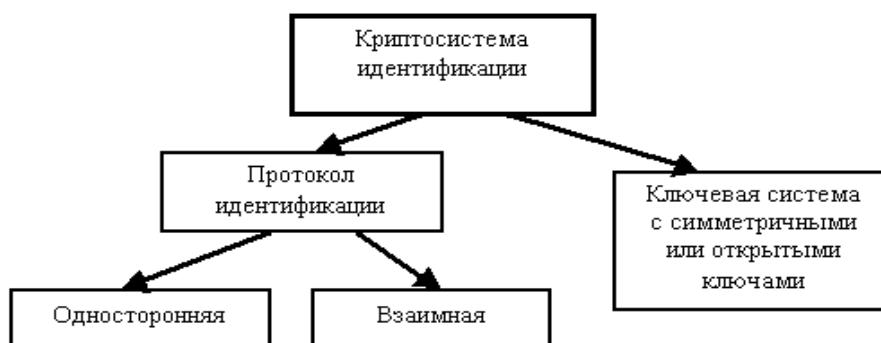
Система управления ключами – определяет порядок использования, смены, хранения и архивирования, резервного копирования и восстановления, замены или изъятия из обращения скомпрометированных, а также уничтожения старых ключей. Целью управления ключами является нейтрализация таких угроз, как: компрометация конфиденциальности секретных ключей, компрометация аутентичности секретных или открытых ключей, несанкционированное использование секретных или открытых ключей, например использование ключа, срок действия которого истек.

Система ключевая симметричной криптосистемы – основана на использовании симметричных (секретных) ключей. Основными проблемами таких систем являются построение системы установки ключей и обеспечение их сохранности для сетей с большим числом абонентов.

Система ключевая асимметричной криптосистемы – основана на использовании асимметричных ключей, состоящих из пары – открытого и секретного (закрытого) ключей. Основными проблемами таких систем являются построение *системы управления ключами*, как правило, представляющей собой инфраструктуру управления сертификатами открытых ключей, включающую центры регистрации и сертификации. Функции обоих центров могут объединяться одним удостоверяющим центром.

Стойкость криптографическая – свойство криптографической системы, характеризующее ее способность противостоять атакам противника, как правило, с целью получить ключ, открытое сообщение или навязать ложное сообщение.





Элементы криптосистем

Алгоритм имитозащищающего кодирования информации – алгоритм преобразования информации (как правило, основан на внесении и использовании избыточности) с целью контроля *целостности*. В отличие от алгоритма формирования цифровой подписи, использует симметричные криптографические системы. В качестве такого преобразования может выступать *код аутентификации*, автоматное и другие преобразования, либо *алгоритм шифрования*.

Алгоритм проверки цифровой подписи – алгоритм, в качестве исходных данных которого используются подписанное сообщение, ключ проверки и параметры *схемы цифровой подписи*, а результатом является заключение о правильности или ошибочности *цифровой подписи*.

Алгоритм расшифрования – алгоритм, реализующий *функцию расшифрования*.

Алгоритм формирования цифровой подписи – алгоритм, в качестве исходных данных которого используются сообщение, ключ подписи и параметры *схемы цифровой подписи*, а в результате формируется *цифровая подпись*.

Алгоритм шифрования – алгоритм, реализующий *функцию шифрования*¹.

Жизненный цикл ключей – последовательность стадий, которые проходят ключи от момента генерации до уничтожения. Включает такие стадии, как: генерация ключей, регистрация пользователей и ключей, инициализация ключей, период действия, хранение ключа, замена ключа, архивирование, уничтожение ключей, восстановление ключей, отмена ключей.

Имитовставка – проверочная комбинация, добавляемая к сообщению для проверки целостности.

Имитостойкость – способность противостоять активным атакам со стороны противника, целью которых является навязывание ложного или подмена передаваемого сообщения или хранимых данных.

Код аутентификации – алгоритм имитозащитающего кодирования информации (как правило, вычисляет значение имитовставки). К кодам аутентификации предъявляются требования: большая сложность вычисления значения кода аутентификации для заданного сообщения без знания ключа; большая сложность подбора для заданного сообщения с известным значением кода аутентификации другого сообщения с известным значением кода аутентификации без знания ключа. Без знания секретного ключа вероятность успешного навязывания противником искаженной или ложной информации мала.

Открытое распределение ключей (согласование ключа, выработка общего значения ключа) – протокол, позволяющий двум абонентам выработать общий секретный ключ путем обмена сообщениями по открытому каналу связи без передачи какой-либо общей секретной информации, распределяемой заранее. Важным преимуществом открытого распределения является то, что ни один из абонентов заранее не может определить значение ключа, так как ключ зависит от сообщений, передаваемых в процессе обмена.

Помехоустойчивость – способность сохранять устойчивую работу при наличии помех в канале связи.

Протокол – распределенный алгоритм, в котором участвуют две или более стороны, обменивающиеся между собой сообщениями.

Протокол идентификации – протокол аутентификации сторон, участвующих во взаимодействии и не доверяющих друг другу. Различают протоколы односторонней и взаимной идентификации. Протоколы идентификации, как правило, основаны на известной обеим сторонам информации (пароли, личные идентификационные номера (PIN), ключи). В дополнение к протоколу идентификации могут использоваться некоторые физические приборы, с помощью которых и проводится идентификация (магнитная или интеллектуальная пластиковая карта, или прибор, генерирующий меняющиеся со временем пароли), а также физические параметры, составляющие неотъемлемую принадлежность доказывающего (подписи, отпечатки пальцев, характеристики голоса, геометрия руки и т. д.).

Протокол криптографический – протокол, предназначенный для выполнения функций криптографической системы, в процессе выполнения которого стороны используют криптографические алгоритмы.

Протокол распределения ключей – протокол, в результате выполнения которого взаимодействующие стороны (участники, группы участников) получают необходимые для функционирования криптографической системы ключи. Различают следующие типы протоколов распределения ключей: протоколы передачи (уже сгенерированных) ключей; протоколы (совместной) выработки общего ключа (открытое распределение ключей); схемы предварительного распределения ключей. В зависимости от порядка взаимодействия сторон

выделяют *двусторонние протоколы*, в которых стороны осуществляют передачу ключей при непосредственном взаимодействии, или, иначе, протоколы типа "точка-точка", и *протоколы с централизованным распределением ключей*, предусматривающие наличие третьей стороны, играющей роль доверенного центра.

Схема цифровой подписи состоит из двух алгоритмов, один – для формирования, а второй – для проверки подписи. Надежность схемы цифровой подписи определяется сложностью следующих трех задач для лица, не являющегося владельцем секретного ключа: *подделки подписи*, то есть вычисления значения подписи под заданным документом; *создания подписанного сообщения*, то есть нахождения хотя бы одного сообщения с правильным значением подписи; *подмены сообщения*, то есть подбора двух различных сообщений с одинаковыми значениями подписи.

Схема предварительного распределения ключей – состоит из двух алгоритмов: распределения исходной ключевой информации и формирования ключа. С помощью первого алгоритма осуществляется генерация исходной ключевой информации. Эта информация включает открытую часть, которая будет передана всем сторонам или помещена на общедоступном сервере, а также секретные части каждой стороны. Вторым алгоритмом предназначен для вычисления действующего значения ключа для взаимодействия между абонентами по имеющейся у них секретной и общей открытой части исходной ключевой информации. Применяется для уменьшения объема хранимой и распределяемой секретной ключевой информации. Схема предварительного распределения ключей должна быть устойчивой, то есть учитывать возможность раскрытия части ключей при компрометации, обмане или сговоре абонентов, и гибкой – допускать возможность быстрого восстановления путем исключения скомпрометированных и подключения новых абонентов.

Функция криптографическая – функция, необходимая для реализации *криптографической системы*, например, генерация ключей и псевдослучайных последовательностей, обратимое преобразование, однонаправленная функция, вычисление и проверка значений имитовставки и цифровой подписи, вычисление значения хэш-функции и т. п., обладают определенными криптографическими свойствами, влияющими на криптографическую стойкость: зависимость от ключа, сложность обращения и др.

Функция расшифрования – осуществляет преобразование множества открытых сообщений в множество зашифрованных сообщений, зависящее от ключа, является обратным к преобразованию, осуществляемому *функцией шифрования*.

Функция шифрования – осуществляет преобразование множества открытых сообщений в множество зашифрованных сообщений, зависящее от ключа.

Цифровая подпись (сообщения или электронного документа) – представляет собой конечную цифровую последовательность, зависящую от самого сообщения или документа и от секретного ключа, известного только подписывающему субъекту, предназначенная для установления авторства. Предполагается, что цифровая подпись должна быть легко проверяемой без получения доступа к секретному ключу. При возникновении спорной ситуации, связанной с отказом подписывающего от факта подписи некоторого сообщения либо с попыткой подделки подписи, третья сторона должна иметь возможность разрешить спор. Цифровая подпись позволяет решить следующие три задачи: осуществить аутентификацию источника данных, установить целостность сообщения или электронного документа, обеспечить невозможность отказа от факта подписи конкретного сообщения.

Шифр – семейство обратимых преобразований множества открытых сообщений в множество зашифрованных сообщений и обратно, каждое из которых определяется некоторым параметром, называемым ключом. Математическая модель шифра включает две функции: шифрования и расшифрования, и модель множества открытых сообщений. В зависимости от способа представления открытых сообщений различают блочные, поточные и другие шифры. Основными требованиями, определяющими качество шифра, являются: криптографическая стойкость, имитостойкость, помехоустойчивость и др. [25-29].

История развития криптографии

История криптографии насчитывает не одно тысячелетие. Уже в исторических документах древних цивилизаций – Индии, Египте, Китае, Месопотамии – имеются сведения о системах и способах составления зашифрованного письма. Видимо, первые системы шифрования появились одновременно с письменностью в четвертом тысячелетии до нашей эры.

В древнеиндийских рукописях приводится более шестидесяти способов письма, среди которых есть и такие, которые можно рассматривать как криптографические. Имеется описание системы замены гласных букв согласными и наоборот. Один из сохранившихся зашифрованных текстов Месопотамии представляет собой табличку, написанную клинописью и содержащую рецепт изготовления глазури для гончарных изделий. В этом тексте использовались редко употребляемые значки, игнорировались некоторые буквы, употреблялись цифры вместо имен. В рукописях Древнего Египта зашифровались религиозные тексты и медицинские рецепты. Шифрование использовалось в Библии. Некоторые фрагменты библейских текстов зашифрованы с помощью шифра, который назывался *атбаиш*. Правило зашифрования состояло в замене i -й буквы алфавита ($i = \overline{1, n}$) буквой с номером $n -$

$i + 1$, где n – число букв алфавита. Происхождение слова *атбаиш* объясняется принципом замены букв. Это слово составлено из букв Алеф, Тае, Бет и Шин, то есть первой и последней, второй и предпоследней букв древнесемитского алфавита.

Развитию криптографии способствовал переход от идеографического письма, основанного на использовании огромного числа иероглифов, к фонетическому письму. В древнем семитском алфавите во втором тысячелетии до нашей эры было уже 30 знаков. Ими обозначались согласные звуки, а также некоторые гласные и слоги. Упрощение письма стимулировало развитие криптографии.

В Древней Греции криптография уже широко использовалась в разных областях деятельности, в особенности в государственной сфере. Плутарх сообщает, что жрецы, например, хранили в форме тайнописи свои прорицания. В Спарте в V – IV веках до н. э. использовалось одно из первых шифровальных приспособлений – *Сцитала*. Это был жезл цилиндрической формы, на который наматывалась лента пергамента. Кроме жезла могли использоваться рукоятки мечей, кинжалов копий и т.д. Вдоль оси цилиндра на пергамент построчно записывался текст, предназначенный для передачи. После записи текста лента сматывалась с жезла и передавалась адресату, который имел точно такую же Сциталу. Ясно, что такой способ шифрования осуществлял перестановку букв сообщения. Ключом шифра служит диаметр Сциталы. Известен также и метод вскрытия такого шифра, приписываемый Аристотелю. Предлагалось заточить на конус длинный брус и, обернув вокруг него ленту, начать сдвигать ее по конусу от малого диаметра до самого большого. В том месте, где диаметр конуса совпадал с диаметром Сциталы, буквы текста сочетались в слоги и слова. После этого оставалось лишь изготовить цилиндр нужного диаметра.

Другим шифровальным приспособлением времен Спарты была *табличка Энея*. На небольшой табличке горизонтально располагался алфавит, а по ее боковым сторонам имелись выемки для наматывания нити. При зашифровании нить закреплялась у одной из сторон таблички и наматывалась на нее. На нити делались отметки (например, узелки) в местах, которые находились напротив букв данного текста. По алфавиту можно было двигаться лишь в одну сторону, то есть делать по одной отметке на каждом витке. После зашифрования нить сматывалась и передавалась адресату. Этот шифр представляет собой шифр замены букв открытого текста знаками, которые означали расстояния между отметками на нити. Ключом являлись геометрические размеры таблички и порядок расположения букв алфавита. Это был довольно надежный шифр, история не сохранила документов, подтверждающих сведения о методах его вскрытия.

Греческий писатель Полибий использовал систему сигнализации, которая была широко принята как метод шифрования. Он записывал буквы алфавита в квадратную таблицу и заменял их координатами: парами чисел (i, j) , где i – номер строки, j – номер столбца. Применительно к латинскому алфавиту *квадрат Полибия* имеет следующий вид.

Таблица 1.1. Квадрат Полибия

	1	2	3	4	5
1	A	B	C	D	E
2	F	G	H	I, J	K
3	L	M	N	O	P
4	Q	R	S	T	U
5	V	W	X	Y	Z

Пары (i, j) передавались с помощью факелов. Например, для передачи буквы O нужно было взять 3 факела в правую руку и 4 факела – в левую.

Подобные шифровальные приспособления с небольшими изменениями просуществовали до эпохи военных походов Юлия Цезаря. Положение меняется в эпоху расцвета Рима, который первоначально представлял собой лишь небольшую гражданскую общину, со временем разросся, подчинив себе сначала Италию, а затем и все Средиземноморье. Чтобы управлять наместниками в многочисленных провинциях, шифрованная связь для римских органов власти стала жизненно необходимой. Особую роль в сохранении тайны сыграл способ шифрования, предложенный Юлием Цезарем и изложенный им в "Записках о галльской войне" (I в. до н. э.). Вот что пишет о нем Гай Светоний: "...существуют и его письма к Цицерону и письма к близким о домашних делах: в них, если нужно было сообщить что-нибудь негласно, он пользовался тайнописью, т. е. менял буквы так, чтобы из них не складывалось ни одного слова. Чтобы разобрать и прочитать их, нужно читать всякий раз четвертую букву вместо первой, например. D вместо A и так далее". Таким образом, Цезарь заменял буквы в соответствии с подстановкой, нижняя строка которой представляет собой алфавит открытого текста, сдвинутый циклически на три буквы влево.

Со времен Цезаря до XV в. шифровальное дело претерпело много изменений, однако нам мало известно о методах и системах шифрования, применяемых в этот период времени. В мрачные годы средневековья практика шифрования сохранялась в строжайшей тайне. Так, в годы крестовых походов шифровальщики, служившие у Папы Римского, после года работы подлежали физическому уничтожению.

В эпоху Возрождения в итальянских городах-государствах параллельно с расцветом культуры и науки активно развивается криптография. Нередко ученые зашифровывали на-

учные гипотезы, чтобы не прослыть еретиком и не подвергнуться преследованиям инквизиции.

Научные методы в криптографии впервые появились, по-видимому, в арабских странах. Арабского происхождения и само слово *шифр*. О тайнописи и ее значении говорится даже в сказках "Тысячи и одной ночи". Первая книга, специально посвященная описанию некоторых шифров, появилась в 855 г., она называлась "Книга о большом стремлении человека разгадать загадки древней письменности". В 1412 г. издается 14-томная энциклопедия, содержащая систематический обзор всех важнейших областей человеческого знания, – "Шауба аль-Аща". Ее автор – Шехаб аль-Кашканди. В этой энциклопедии есть раздел о криптографии под заголовком "относительно сокрытия в буквах тайных сообщений", в котором приводятся семь способов шифрования. Там же дается перечень букв в порядке частоты их употребления в арабском языке на основе изучения текста Корана, а также приводятся примеры раскрытия шифров *методом частотного анализа* встречаемости букв.

В XIV в. появилась книга о системах тайнописи, написанная сотрудником тайной канцелярии Папы Римского Чикко Симонетти. В этой книге приводятся шифры замены, в которых гласным буквам соответствует несколько значковых выражений. Такие шифры позже стали называться шифрами *многозначной замены* или *омофонами*. Они получили развитие в XV в. Так, в книге "Трактат о шифрах" Габриеля де Лавинды – секретаря папы Климентия XII – приводится описание шифра *пропорциональной замены*, в котором каждой букве ставится в соответствие несколько эквивалентов, число которых пропорционально частоте встречаемости буквы в открытом тексте. В 1469 г. был предложен подобный же шифр, получивший название "миланский ключ". Появление омофонов свидетельствовало о том, что к тому времени уже хорошо осознавали слабости шифров простой замены. Такая модификация шифра разрушала статистику букв открытого сообщения, что явилось заметным шагом в развитии криптографии.

Еще один значительный шаг вперед криптография сделала благодаря труду Леона Альберти. Известный философ, живописец, архитектор, он в 1466 г. написал труд о шифрах. В этой работе был предложен шифр, основанный на использовании *шифровального диска*. Сам Альберти назвал его шифром, "достойным королей".

Шифровальный диск представлял собой пару соосных дисков разного диаметра (рисунок 1). Большой из них – неподвижный, его окружность разделена на 24 равных сектора, в которые вписаны 20 букв латинского алфавита в их естественном порядке и 4 цифры (от 1 до 4). При этом из 24-буквенного алфавита были удалены 4 буквы, без которых можно было обойтись, подобно тому, как в русском языке обходятся без Ъ, Ё, Й. Меньший диск –

подвижный, по его окружности, разбитой также на 24 сектора, были вписаны все буквы смешанного латинского алфавита.



Рис. 1.1. Диск Альберти

Имея два таких прибора, корреспонденты договаривались о первой индексной букве на подвижном диске. При шифровании сообщения отправитель ставил индексную букву против любой буквы большего диска. Он информировал корреспондента о таком положении диска, записывая эту букву внешнего диска в качестве первой буквы шифртекста. Очередная буква открытого текста отыскивалась на неподвижном Диске и стоящая против нее буква меньшего диска являлась результатом ее зашифрования. После того как были зашифрованы несколько букв текста, положение индексной буквы изменялось, о чем также каким-либо образом передавалось корреспонденту.

Такой шифр имел две особенности, которые делают изобретение Альберти событием в истории криптографии. Во-первых, в отличие от шифров простой замены шифровальный диск использовал не один, а несколько алфавитов для зашифрования. Такие шифры получили название *многоалфавитных*. Во-вторых, шифровальный диск позволял использовать так называемые *коды с перешифрованием*, которые получили широкое распространение лишь в конце XIX века, то есть спустя четыре столетия после изобретения Альберта. Для этой цели на внешнем диске имелись цифры. Альберта составил код, состоящий из 336 кодовых групп, занумерованных от 11 до 4444. Каждому кодовому обозначению соответствовала некоторая законченная фраза. Когда такая фраза встречалась в открытом сообщении, она заменялась соответствующим кодовым обозначением, а с помощью диска цифры зашифровывались как обычные знаки открытого текста, превращаясь в буквы.

Богатым на новые идеи в криптографии оказался XVI век. Многоалфавитные шифры получили развитие в вышедшей в 1518 г. первой печатной книге по криптографии под названием "Полиграфия". Автором книги был один из самых знаменитых ученых того времени аббат Иоганнес Тритемий. В этой книге впервые в криптографии появляется квадратная таблица. Шифралфавиты записаны в строки таблицы один под другим, причем каждый из них сдвинут на одну позицию влево по сравнению с предыдущим.

Тритемий предлагал использовать эту таблицу для многоалфавитного зашифрования самым простым из возможных способов: первая буква текста шифруется первым алфавитом, вторая буква – вторым и т. д. В этой таблице не было отдельного алфавита открытого текста, для этой цели служил алфавит первой строки. Таким образом, открытый текст, начинающийся со слов HUNC CAVETO VIRUM..., приобретал вид HXPF GFBMCZ FUEIB....

Преимущество этого метода шифрования по сравнению с методом Альберти состоит в том, что с каждой буквой задействуется новый алфавит. Альберти менял алфавиты лишь после трех или четырех слов. Поэтому его шифртекст состоял из отрезков, каждый из которых обладал закономерностями открытого текста, которые помогали вскрыть криптограмму. Побуквенное зашифрование не дает такого преимущества. *Шифр Тритемия* является также первым нетривиальным примером *периодического шифра*. Так называется многоалфавитный шифр, правило зашифрования которого состоит в использовании периодически повторяющейся последовательности простых замен.

В 1553 г. Джованни Баттиста Белазо предложил использовать для многоалфавитного шифра буквенный, легко запоминаемый ключ, который он назвал *паролем*. Паролем могло служить слово или фраза. Пароль периодически записывался над открытым текстом. Буква пароля, расположенная над буквой текста, указывала на алфавит таблицы, который использовался для зашифрования этой буквы. Например, это мог быть алфавит из таблицы Тритемия, первой буквой которого являлась буква пароля. Однако Белазо, как и Тритемий, использовал в качестве шифралфавитов обычные алфавиты.

Воскресить смешанные алфавиты, которые применял Альберти, и объединить идеи Альберти с идеями Тритемия и Белазо в современную концепцию многоалфавитной замены выпало на долю итальянца Джованни де ла Порта. Ему было 28 лет, когда он в 1563 г. опубликовал книгу "О тайной переписке". По сути, эта книга являлась учебником по криптографии, содержащим криптографические познания того времени. Порта предложил использовать квадратную таблицу с периодически сдвигаемым смешанным алфавитом и паролем. Он советовал выбирать длинный ключ. Впервые им был предложен *шифр простой биграммной замены*, в котором пары букв представлялись одним специальным графическим

символом. Они заполняли квадратную таблицу размеров 20x20, строки и столбцы которой занумерованы буквами алфавита

A B C D E F G H I L M N O P Q R S T U Z

Например, биграмма EA заменялась символом Δ, биграмма LF – символом ⊕ и т. д. В своей книге Порты ввел многоалфавитный шифр, определяемый таблицей 6.2.

Таблица 6.2. Таблица Порты

A	a	b	c	d	e	f	g	h	i	k	l	m
B	n	o	p	q	r	s	t	u	x	y	z	w
C	a	b	c	d	e	f	g	h	i	k	l	m
D	o	p	q	r	s	t	u	x	y	z	w	n
E	a	b	c	d	e	f	g	h	i	k	l	m
F	p	q	r	s	t	u	x	y	z	w	n	o
G	a	b	c	d	e	f	g	h	i	k	l	m
H	q	r	s	t	u	x	y	z	w	n	o	p
I	a	b	c	d	e	f	g	h	i	k	l	m
K	r	s	t	u	x	y	z	w	n	o	p	q
L	a	b	c	d	e	f	g	h	i	k	l	m
M	s	t	u	x	y	z	w	n	o	p	q	r
N	a	b	c	d	e	f	g	h	i	k	l	m
O	t	u	x	y	z	w	n	o	p	q	r	s
P	a	b	c	d	e	f	g	h	i	k	l	m
Q	u	x	y	z	w	n	o	p	q	r	s	t
R	a	b	c	d	e	f	g	h	i	k	l	m
S	x	y	z	w	n	o	p	q	r	s	t	u
T	a	b	c	d	e	f	g	h	i	k	l	m
U	y	z	w	n	o	p	q	r	s	t	u	x
X	a	b	c	d	e	f	g	h	i	k	l	m
Y	z	w	n	o	p	q	r	s	t	u	x	y
Z	a	b	c	d	e	f	g	h	i	k	l	m
W	w	n	o	p	q	r	s	t	u	x	y	z

Шифрование осуществляется при помощи лозунга, который пишется над открытым текстом. Буква лозунга определяет алфавит (заглавные буквы первого столбца), расположенная под ней буква открытого текста ищется в верхнем или нижнем полуалфавите и заменяется соответствующей ей буквой второго полуалфавита. Например, фраза, начинающаяся словами HUNC CAVETO VIRUM..., будет зашифрована при помощи лозунга DE LA PORTA в XFHP YTMOGA FQEAS.

Еще одно важное усовершенствование многоалфавитных систем, состоящее в идее использования в качестве ключа текста самого сообщения или же зашифрованного текста, принадлежит Джероламо Кардано и Блезу де Виженеру. Такой шифр был назван *самоключом*. В книге Виженера "Трактат о шифрах" самоключ представлен следующим образом. В простейшем случае за основу бралась таблица Тритемия с добавленными к ней в качестве

первой строки и первого столбца алфавитами в их естественном порядке. Позже такая таблица стала называться *таблицей Виженера*. Подчеркнем, что в общем случае таблица Виженера состоит из циклически сдвигаемых алфавитов, причем первая строка может быть произвольным смешанным алфавитом (см. табл. 6..3).

Таблица 6.3. Таблица Виженера

	A	B	C	D	E	F	G	H	I	K	L	M	N	O	P	Q	R	S	T	U	X	Y	Z	W
A	A	B	C	D	E	F	G	H	I	K	L	M	N	O	P	Q	R	S	T	U	X	Y	Z	W
B	B	C	D	E	F	G	H	I	K	L	M	N	O	P	Q	R	S	T	U	X	Y	Z	W	A
C	C	D	E	F	G	H	I	K	L	M	N	O	P	Q	R	S	T	U	X	Y	Z	W	A	B
D	D	E	F	G	H	I	K	L	M	N	O	P	Q	R	S	T	U	X	Y	Z	W	A	B	C
E	E	F	G	H	I	K	L	M	N	O	P	Q	R	S	T	U	X	Y	Z	W	A	B	C	D
F	F	G	H	I	K	L	M	N	O	P	Q	R	S	T	U	X	Y	Z	W	A	B	C	D	E
G	G	H	I	K	L	M	N	O	P	Q	R	S	T	U	X	Y	Z	W	A	B	C	D	E	F
H	H	I	K	L	M	N	O	P	Q	R	S	T	U	X	Y	Z	W	A	B	C	D	E	F	G
I	I	K	L	M	N	O	P	Q	R	S	T	U	X	Y	Z	W	A	B	C	D	E	F	G	H
K	K	L	M	N	O	P	Q	R	S	T	U	X	Y	Z	W	A	B	C	D	E	F	G	H	I
L	L	M	N	O	P	Q	R	S	T	U	X	Y	Z	W	A	B	C	D	E	F	G	H	I	K
M	M	N	O	P	Q	R	S	T	U	X	Y	Z	W	A	B	C	D	E	F	G	H	I	K	L
N	N	O	P	Q	R	S	T	U	X	Y	Z	W	A	B	C	D	E	F	G	H	I	K	L	M
O	O	P	Q	R	S	T	U	X	Y	Z	W	A	B	C	D	E	F	G	H	I	K	L	M	N
P	P	Q	R	S	T	U	X	Y	Z	W	A	B	C	D	E	F	G	H	I	K	L	M	N	O
Q	Q	R	S	T	U	X	Y	Z	W	A	B	C	D	E	F	G	H	I	K	L	M	N	O	P
R	R	S	T	U	X	Y	Z	W	A	B	C	D	E	F	G	H	I	K	L	M	N	O	P	Q
S	S	T	U	X	Y	Z	W	A	B	C	D	E	F	G	H	I	K	L	M	N	O	P	Q	R
T	T	U	X	Y	Z	W	A	B	C	D	E	F	G	H	I	K	L	M	N	O	P	Q	R	S
U	U	X	Y	Z	W	A	B	C	D	E	F	G	H	I	K	L	M	N	O	P	Q	R	S	T
X	X	Y	Z	W	A	B	C	D	E	F	G	H	I	K	L	M	N	O	P	Q	R	S	T	U
Y	Y	Z	W	A	B	C	D	E	F	G	H	I	K	L	M	N	O	P	Q	R	S	T	U	X
Z	Z	W	A	B	C	D	E	F	G	H	I	K	L	M	N	O	P	Q	R	S	T	U	X	Y
W	W	A	B	C	D	E	F	G	H	I	K	L	M	N	O	P	Q	R	S	T	U	X	Y	Z

Первая строка служит алфавитом открытого текста, а первый столбец – алфавитом ключа. Для зашифрования открытого сообщения ($T_o = t_1t_2\dots$) Виженер предлагал в качестве *ключевой последовательности* (Γ) использовать само сообщение (T_o) с добавленной к нему в качестве первой буквы (t_0), известной отправителю и получателю (этим идея Виженера отличалась от идеи Кардано, у которого не было начальной буквы и система которого не обеспечивала однозначности расшифрования). Последовательности букв подписывались друг под другом:

$$\Gamma = t_0t_1t_2\dots t_{i-1}\dots$$

$$T_o = t_1t_2t_3\dots t_i\dots$$

$$T_u = s_1s_2s_3... s_i...$$

При этом пара букв, стоящих друг под другом в Γ и T_o , указывала, соответственно, номера строк и столбцов таблицы, на пересечении которых находится знак s_i шифрованного текста (T_u). Например, фраза HUNC CAVETO VIRUM..., использованная в предыдущих примерах, и начальная буква P дают шифртекст YCHP ECUWZH IDAMG.

Во втором варианте Виженер предлагал в качестве ключевой последовательности использовать шифрованный текст:

$$\Gamma = s_0s_1s_2... s_{i-1}...$$

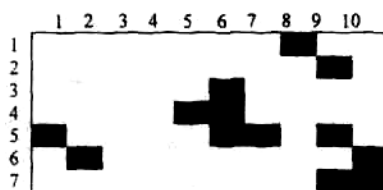
$$T_o = t_1t_2t_3... t_i...$$

$$T_u = s_1s_2s_3... s_i...$$

Самоключ Виженера был незаслуженно забыт на долгое время, а под *шифром Виженера* до сих пор понимают самый простой вариант с коротким ключевым словом и с таблицей, состоящей из обычных алфавитов.

Кардано принадлежит также идея *поворотной решетки* как средства шифрования. Изначально обычная решетка представляла собой лист из твердого материала, в котором через неправильные интервалы сделаны прямоугольные вырезы высотой для одной строчки и различной длины. Накладывая эту решетку на лист писчей бумаги, можно было записывать в вырезы секретное сообщение. После этого, сняв решетку, нужно было заполнить оставшиеся свободные места на листе бумаги неким текстом, маскирующим секретное сообщение. Подобным стенографическим методом маскировки сообщения пользовались многие известные исторические лица, например кардинал Ришелье во Франции и русский дипломат и писатель А. Грибоедов. Так, Ришелье использовал прямоугольник размера 7×10. Для длинных сообщений прямоугольник использовался несколько раз. Прорези трафарета размещались в позициях:

(1,8), (2,9), (3,6), (4,5), (4,6), (5,1), (5,6), (5,7), (5,9), (6,2), (6,10), (7,9), (7,10).



Следующий текст выглядит как невинное любовное письмо:

	1	2	3	4	5	6	7	8	9	10	
1	I		L	O	V	E		Y	O	U	
2	I		H	A	V	E		Y	O	U	
3	D	E	E	P		U	N	D	E	R	
4	M	Y		S	K	I	N		M	Y	
5	L	O	V	E		L	A	S	T	S	
6	F	O		R	E	V	E	R		I	N
7	H	Y		P	E	R	S	P	A	C	E

Однако, используя трафарет Ришелье, получим зловещую команду:

YOU KILL AT ONES

Кардано использовал квадратную решетку, которая своими вырезами однократно покрывает всю площадь квадрата при ее самосовмещениях. На основе такой решетки он построил шифр перестановки.

Нельзя не упомянуть в историческом обзоре имени Матео Ардженти, работавшего в области криптографии в начале XVII в. Он составил руководство по криптографии на 135 листах, изданное в переплете из телячьей кожи. В этой книге впервые предложено использовать некоторое слово в качестве мнемонического ключа для смешанного алфавита. Началом смешанного алфавита служило ключевое слово (как правило, без повторяющихся букв), за которым следовали остальные буквы в их естественном порядке. Например, ключевое слово PIETRO дает смешанный латинский алфавит

PIETRO ABCDFGHLMNQS UZ

Такие смешанные алфавиты часто использовались в качестве алфавитов шифртекста в шифрах простой замены.

С целью усложнения шифра простой замены Ардженти вводил *пустышки*, которые добавлялись в зашифрованное сообщение, использовал шифробозначения разной значности. Для некоторых частых сочетаний букв текста вводил отдельные обозначения, придавал частым буквам несколько обозначений. Позже подобные идеи получили широкое распространение. Приведем пример *шифра Ардженти* (см. табл. 6.4).

Таблица 6.4. Шифр Ардженти

	A	B	C	D	E	F	G	H	I	L	M	N	O	P
1	86	02	20	62	22	06	60	3	24	26	84	9	66	

Q	R	S	T	U	Z	ET	CON	NON	CHE	∅
68	28	42	80	04	88	08	64	00	44	5
				40						7

Слово ARGENTI может быть зашифровано многими способами, например так:

5128068285480377

или же так:

172850675628455803

Наибольшим достижением Ардженти считается разработанный им *буквенный код* – один из шифров замены, в котором буквы, слоги, слова и целые фразы заменялись группами букв. Необходимым количеством словарных величин в коде в то время считалось 1200.

В истории криптографии XVII – XVIII века называют эрой "*черных кабинетов*". В этот период во многих государствах Европы, в первую очередь во Франции, получили развитие дешифровальные подразделения, названные "черными кабинетами". Первый из них образован по инициативе кардинала Ришелье при дворе короля Людовика XIII. Его возглавил первый профессиональный криптограф Франции Антуан Россиньоль. Следует отметить, что некоторые оригинальные идеи, возникшие в криптографии в этот период, связаны с именем самого Ришелье, который использовал, например, для секретной переписки с королем оригинальный шифр перестановки с переменным ключом. Его использование становится понятным из следующего примера:

Шифр Ришелье

Ключ: 2741635; 15243; 671852493; 07; 28615; 943; ...

Открытый текст: LETTER SENT TO THE EMPEROR GIVING FULL DETAIL

Ключ: (2741635) (15243) (671852493) (07) (28615) (943) (27 41635)

Шифртекст: TLRTSEE ETOTN EPOEMTHER NI LUGIG VFR TLIE SAD

Известно, что Ришелье пользовался также кодами. Попутно отметим, что свой несложный код был и у знаменитого Наполеона:

Малый шифр Наполеона (Petit Chiffre)

(реконструирован Базери)

A – 15, ar – 25, al – 39

B – 37, bu – 3, bo – 35, bi – 29

C – 6, ca – 32, ce – 20

D – 23, de – 52

E – 53, es – 82, et – 50, en – 68

F – 55, fa – 69, fe – 58, fo – 71

G – 81, ga – 51

H – 85, hi – 77

I – 119, jai – 122

J – 87, jai – 123

K – ?

L – 96, lu – 103, le – 117, la – 106

M – 114, ma – 107

N – 115, ne – 94, ni – 116

O – 90, ot – 153

P – 137, po – 152

Q – 173, que – 136

R – 169, ra – 146, re – 126, ri – 148,
 S – 167, sa – 171, se – 177, si – 134, so – 168, su – 174
 T – 176, ti – 145, to – 157
 U – 138
 V – 164, ve – 132, vi – 161, vo – 175
 W, X, Y – ?
 Z – 166

В то время в Европе получили широкое распространение шифры, называемые *номенклаторами*, объединявшие в себе простую замену и код. В простейших номенклаторах код состоял из нескольких десятков слов или фраз с двухбуквенными кодовыми обозначениями. Со временем списки заменяемых слов в номенклаторах увеличились до двух или трех тысяч эквивалентов слогов и слов. В царской России XVIII в. закодированное открытое сообщение шифровалось далее простой заменой.

Кстати, несколько слов о русской криптографии. Уже с XIV в. в Новгороде существовала техника тайного письма. Использовались в основном шифры простой замены. Благодаря торговым связям Новгорода с Германией в России становятся известными многие западные разработки, в том числе новые системы шифрования. Учреждение постоянной почтовой связи России с Европой дало возможность развития шифрованной переписке. Благодаря привлечению Петром I для разработки проектов развития образования и государственного устройства России знаменитого Готфрида Вильгельма Лейбница, который известен и как криптограф, в Петербурге появилась цифирная палата, задачами которой было развитие и использование систем шифрования.

Когда Россиньоль начинал свою карьеру, в номенклаторах как элементы открытого текста, так и элементы кода располагались в алфавитном порядке (или в алфавитном и числовом порядке, если код был цифровой). Россиньоль заметил, что такой "параллелизм" открытого текста и кода облегчал восстановление открытого текста. Если, например, он устанавливал, что в английской депеше 137 заменяет FOR, а 168 – IN, то он уже знал, что 21 не может заменять TO, так как цифровые кодовые обозначения для слов, начинающихся с T, должны быть больше, нежели для слов, начинающихся с I. Обнаружив такую слабость, Россиньоль перемешивал кодовые элементы по отношению к открытому тексту. На одном листе он располагал элементы открытого текста в алфавитном порядке, а кодовые элементы – вразброс, на другом листе для облегчения расшифрования кодовые элементы стояли в алфавитном порядке, тогда как их открытые эквиваленты были разбросаны. Это явилось значительным усовершенствованием подобных шифрсистем. Однако составление неалфа-

витных номенклаторов обходилось очень дорого, и, таким образом, по соображениям экономии и в ущерб надежности многие номенклаторы регрессировали к упрощенному алфавитному типу.

В Англии тоже был свой "черный кабинет". В его работе в XVII в. заметное место занимал Джон Валлис, известный как крупнейший английский математик до Исаака Ньютона. Работы по вскрытию шифров для парламента привели к назначению Валлиса в 1649 г. в Оксфорд профессором геометрии в возрасте 32 лет. В своем труде "Арифметика бесконечного" он сделал выводы, которые послужили Ньютону стартовой площадкой для разработки интегрального исчисления. Валлис ввел знак ∞ для бесконечности и первый путем интерполяции вычислил число π . Кстати, само это обозначение также принадлежит ему.

В Германии начальником первого дешифровального отделения был граф Гронсфельд, создавший один из вариантов усовершенствования шифра Виженера. Он взял числовой, легко запоминаемый лозунг. Вместо таблицы Виженера использовался один несмешанный алфавит. При шифровании знаки открытого текста выписывались под цифрами лозунга. Очередная буква открытого текста заменялась буквой алфавита, отстоящей от нее вправо на количество букв, равное соответствующей цифре лозунга.

Шифр Гронсфельда

Открытый текст: GERMANY

Лозунг: 1 3 5 7 9

Алфавит: ABCDEFGHIJKLMNOPQRSTUVWXYZ

Для удобства выпишем алфавит с порядковыми номерами букв:

A	B	C	D	E	F	G	H	I	J	K	L	M
1	2	3	4	5	6	7	8	9	10	11	12	13

N	O	P	Q	R	S	T	U	V	W	X	Y	Z
14	15	16	17	18	19	20	21	22	23	24	25	26

и лозунг над текстом:

1	3	5	7	9	1	3
G	E	R	M	A	N	Y

Теперь легко получить зашифрованный текст: ННWTJOB

Любопытен опыт использования криптографии при составлении *астрономических анаграмм*. Одно из таких применений связано с открытием колец Сатурна [Пер66].

В годы жизни Галилео Галилея существовал обычай закреплять за собой право на первенство в каком-либо открытии своеобразным способом. Напав на открытие, которое нуждается в дальнейшем подтверждении, ученый из опасения, чтобы его не опередили другие, прибегал к помощи *анаграммы* (перестановке букв); он кратко объявлял о сущности своего открытия в форме анаграммы, истинный смысл которой был известен лишь ему одному. Это давало ученому возможность не спеша проверить свое открытие, а в случае появления другого претендента – доказать свое первенство. Когда же он окончательно убеждался в правильности первоначальной догадки, он раскрывал секрет анаграммы. Заметив в свою несовершенную подзорную трубу, что Сатурн имеет по бокам какие-то придатки, Галилей поспешил сделать заявку на это открытие и опубликовал следующий набор букв:

SMAISMRMIELMEPOETALEUMIBUVNEUGTTAVIRAS

Задача восстановления открытого текста (без какой-либо дополнительной информации об использованном преобразовании) требует перебора $\frac{39!}{3!5!4!4!2!2!5!3!3!2!2!}$ возможных перестановок букв криптограммы (это – число перестановок с повторениями). Приведенное число имеет в своей записи примерно 35 цифр.

Современник итальянского ученого Иоганн Кеплер с присущим ему беспримерным терпением затратил немало труда на то, чтобы проникнуть в сокровенный смысл заявки Галилея, и ему казалось, что он добился этого, когда из опубликованных букв (опустив две из них) составил такую латинскую фразу:

SALVE, UMBISTINEUM GEMINAUM MARTIA PROLES

(Привет вам, близнецы, Марса порождение)

Кеплер был убежден, что Галилей открыл те два спутника Марса, существование которых подозревал он сам (они в действительности и были открыты, но спустя два с половиной века). Однако остроумие Кеплера на этот раз не привело к цели. Когда Галилей раскрыл, наконец, секрет своей заявки, оказалось, что фраза (если двумя буквами пренебречь) такова:

ALTISSIMUM PLANETAM TERGEMINUM OBSERVAVI

(Высочайшую планету тройною наблюдал)

Из-за слабости своей трубы Галилей не мог понять истинного значения этого "тройного" образа Сатурна, а когда спустя несколько лет боковые придатки планеты совершенно исчезли", Галилей решил, что ошибся, и никаких придатков У Сатурна нет. Открыть кольца Сатурна удалось только через полвека Гюйгенсу. Подобно Галилею, он не сразу опубликовал свое открытие, а скрыл догадку под тайнописью:

AAAAAAACCCCCDEEEEGHIIIIILLLLMMNNNNNNNNN

OOOOPPQRRSTTTTTUUUUU

Спустя три года, убедившись в правильности своей догадки, Гюйгенс обнародовал смысл заявки:

Annulo cingitur, tenui, piano, nusquam cohaerente,
ad eclipticam inclinato

(Кольцом окружен тонким, плоским, нигде
не прикасающимся к эклиптике)

В целом можно сказать, что XVII и XVIII века не дали новых идей в криптографии. Эра "черных кабинетов" закончилась в 40-х годах XIX в. в период революционного подъема.

Много новых идей в криптографии принес XIX в. Изобретение в середине XIX в. телеграфа и других технических видов связи дало новый толчок развитию криптографии. Информация передавалась в виде токовых и бестоковых посылок, то есть представлялась в двоичном виде. Поэтому возникла проблема "рационального" представления информации, которая решалась с помощью кодов. Коды позволяли передать длинное слово или целую фразу двумя-тремя знаками. Появилась потребность в высокоскоростных способах шифрования и в корректирующих кодах, необходимых в связи с неизбежными ошибками при передаче сообщений.

Однако еще до изобретения телеграфа появился ряд интересных шифровальных устройств. Приблизительно в 1800 г. была создана одна шифровальная система, занимающая особое место в истории криптографии. Речь идет о "дисковом шифре" Т.Джефферсона – первого государственного секретаря США, ставшего позже третьим президентом.

Дисковый шифратор Т.Джефферсона состоял из 25 – 36 деревянных дисков одинакового размера, насаженных на общую ось (рисунок 6.2).

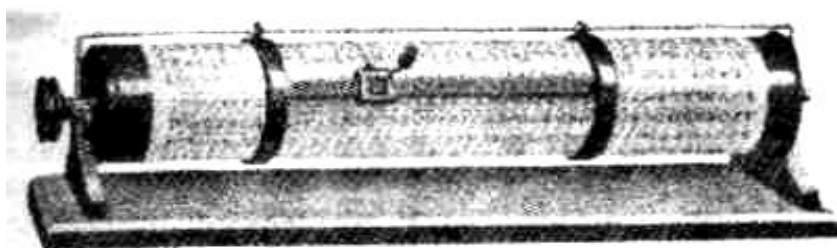


Рис.6.2. Дисковый шифратор Т. Джефферсона

На одном конце оси имелась неподвижная головка, на другом – резьба и гайка, с помощью которой все диски фиксировались в любом нужном угловом положении. Имелась также прямолинейная рейка, способная вращаться на оси и позволяющая выделить строку

букв на дисках, параллельную оси. На боковой поверхности каждого диска, разделенной на 26 равных частей, наносились буквы смешанных английских алфавитов. Для зашифрования части сообщения (длина которой равнялась числу дисков на оси) под рейку, находящуюся в фиксированном угловом положении, подводилась первая буква сообщения, найденная на первом диске, затем – вторая буква сообщения, найденная на втором диске, и т.д., так чтобы все подобранные буквы оказались в одной строке. Положение дисков фиксировалось гайкой, после чего рейка подводилась под любую другую строку цилиндра, буквы которой составляли зашифрованный текст. При расшифровании буквы зашифрованного текста, набранные на последовательных дисках, подводились аналогичным образом под рейку, положение дисков фиксировалось гайкой, после чего с помощью рейки просматривались образовавшиеся строки цилиндра, среди которых несложно было найти открытое сообщение.

Кажущаяся некорректность, связанная с возможностью неоднозначности расшифрования, устраняется достаточно большим числом используемых дисков. Это замечание относится, конечно, лишь к осмысленным текстам. При зашифровании неосмысленных текстов требовалась дополнительная информация о величине сдвига рейки, без чего однозначное расшифрование невозможно.

Такая шифрсистема имеет огромное количество ключевых элементов. К ним относятся: расположение букв алфавита на дисках, расстановка дисков на оси, выбор набора дисков из имеющегося запаса. Дисковый шифр можно отнести по типу к многоалфавитной замене. Его особенностью является поблочный характер зашифрования, при котором каждый участок текста (блок) шифруется независимо от других. Позже такие шифры стали называться *блочными шифрами*.

Вместо того чтобы (пользуясь служебным положением) внедрить свое замечательное изобретение в практику, Джефферсон, по-видимому, отложил его в архив и предал забвению. Шифр был обнаружен в его бумагах в библиотеке конгресса лишь в 1922 г., по иронии судьбы именно в том году, когда в армии США начали применять почти аналогичную систему, изобретенную независимо от Джефферсона.

В 1817г. другой американец Десиус Уодсворт сконструировал шифровальное устройство, которое также внесло новый принцип в криптографию. Его нововведение состояло в том, что он сделал алфавиты открытого и зашифрованного текстов различных длин. Устройство, с помощью которого он это осуществил, представляло собой диск, на котором были расположены два подвижных кольца с алфавитами. Внешний алфавит состоял из 26 букв и 7 цифр (от 2 до 8). Внутренний алфавит состоял лишь из 26 букв. Диск имел подобие неподвижной часовой стрелки, в двух прорезях которой появлялись расположенные друг под другом буквы алфавитов. На внутреннем кольце указывалась буква открытого текста, на

внешнем кольце – соответствующая буква шифртекста. Оба кольца могли вращаться и были связаны друг с другом с помощью двух шестерен, одна из которых имела 33 зубца, а другая – 26. Буквы и цифры внешнего кольца были съемными и могли быть собраны в любом порядке. Перед зашифрованием корреспонденты договаривались относительно взаимного начального положения обоих колец. Для установки дисков в такое положение шестерни можно было разъединить. Проследим на примере слова "введение" процесс зашифрования.

Сначала внутреннее кольцо поворачивалось до тех пор, пока в прорези стрелки не показывалась буква "в". Стоящая в другой прорези буква внешнего кольца записывалась в качестве первой буквы шифртекста. Затем внутреннее кольцо вращалось до тех пор, пока буква "в" вновь не показывалась в прорези. Это вращение посредством шестерен передавалось на внешнее кольцо, но из-за различия в числе букв алфавитов оно совершало лишь $26/33$ полного оборота, в то время как

внутреннее кольцо совершало полный оборот. Значит, второй знак шифртекста располагался во внешнем алфавите на расстоянии семи мест вперед от первого знака, несмотря на то, что оба знака представляли одну и ту же букву открытого текста. Если этот процесс зашифрования осуществлять дальше, то эквиваленты шифртекста для буквы "в" начнут повторяться лишь после того, как будут использованы все 33 буквы и цифры внешнего алфавита. Это объясняется тем, что числа 26 и 33 не имеют общих делителей, благодаря которым такое повторение могло бы произойти раньше. Следующие буквы открытого текста шифровались аналогично.

Такая шифрсистема реализует периодическую многоалфавитную замену. Различие чисел букв алфавитов открытого и шифрованного текстов приводит к существенным отличиям этой системы от предыдущих многоалфавитных систем. Так, в устройстве Уодсворда используется 33 шифралфавита, а не 24 или 26, как в системах Тритемия или Виженера. Важнее то, что эти алфавиты используются не непосредственно один за Другим, а в произвольном порядке, который зависит от букв открытого текста. Этот произвольный порядок служит гораздо более надежной защитой шифра, чем правильная последовательность использования алфавитов, как в системе Тритемия.

Идея Уодсворда была незаслуженно забыта. Славу открытия приписывают английскому ученому Чарлзу Уитстону, который значительно позже и независимо изобрел свое устройство на том же принципе. Основное отличие заключалось в том, что в устройстве Уитстона алфавиты были неподвижными, но зато имелась пара подвижных стрелок, соединенных шестеренками.

Уитстон более известен как ученый, предложивший идею электрического телеграфа, изобретатель концертино, автор первых стереоскопических рисунков. Он высказал гипотезу

о создании говорящих машин, разработал метод точного измерения электрического сопротивления, который называется "мостик Уитстона".

Впервые свое устройство Уитстон продемонстрировал на Всемирной выставке в Париже в 1876 г. На внешнем кольце находился алфавит открытого текста, состоящий из 27 элементов: 26 букв, расположенных в обычном порядке, и знака пробела между словами. Внутренний алфавит состоял из 26 букв, расположенных в произвольном порядке.

Уитстон изобрел шифр, который позже стали называть *шифром Плейфера*. Дело в том, что Лион Плейфер, заместитель председателя Палаты общин, министр почт, председатель Британской ассоциации развития науки, был другом Уитстона, был похож на него, так что их часто путали. В 1854 г. Плейфер продемонстрировал систему шифрования, которую он назвал "недавно открытый симметричный шифр Уитстона". Это был первый из известных биграммных буквенных шифров (напомним, что биграммный шифр Порты был значковым). То обстоятельство, что Плейфер популяризировал изобретение Уитстона, сохранило его имя в названии шифра. Этот шифр использовался англичанами в период I мировой войны.

Во второй половине XIX в. появился весьма устойчивый способ усложнения числовых кодов – *гаммирование*. Он заключался в перешифровании закодированного сообщения с помощью некоторого ключевого числа, которое и называлось *гаммой*. Шифрование с помощью гаммы состояло в сложении всех кодированных групп сообщения с одним и тем же ключевым числом. Эту операцию стали называть "*наложением гаммы*". Например, результатом наложения гаммы 6413 на кодированный текст 3425 7102 8139 являлась числовая последовательность 9838 3515 4552:

$$\begin{array}{r} 3425\ 7102\ 8139 \\ +\ 6413\ 6413\ 6413 \\ \hline 9838\ 3515\ 4552 \end{array}$$

Единицы переноса, появляющиеся при сложении между кодовыми группами, опускались. "*Снятие гаммы*" являлось обратной операцией:

$$\begin{array}{r} 9838\ 3515\ 4552 \\ -\ 6413\ 6413\ 6413 \\ \hline 3425\ 7102\ 8139 \end{array}$$

В 1888 г. француз маркиз де Виари в одной из своих научных статей, посвященных криптографии, обозначил греческой буквой X любую букву шифрованного текста, греческой буквой Г любую букву гаммы и строчной буквой с любую букву открытого текста. Он, по сути, доказал, что алгебраическая формула

$$X = (c + \Gamma) \bmod 26 \quad (1)$$

воспроизводит зашифрование по Виженеру при замене букв алфавита числами согласно следующей таблице:

A	B	C	D	E	F	G	H	I	J	K	L	M
0	1	2	3	4	5	6	7	8	9	10	11	12

N	O	P	Q	R	S	T	U	V	W	X	Y	Z
13	14	15	16	17	18	19	20	21	22	23	24	25

Тем самым была заложена алгебраическая основа для исследования шифров замены типа шифра Виженера. Используя уравнение шифрования, можно было отказаться от громоздкой таблицы Виженера.

Позже лозунговая гамма стала произвольной последовательностью, а шифр с уравнением шифрования (1) стал называться *шифром гаммирования*.

Еще одним известным криптографом того времени был голландец Керкгоффс. Его полным именем было Жан-Вильгельм-Губерт-Виктор-Франсуа-Александр-Огюст Керкгоффс ван Ньювенгоф. Разносторонний ученый, преподававший 6 иностранных языков, историю и математику, он в возрасте 47 лет написал книгу "Военная криптография". В ней сформулированы 6 конкретных требований к шифрам, два из которых относятся к стойкости шифрования, а остальные – к эксплуатационным качествам. Одно из них ("компрометация системы не должна причинять неудобств корреспондентам") стало называться "*правилом Керкгоффса*". Суть этого правила состоит в том, что *стойкость* (или надежность) шифра определяется лишь секретностью ключа. Другими словами, оценка качества шифра (на основе некоторого зашифрованного текста) должна проводиться при условии, что о данном шифре известно все, кроме использованного ключа.

XX век "прославился" двумя мировыми войнами. Эти войны оставили свой отпечаток на всех процессах, происходивших в человеческом обществе. Они не могли не сказаться и на развитии криптографии.

В период первой мировой войны в качестве полевых шифров широко использовались ручные шифры, в первую очередь шифры перестановки с различными усложнениями. Это были *вертикальные перестановки*, усложненные перекодировкой исходного алфавита, а также *двойные вертикальные перестановки*.

Первая мировая война явилась поворотным пунктом в истории криптографии: если до войны криптография представляла собой достаточно узкую область, то после войны она стала широким полем деятельности. Причина этого состояла в необычайном росте объема шифрпереписки, передаваемой по различным каналам связи. Криптоанализ стал важнейшим элементом разведки.

Прогресс этой области криптографии характеризовался и изменениями в самом криптоанализе. Эта наука переросла методы индивидуальной работы криптоаналитика над криптограммой. Системы секретной связи перестали быть настолько малочисленными и однородными, что один специалист мог овладеть всеми специализациями. Характер используемых шифров потребовал для их вскрытия скрупулезного анализа переписки, поиска ситуаций, благоприятствующих успешному криптоанализу, знания соответствующей обстановки. Кроме того, криптоанализ обогатился большим опытом использования в годы войны ошибок неопытных или ленивых шифровальщиков. Еще Ф. Бэкон писал, что "в результате неловкости и неискусности тех рук, через которые проходят величайшие секреты, эти секреты во многих случаях оказывались обеспеченными слабейшими шифрами" [Кан67]. Этот печальный опыт привел к необходимости введения строгой дисциплины среди шифровальщиков.

Несмотря на указанные последствия, первая мировая война не породила никаких новых научных идей в криптографии. Наоборот, полностью исчерпали свои возможности ручное шифрование, с одной стороны, и техническая сторона криптоанализа, состоявшая в подсчете частот встречаемости знаков, с другой.

В тот период проявились таланты целого ряда ставших впоследствии известными криптографов. В их числе был Г.О. Ярдли, который вскоре после объявления США войны в 1917 г. *убедил* военное министерство в необходимости создания криптографической службы. В 27 лет он был назначен начальником криптографического отдела (MI-8) разведки военного министерства. При отделе было создано учебное отделение по подготовке криптоаналитиков для американской армии. Отдел MI-8 добился больших успехов в дешифровании дипломатической переписки многих развитых стран. В 1919 г. отдел был преобразован в "черный кабинет" с совместным финансированием от военного министерства и госдепартамента в объеме 100 тыс. долларов в год. Одной из главных задач "черного кабинета" было раскрытие японских кодов, некоторые из которых содержали до 25 тысяч кодовых величин. В период с 1917 г. по 1929 г. специалистам "черного кабинета" удалось дешифровать более 45 тысяч криптограмм различных стран, в том числе и Япония.

Ярдли, желая упрочить успехи, подготовил докладную записку Президенту США о мерах по укреплению своей службы. Однако ставший в то время Государственным секретарем Г. Стимсон был шокирован, узнав о существовании "черного кабинета", и полностью осудил его деятельность. Ему принадлежит знаменитая фраза: "Джентльмены не читают писем друг друга". Финансирование "черного кабинета" было прекращено, и Ярдли лишился работы. Он написал книгу "Американский черный кабинет", в которой рассказал о многих успехах по дешифрованию. Книга была издана большими тиражами в ряде стран и произвела

эффект разорвавшейся бомбы. Позже он написал книгу "Японские дипломатические секреты", в которой приводились многие японские телеграммы. Рукопись этой книги была конфискована по решению суда. Последние годы жизни Ярдли не занимался криптографией. Он умер в 1958 г. и был похоронен с воинскими почестями на Арлингтонском национальном кладбище. В некрологе он был назван "отцом американской криптографии".

Значительный успех в криптографии связан с еще одним американцем – Г. Вернамом. В 1917 г. он, будучи сотрудником телеграфной компании, предложил идею автоматического шифрования телеграфных сообщений. Речь шла о своеобразном наложении гаммы на знаки алфавита, представленные в соответствии с телетайпным кодом Бодо пятизначными "импульсными комбинациями". Например, буква *a* представлялась комбинацией (+ + – – –), а комбинация (+ + – + +) представляла символ перехода от букв к цифрам. На бумажной ленте, используемой при работе телетайпа, знаку "+" отвечало наличие отверстия, а знаку "–" – его отсутствие. При считывании с ленты металлические щупы проходили через отверстия, замыкали электрическую цепь и тем самым посылали в линию импульс тока.

Вернам предложил электромеханически по координатно складывать "импульсы" знаков открытого текста с импульсами" гаммы, предварительно нанесенными на ленту. Сложение проводилось "по модулю 2". Имеется в виду, что если "+" отождествить с 1, а "–" с 0, то сложение определяется двоичной арифметикой:

+	0	1
0	0	1
1	1	0

Например, наложение на знак открытого текста (11001) знака гаммы (01111) давало знак шифртекста (10110). При расшифровании нужно было произвести ту же операцию со знаком шифртекста: $(10110) \oplus (01111) = (11001)$.

Вернам сконструировал и устройство для такого сложения. Замечательно то, что процесс шифрования оказывался полностью автоматизированным, в предложенной схеме исключался шифровальщик. Кроме того, оказывались слитыми воедино процессы шифрования-расшифрования и передачи по каналу связи. Тем самым наряду с традиционной схемой *предварительного шифрования*, когда по каналу передается предварительно зашифрованное сообщение, положено начало *линейному шифрованию*.

В 1918 г. два комплекта соответствующей аппаратуры были изготовлены и испытаны. Испытания дали положительные результаты. Единственное неудовлетворение специалистов-криптографов было связано с гаммой. Дело в том, что первоначально гамма была нанесена на ленту, склеенную в кольцо. Несмотря на то, что знаки гаммы на ленте выбирались случайно, при зашифровании длинных сообщений гамма регулярно повторялась. Этот недостаток так же отчетливо осознавался, как и для шифра Внженера. Уже тогда хорошо

понимали, что повторное использование гаммы недопустимо даже в пределах одного сообщения. Хотя сам Вернам не был математиком, он, может и неосознанно, предлагал однократное использование гаммы. Попытки удлинить гамму приводили к неудобствам в работе с длинным кольцом. Тогда был предложен вариант с двумя лентами, одна из которых шифровала другую, в результате чего получалась гамма, имеющая длину периода, равную произведению длин исходных периодов.

Несмотря на то, что *шифр Вернама* обладал целым рядом достоинств, он не получил широкого распространения. Трудности, связанные с изготовлением, рассылкой и учетом использованной гаммы, особенно в условиях военной связи, при передаче больших объемов сообщений, стали непреодолимыми. Вспомнили о шифре Вернама лишь накануне II мировой войны.

Почти половина XX в. была связана с использованием *колесных шифраторов*. Различные их конструкции были запатентованы примерно в одно и то же время (в период 1917 – 1919 гг.) в разных странах: американцем Э. Х. Хеберном, голландцем Х. Ф. Кохом, немцем А. Шербиусом и шведом А. Г. Даммом.

Чертежи своей схемы на основе *шифрующего диска* Хеберн представил в 1917 г., и уже в следующем году был построен первый дисковый аппарат, получивший одобрение ВМС США. В 1921 г. Хеберн основал первую в США компанию по производству шифрмашин, которую через десять лет ждал бесславный конец, связанный с финансовыми трудностями.

Что представлял собой шифрующий диск? Корпус диска (имевшего размеры хоккейной шайбы) состоял из изоляционного материала, например твердой резины. По окружности каждой из его сторон были вмонтированы на равном расстоянии друг от друга 26 электрических контактов (см. рис. 1.3.). Каждый контакт был соединен внутри корпуса с некоторым контактом на другой стороне. Контакты на *входной стороне* представляли буквы открытого текста, контакты на *выходной стороне* – буквы шифртекста.

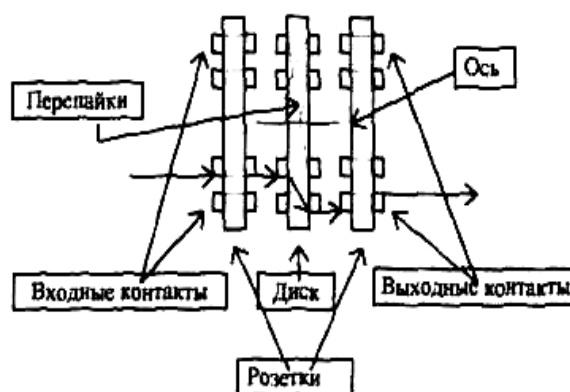


Рисунок 6.3. Шифрующий диск

Диск устанавливался на оси между двумя неподвижными пластинами (*розетками*), каждая из которых также была изготовлена из изолятора и имела 26 контактов, соответствующих расположению контактов на диске. Контакты входной розетки соединялись с клавиатурой пишущей машинки, печатающей буквы открытого текста. Контакты выходной розетки соединялись с выходным устройством, указывающим буквы шифртекста, например, с помощью лампочек. При фиксированном угловом положении диска электрические цепи, соединяющие входные и выходные контакты, реализовывали одноалфавитную замену. При повороте же диска (на углы $2\pi k/26$) схема реализовывала многоалфавитную замену (с 26 простыми заменами).

Рядом с одним диском можно было установить и другие диски. Тем самым схема токопрохождения удлинялась и число возможных простых замен, реализуемых многодисковой схемой значительно возрастало. При движении k дисков по простейшей схеме одометра получался период, равный 26^k , который можно было сделать астрономическим числом.

Подобные шифрмашинки обслуживали значительную часть линий связи высшего командования ВМС США, начиная с 20-х годов.

Х. Ф. Кох предлагал конструкцию шифрующего диска, в котором роль электричества выполняла пневматика. Речь идет о каналах, соединяющих входные и выходные контакты, по которым может проходить поток воздуха, водная или масляная струя и т.п. Любопытно, что подобные дисковые системы на основе пневматики были реально изготовлены и использовались на практике.

Принцип шифрующего диска использовали и шифрмашинки, разработанные А. Шербиусом. Самой знаменитой из них была "*Энигма*", которая в двух отношениях отличалась от других дисковых машин. Во-первых, после блока дисков была расположена неподвижная *обратимая розетка*, контакты которой были попарно соединены друг с другом. Импульс тока, приходивший на этот контакт, заворачивался и вновь проходил через блок дисков в противоположном направлении. Это давало двойное шифрование каждой буквы. Другая особенность "*Энигмы*" заключалась в неравномерном движении дисков, которое управлялось зубчатыми колесами.

В 1923 г. "*Энигма*" выставлялась на конгрессе международного почтового союза, однако это не способствовало ее коммерческому успеху: она не раскупалась. За десять лет фирма Шербиуса, производившая "*Энигму*", не получила прибыли и в 1934 г. была ликвидирована и передала свои активы другой фирме. После прихода к власти в Германии Гитлера началось серьезное перевооружение армии, и немецкие специалисты сочли "*Энигму*" достаточно

удобной и надежной шифрмашинной. В довоенный период и во время II мировой войны "Энигма" широко использовалась в германской армии, ВМС и ВВС. Она была портативной (размером с пишущую машинку), работала от батареи, имела деревянный футляр. Ее серьезный недостаток состоял в том, что она не печатала шифртекст (а имела лишь загорающиеся лампочки, отвечающие буквам), и для быстрой работы требовалось три или даже четыре человека – для чтения и набора на клавиатуре текста сообщения, диктовки высвечивающихся букв шифртекста и их записи.

С "Энигмой" теснейшим образом связан ход многих событий периода II мировой войны. Дело в том, что она являлась источником ценнейших сведений для английских спецслужб, читавших переписку "Энигмы" (в рамках *операции "Ультра"*). Эта информация стоила так дорого, что У. Черчилль пожертвовал городом Ковентри, когда ему стал известен план германской бомбардировки этого английского города. С "Энигмой" связано также появление первой в истории вычислительной машины, сконструированной в 1942 г. для перебора ключевых элементов группой специалистов-криптографов под руководством известного математика А. Тьюринга.

Еще один патент на дисковую машину был выдан А. Г. Дамму в 1919 г. Устройство этой машины было настолько сложным, что никогда не было реализовано. Но его автор основал компанию по производству шифрмашин, которая впоследствии стала прибыльной. Среди вкладчиков капитала были Э. Нобель, племянник знаменитого А. Нобеля, и Ц. Хагелин, управляющий нефтедобывающей компанией братьев Нобелей в России и некоторое время бывший генеральным консулом Швеции в Санкт-Петербурге. До 1927 г. эта компания не имела больших успехов. Их появление было связано с именем сына Ц. Хагелина – Б. Хагелина, родившегося на Кавказе, проучившегося несколько лет в Петербургском университете и получившего позже диплом инженера-механика в Швеции.

В 1925 г. Б. Хагелину удалось модернизировать одну из машин Дамма, снабдив ее клавиатурой и индикаторными лампочками, как у "Энигмы". Это была также колесная машина, работающая, однако, по иному, чем дисковые машины, принципу. Это была машина В-21. Ее работа была основана на матричном *коммутаторе*, в котором электрически изменялось соединение строк и столбцов для преобразования буквы открытого текста в букву шифртекста. Эти изменения определялись группой ключевых колес, каждое из которых имело по ободу выдвинутые или вдвинутые *штифты*. Колеса имели различные числа штифтов, так что период многоалфавитного шифра, реализуемого машиной, был равен произведению чисел штифтов на всех колесах. В 1926 г. Б. Хагелин предложил В-21 шведской армии, которая сделала на нее большой заказ.

В 1927 г. Б. Хагелин возглавил фирму, выкупленную семьей Хагелин. Свою следующую машину В-211 он снабдил печатающим устройством, работавшим со скоростью около 200 знаков в минуту. Она была самой портативной печатающей шифрмашинкой в 1934 г.

В том же году французский генштаб заказал Б. Хагелину карманную печатающую машину, которая могла бы обслуживаться одним человеком. Через некоторое время такая машина была изготовлена. Она реализовывала шифр гаммирования, причем для выработки гаммы была использована идея суммирующего устройства, состоящего из комбинационных линеек, расположенных в цилиндрическом барабане. На линейках рядами были расположены так называемые *рейтеры*. При повороте барабана на 360° рейтеры, вступая во взаимодействие с другими элементами схемы, могли выдвигать некоторые линейки влево, причем число выдвинутых линеек и определяло значение знака гаммы (от 0 до 25) в данный такт шифрования. Во взаимодействие с рейтерами вступали штифты, расположенные на колесах блока дисков, составляющего вторую основную часть машины. Размеры и схема движения дисков обеспечивали период, приблизительно равный $1,01 \cdot 10^8$. Как расположение рейтеров, так и расположение штифтов могло легко меняться, они являлись ключевыми элементами. Это была машина С-36, ставшая впоследствии знаменитой. По размерам она была меньше телефонного аппарата, весила вместе с футляром около двух с половиной килограммов. Французы сразу же сделали заказ на 5000 машин. Позднее машина была существенно усовершенствована, ею заинтересовались в США. В 1939 г. она была взята на вооружение армии США. Под военным наименованием М-209 она использовалась в качестве полевого шифра на протяжении всей II мировой войны. Всего было произведено около 140000 таких машин. Позже фирма Хагелин стала производить широко известные машины С-48, С-52, Т-55 и многие другие. Среди заметных фигур в криптографии первой половины XX в. выделяется У. Фридман, получивший серьезные теоретические результаты в криптоанализе и ставший известным благодаря своим заслугам по вскрытию военных шифров Японии и Германии.

У. Фридман родился в 1891 г. в Кишиневе, в семье переводчика, работавшего в русском почтовом ведомстве. В 1892 г. его семья эмигрировала в США, где отец стал заниматься швейными машинами. У. Фридман в 1914 г. окончил Корнельский университет по специальности генетика. В городе Итака, где проживала семья Фридмана, крупный бизнесмен Д. Фабиан имел собственные лаборатории по акустике, генетике и криптографии. Любопытно, что криптографией Д. Фабиан увлекся, пытаясь доказать, что автором пьес У. Шекспира являлся Ф. Бэкон.

В 1915 г. Д. Фабиан нанял на работу в свое поместье Ривербэнк специалиста по генетике. Им стал У. Фридман. Вскоре он увлекся криптографией и проявил себя в этом деле. Через

некоторое время У. Фридман уже возглавлял в Ривербэнкских лабораториях два отдела – генетики и шифров.

Помимо криптоаналитической работы У. Фридман занимался преподаванием в классе, состоявшем из армейских офицеров, присланных в Ривербэнк для изучения криптографии. До 1918 г. им был подготовлен цикл из семи лекций, восьмую он написал после возвращения со службы в качестве дешифровальщика в американских экспедиционных силах (шла I мировая война). Известные все вместе как *Ривербэнкские публикации*, эти работы являются серьезным вкладом в теоретическую криптографию.

Наибольший интерес с точки зрения современной криптографии представляют лекции "Методы раскрытия шифров с длинной связной гаммой" и "Индекс совпадения и его применения в криптографии" ([Fri20]). В первой из них предлагается бесключевой метод чтения при использовании неравновероятной гаммы. Во второй излагается так называемый *k-тест*, позволяющий выяснить, можно ли подписать друг под другом две (или более) криптограммы (или отрезки криптограмм) так, чтобы буквы в каждой колонке оказались бы зашифрованы одинаковыми знаками гаммы.

Поступив в 1921 г. на службу в войска связи, У. Фридман успешно применял свои методы для вскрытия машинных шифров. Когда была создана служба радиоразведки, У. Фридман стал ее главой и продолжил свои разработки, самой значимой из которых было вскрытие японской *пурпурной шифрмашинны*. В 1929 г. он стал широко известен как один из ведущих криптографов мира, когда "Британская энциклопедия" поместила его статью "О кодах и шифрах". С основными результатами У. Фридмана можно познакомиться в четырехтомнике "Военная криптография".

Выдающиеся результаты в применении математических методов в криптографии принадлежат Клоду Шеннону. К. Шеннон получил образование по электронике и математике в Мичиганском университете, где и начал проявлять интерес к теории связи и теории шифров. В 1940 г. он получил степень доктора по математике, в течение года обучался в Принстонском институте усовершенствования, после чего был принят на службу в лабораторию компании "Bell Telephone".

К 1944 г. К. Шеннон завершил разработку теории секретной связи. В 1945 г. им был подготовлен секретный доклад "Математическая теория криптографии", который был рас-секречен в 1949 г.

В данной работе излагается теория так называемых *секретных систем*, служащих фактически математической моделью шифров. Помимо основных алгебраических (или функциональных) свойств шифров, постулируемых в модели, множества сообщений и ключей наделяются соответствующими априорными вероятностными свойствами, что позволяет

формализовать многие постановки задач синтеза и анализа шифров. Так, и сегодня при разработке новых классов шифров широко используется принцип Шеннона *рассеивания и перемешивания*, состоящий в использовании при шифровании многих итераций "рассеивающих" и "перемешивающих" преобразований.

Разработанные К. Шенноном концепции *теоретической и практической секретности* (или стойкости) позволяют количественно оценивать криптографические качества шифров и пытаться строить в некотором смысле *идеальные* или *совершенные шифры*. Моделируется также и язык открытых сообщений. А именно, предлагается рассматривать язык как *вероятностный процесс*, который создает дискретную последовательность символов в соответствии с некоторой вероятностной схемой.

Центральной в работах К. Шеннона является концепция *избыточной информации*, содержащейся в текстовых сообщениях. Избыточность означает, что в сообщении содержится больше символов, чем в действительности требуется для передачи содержащейся в нем информации. Например, всего лишь десять английских слов – the, of, and, to, a, in, that, it, is, i – составляют более 25% любого (английского) текста. Легко понять, что их можно изъять из текста без потери информации, так как их легко восстановить по смыслу (или по контексту). Фактически К. Шеннон показал, что успех криптоанализа определяется тем, насколько избыточность, имеющаяся в сообщении, "переносится" в зашифрованный текст. Если шифрование "стирает" избыточность, то восстановить текст сообщения по криптограмме становится принципиально невозможно.

Задачу дешифрования К. Шеннон рассматривает как задачу вычисления апостериорных знаний противника о шифре после перехвата криптограммы. Дело в том, что вероятности сообщений и ключей составляют априорные знания противника, которыми он располагает в соответствии с правилом Керкгоффса. После перехвата криптограммы он может (по крайней мере, в принципе, поскольку множества сообщений и ключей *конечны*) вычислить апостериорные вероятности возможных ключей и сообщений, которые могли быть использованы при составлении данной криптограммы. Вот эти вероятности и составляют апостериорные знания противника. С этой точки зрения показателен следующий пример.

Пусть для зашифрования нормативного английского языка применяется шифр простой замены, в котором каждый из $26!$ ключей может быть выбран с равной вероятностью. Пусть противник знает об источнике сообщений лишь то, что он создает английский текст. Тогда априорными вероятностями различных сообщений из N букв являются их относительные частоты в нормативном тексте. Если же противник перехватил криптограмму из N букв, то он может вычислить условные вероятности открытых текстов и ключей, которые могут создать такую криптограмму. Если N достаточно велико, скажем $N = 50$, то *обычно* имеется

единственное сообщение (и единственный ключ) с условной вероятностью, близкой к единице (это – само сообщение, подвергнутое шифрованию), в то время как все другие сообщения имеют суммарную вероятность, близкую к нулю. Таким образом, имеется, по существу, единственное "решение" такой криптограммы. Для меньших значений N , скажем $N = 10$, обычно найдется несколько пар сообщений и ключей, вероятности которых сравнимы друг с другом, то есть нет ни одного сообщения (и ключа) с вероятностью, близкой к единице. В этом случае "решение" криптограммы неоднозначно.

Понятие *совершенной секретности* К. Шеннон определяет требованием, чтобы апостериорные знания противника в точности совпадали бы с априорными знаниями. Он приводит пример *совершенного шифра*, которым является шифр Вернама (со случайной равновероятной гаммой). Следует подчеркнуть, что все рассуждения о стойкости шифров К. Шеннон проводит лишь для одной постановки задачи криптоанализа: когда противник располагает лишь *одной криптограммой* и требуется найти текст сообщения. Для других постановок задач требуются отдельные исследования.

Теоретической мерой секретности (или стойкости) по К. Шеннону является энтропийная характеристика – *неопределенность шифра по открытому сообщению*, которая измеряет (в статистическом смысле), насколько "близка" *средняя* криптограмма из N букв к единственному "решению". Он выводит формулу для приближенного вычисления минимального N , при котором находится единственное "решение". Такая величина получила название *расстояния единственности*. Формула для расстояния единственности связывает между собой неопределенность шифра по открытому тексту и избыточность текста. Чем большим оказывается расстояние единственности, тем более шифр приближается к совершенному шифру, для которого формально расстояние единственности равно ∞ .

Наконец, К. Шеннон вводит понятие *рабочей характеристики* шифра, подходя к практической оценке стойкости. Он формулирует также основные критерии оценки качества секретных систем с позиций практики их использования.

Как видим, К. Шеннону удалось решить фундаментальные проблемы в теоретической криптографии. Его работы стимулировали бурный рост научных исследований по теории информации и криптографии.

В работах К. Шеннона по исследованию свойств языка важную роль играет величина удельной энтропии H на букву текста, другими словами, среднее количество информации, передаваемой буквой открытого текста. Предложенный им метод экспериментов с угадыванием очередной буквы английского текста по предыдущим буквам оказался неэффективным при получении оценок величины H для других языков. Метод "отгадывания" развил в своих работах А. Н. Колмогоров. Достаточно точные приближения

параметра H для русского и французского языков получил Б. Б. Пиотровский. Он указал на существенную разницу между значениями H для текстов различного характера (литературных, деловых, разговорной речи).

Понятие "количества информации", содержащейся в тексте, базировалось, по К. Шеннону, лишь на частотных характеристиках. В своих фундаментальных работах 60-х годов А.Н. Колмогоров подошел к определению количества информации с учетом *смыслового содержания* текста, что позволило уточнить приближение величины H для литературных текстов. Необходимо также отметить, что еще задолго до К. Шеннона частотные характеристики языка изучал выдающийся русский ученый А. А. Марков. Сегодня часто используются так называемые *марковские модели* открытых текстов, учитывающие зависимости букв текста от предыдущих букв.

Следующая страница в истории криптографии XX в. посвящена телефонным шифраторам, которые были разработаны в 30-х годах и стали широко использоваться во время II мировой войны. В России разработка телефонного шифратора велась под руководством В. А. Котельникова, ставшего впоследствии академиком, ученым с мировым именем. Ему принадлежит знаменитая *теорема дискретизации* (или *теорема отсчетов*), лежащая в основе теории *цифровой обработки сигналов*.

Идея телефонного шифратора была запатентована Д.Х.Роджерсом еще в 1881 г., спустя пять лет после изобретения Беллом телефона. Идея состояла в передаче телефонного сообщения по нескольким (в простейшем случае – по двум) цепям поочередными импульсами в некоторой быстро изменяющейся последовательности. Предлагалось разнести такие линии на значительное расстояние друг от друга с тем, чтобы устранить возможность подключения сразу ко всем одновременно. Подключение же к одной из них позволяло бы слышать лишь отдельные неразборчивые сигналы.

В более поздних разработках предлагались различные преобразования непосредственно самой речи. Звуки речи преобразуются телефоном в непрерывный электрический сигнал, который с помощью соответствующих устройств изменяется шифратором по законам электричества. К числу возможных изменений относятся: *инверсия*, *смещение*, или *деление диапазона частот*, *шумовые маскировки*, *временные перестановки* частей сигнала, а также различные комбинации перечисленных преобразований. Естественно, каждое из указанных преобразований производится под управлением ключа, который имеется у отправителя и получателя. Наиболее просто реализуемым являлось преобразование инверсии. Сложнее реализовались временные перестановки. Для их осуществления речевой сигнал в некоторый промежуток времени предварительно записывался на магнитофонной ленте. Запись делилась на отрезки длительностью в доли секунд. Отрезки с помощью нескольких магнитных

головок разносились и перемешивались, в результате чего в канале слышалась хаотическая последовательность звуков. Использовалась также движущаяся магнитная головка, которая в зависимости от направления движения считывала сигналы быстрее или медленнее, чем они были записаны на ленте. В результате тон сигналов становился выше или ниже обычного, в канале быстро чередовались высокие и низкие звуки, не воспринимаемые ухом. Следует отметить, что одной из самых сложных проблем, которые возникали при разработке телефонных шифраторов, была *проблема узнавания* восстановленной после расшифрования речи.

В США первый *телефонный шифратор*, под названием АЗ, был принят в эксплуатацию в 1937 г. Именно он доставил президенту Рузвельту известие о начале II мировой войны утром 1 сентября 1939 г. по вызову американского посла в Париже. АЗ осуществлял инверсию и перестановку 5 поддиапазонов частот. Из 3840 возможных комбинаций ($5! \cdot 2^5$) фактически использовались лишь 6, которые менялись 36 раз за каждые 20 секунд. Слабость используемой криптографии компенсировалась регулярным изменением частот передачи.

В настоящее время аналоговая телефония уступает место цифровой телефонии. Тем самым и многие технические проблемы, связанные с криптографическими преобразованиями аналоговых сигналов, отпадают за ненадобностью. Дело в том, что *оцифрованный* сигнал является дискретным и, следовательно, к нему можно применить хорошо разработанную надежную "дискретную криптографию".

Во второй половине XX в. вслед за развитием элементной базы вычислительной техники, появились *электронные шифраторы*, разработка которых потребовала серьезных теоретических исследований во многих областях прикладной и фундаментальной математики, в первую очередь алгебре, теории вероятностей и математической статистике. Сегодня именно электронные шифраторы составляют подавляющую долю средств шифрования. Они удовлетворяют всё возрастающим требованиям по надежности и скорости шифрования. Прогресс в развитии вычислительной техники сделал возможными *программные реализации* криптографических алгоритмов, которые все увереннее вытесняют во многих сферах традиционные аппаратные средства.

В семидесятых годах произошло два события, серьезно повлиявших на дальнейшее развитие криптографии. Во-первых, был принят (и опубликован!) первый *стандарт шифрования данных* (DES), "легализовавший" принцип Керкгоффа в криптографии. Во-вторых, после работы американских математиков У. Диффи и М. Хеллмана [36] родилась "новая криптография" – *криптография с открытым ключом*. Оба этих события были рождены потребностями бурно развивающихся средств коммуникаций, в том числе локальных и глобальных компьютерных сетей, для защиты которых потребовались легко

доступные и достаточно надежные криптографические средства. Криптография стала широко востребоваться не только в военной, дипломатической, государственной сферах, но также в коммерческой, банковской и других сферах.

Вслед за идеей Диффи и Хеллмана, связанной с гипотетическим понятием *однонаправленной (или односторонней) функции с секретом*, появились "кандидат" на такую функцию и реально осуществленная шифрсистема RSA с открытым ключом. Такая система была предложена в 1978 г Райвестом, Шамиром и Адлеманом. Парадоксальным казалось то, что в RSA для зашифрования и расшифрования используются разные ключи, причем ключ зашифрования может быть *открытым*, то есть всем известным. Вслед за RSA появился целый ряд других систем. В связи с несимметричным использованием ключей стал использоваться термин *асимметричная шифрсистема*, в то время как традиционные шифрсистемы стали называться *симметричными*.

Наряду с идеей открытого шифрования Диффи и Хеллман предложили идею *открытого распределения ключей*, позволяющую избавиться от защищенного канала связи при рассылке криптографических ключей. Их идея основывалась на сложности решения задачи *дискретного логарифмирования*, то есть задачи, являющейся обратной для задачи возведения в степень в конечном поле большого порядка.

Позволим себе остановиться в нашем кратком историческом обзоре на этом этапе развития криптографии, поскольку рассказ о ее современных проблемах и приложениях заставил бы нас приступить к изложению основного содержания книги.

6.1. Теория классических шифров

Основные характеристики открытого текста

Криптография занимается защитой сообщений, содержащихся на некотором материальном носителе. При этом сами сообщения представляют собой последовательности знаков (или *слова*) некоторого *алфавита*. Различают *естественные* алфавиты, например русский или английский, и *специальные* алфавиты (цифровые, буквенно-цифровые), например двоичный алфавит, состоящий из символов 0 и 1. В свою очередь, естественные алфавиты также могут отличаться друг от друга даже для данного языка.

Алфавиты открытых сообщений

Наиболее привычны буквенные алфавиты, например русский, английский и т. д. Приведем сведения об алфавитах некоторых естественных европейских языков.

Полный русский алфавит состоит из 33 букв:

А Б В Г Д Е Ё Ж З И Й К Л М Н О П Р С Т У Ф Х Ц Ч Ш Щ Ъ Ы Ь Э Ю Я

Вместе с тем используются и сокращенные русские алфавиты, содержащие 32, 31 или 30

букв. Можно отождествить буквы Е и Ё, И и Й, Ь и Ъ. часто бывает удобно включить в алфавит знак пробела между словами, в качестве которого можно взять, например, символ “–”.

Английский алфавит состоит из 26 букв:

A B C D E F G H I J K L M N O P Q R S T U V W X Y Z

Иногда используется сокращенный 25-буквенный алфавит, в котором отождествлены буквы I и J.

Любопытно отметить, что полинезийский язык Самоа имеет алфавит, содержащий всего 16 букв, из которых около 60% – гласных. В арабском языке и иврите согласные вообще не используются. Они опускаются в письменном тексте и восстанавливаются читателем по смыслу.

В вычислительной технике распространены 128-битовые и 256-битовые алфавиты, использующие представление знаков алфавита в виде 7- или 8-значных двоичных комбинаций.

Наиболее известен код ASCII (American Standart Code for Information Interchange) – американский стандартный код информационного обмена.

Буквенный алфавит, в котором буквы расположены в их естественном порядке, обычно называют *нормальным алфавитом*. В противном случае говорят о *смешанных алфавитах*. В свою очередь, смешанные алфавиты делят на *систематически перемешанные алфавиты* и *случайные алфавиты*. К первым относят алфавиты, полученные из нормального на основе некоторого правила, ко вторым – алфавиты, буквы которых следуют друг за другом в хаотическом (или случайном) порядке.

Смешанные алфавиты обычно используются в качестве нижней строки подстановки, представляющей собой ключ шифра простой замены. Для запоминания ключа (это надежнее, чем хранение ключа на некотором носителе) применяется несложная процедура перемешивания алфавита, например, основанная на *ключевом слове*.

Частотные характеристики текстовых сообщений

Криптоанализ любого шифра невозможен без учета особенностей текстов сообщений, подлежащих шифрованию. Глубинные закономерности текстовых сообщений исследуются в теории информации. Наиболее важной для криптографии характеристикой текстов является избыточность текста, введенная К. Шенноном. Именно избыточность открытого текста, проникающая в шифртекст, является основной слабостью шифра.

Более простыми характеристиками текстов, используемыми в криптоанализе, являются такие характеристики, как *повторяемость* букв, пар букв (*биграмм*) и вообще *m-ок* (*m-грамм*), *сочетаемость* букв друг с другом, чередование гласных и согласных и некоторые

другие. Такие характеристики изучаются на основе эмпирических наблюдений текстов достаточно большой длины.

Для установления статистических закономерностей проводилась большая серия экспериментов по оценке вероятностей появления в открытом тексте фиксированных m -грамм (для небольших значений m).

Суть экспериментов состоит в подсчете чисел вхождений каждой из n^m возможных m -грамм в достаточно длинных открытых текстах $T = t_1 t_2 \dots t_l$, составленных из букв алфавита $\{a_1, a_2, \dots, a_n\}$. При этом просматриваются подряд идущие m -граммы текста:

$$t_1 t_2 \dots t_m, t_2 t_3 \dots t_{m+1}, \dots, t_{l-m+1} t_{l-m+2} \dots t_l.$$

Если $\mathcal{G}(a_{i1} a_{i2} \dots a_{im})$ – число появлений m -граммы $a_{i1} a_{i2} \dots a_{im}$ в тексте T , а L – общее число подсчитанных m -грамм, то опыт показывает, что при достаточно больших L частоты

$$\frac{\mathcal{G}(a_{i1} a_{i2} \dots a_{im})}{L} \quad (2)$$

для данной m -граммы мало отличаются друг от друга. В силу этого относительную частоту (1.1) считают приближением вероятности $P(a_{i1} a_{i2} \dots a_{im})$ появления данной m -граммы в случайно выбранном месте текста (такой подход принят при статистическом определении вероятности). Например, при $m = 1$ хорошее приближение вероятностей появления букв достигается на текстах длины в несколько тысяч букв.

Некоторая разница значений частот в приводимых в различных источниках таблицах объясняется тем обстоятельством, что частоты существенно зависят не только от длины текста, но и от его характера. Так, в технических текстах редкая буква Φ может стать довольно частой в связи с частым использованием таких слов, как *функция*, *дифференциал*, *диффузия*, *коэффициент* и т. п.

Еще большие отклонения от нормы в частоте употребления отдельных букв наблюдаются в некоторых художественных произведениях, особенно в стихах. Поэтому для надежного определения средней частоты буквы желательно иметь набор различных текстов, заимствованных из различных источников. Вместе с тем, как правило, подобные отклонения незначительны, и в первом приближении ими можно пренебречь.

В связи с этим подобные таблицы, используемые в криптографии, должны составляться с учетом *характера переписки*.

Для русского языка частоты (в порядке убывания) знаков алфавита, в котором отождествлены Е с Ё , Ь с Ъ , а также имеется знак пробела (–) между словами.

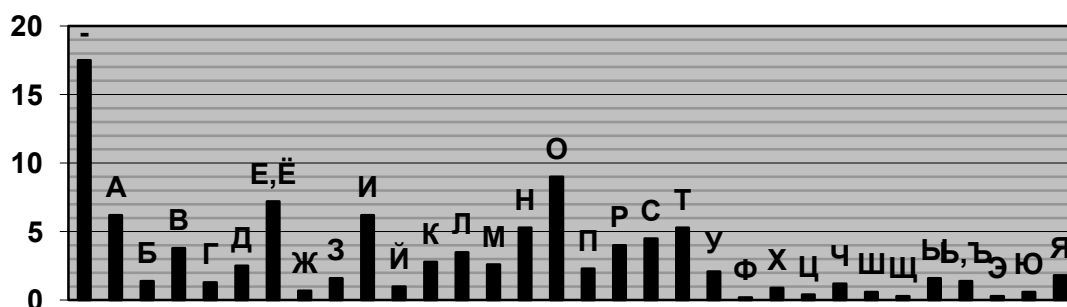


Рис. 1.4. Диаграмма частот букв русского языка

Имеется мнемоническое правило запоминания десяти наиболее частых букв русского алфавита. Эти буквы. СОС13Вляюг нелепое слово СЕНОВАЛИТР. Можно также предложить аналогичный способ запоминания частых букв английского языка, например, с помощью слова TETRIS-HONDA.

Устойчивыми являются также частотные характеристики биграмм, триграмм и четырехграмм осмысленных текстов.

Хорошие таблицы k -грамм легко получить, используя тексты электронных версий многих книг, содержащихся на CD-дисках.

Для получения более точных сведений об открытых текстах можно строить и анализировать таблицы k -грамм при $k > 2$, однако для учебных целей вполне достаточно ограничиться биграммами. Неравновероятность k -грамм (и даже слов) тесно связана с характерной особенностью открытого текста – наличием в нем большого числа повторений отдельных фрагментов текста: корней, окончаний, суффиксов, слов и фраз. Так, для русского языка такими привычными фрагментами являются наиболее частые биграммы и триграммы:

СТ, НО, ЕН, ТО, НА, ОВ, НИ, РА, ВО, КО, СТО, ЕНО, НОВ, ТОВ, ОВО, ОВА

Полезной является информация о сочетаемости букв, то есть о предпочтительных связях букв друг с другом, которую легко извлечь из таблиц частот биграмм.

Имеется в виду таблица, в которой слева и справа от каждой буквы расположены наиболее предпочтительные “соседи” (в порядке убывания частоты соответствующих биграмм). В таких таблицах обычно указывается также доля гласных и согласных букв (в процентах), предшествующих (или следующих за) данной букве.

При анализе сочетаемости букв друг с другом следует иметь в виду зависимость появления букв в открытом тексте от значительного числа предшествующих букв. Для анализа этих закономерностей используют понятие *условной вероятности*.

Наблюдения над открытыми текстами показывают, что для условных вероятностей выполняются неравенства

$$p(a_{i1}) \neq p(a_{i1} / a_{i2}), p(a_{i1} / a_{i2}) \neq p(a_{i1} / a_{i2} a_{i3}), \dots$$

Систематически вопрос о зависимости букв алфавита в открытом тексте от предыдущих букв исследовался известным русским математиком А. А. Марковым (1856 – 1922). Он доказал, что появления букв в открытом тексте нельзя считать независимыми друг от друга. В связи с этим А. А. Марковым отмечена еще одна устойчивая закономерность открытых текстов, связанная с чередованием гласных и согласных букв. Им были подсчитаны частоты встречаемости биграмм вида гласная-гласная (g, g), гласная-согласная (g, c), согласная-гласная (c, g), согласная-согласная (c, c) в русском тексте длиной в 10^5 знаков.

Из этой таблицы видно, что для русского языка характерно чередование гласных и согласных, причем относительные частоты могут служить приближениями соответствующих условных и безусловных вероятностей:

$$p(g/c) \approx 0,663, \quad p(c/g) \approx 0,872$$

$$p(g) \approx 0,432, \quad p(c) \approx 0,568.$$

После А. А. Маркова зависимость появления букв текста вслед за несколькими предыдущими исследовал методами теории информации К. Шеннон. Фактически им было показано, в частности, что такая зависимость ощутима на глубину приблизительно в 30 знаков, после чего она практически отсутствует.

Приведенные выше закономерности имеют место для обычных “читаемых” открытых текстов, используемых при общении людей. Как уже отмечалось ранее, эти закономерности играют большую роль в криптоанализе. В частности, они используются при построении формализованных критериев на открытый текст, позволяющих применять методы математической статистики в задаче распознавания открытого текста в потоке сообщений. При использовании же специальных алфавитов требуются аналогичные исследования частотных характеристик “открытых текстов”, возникающих, например, при межмашинном обмене информацией или в системах передачи данных. В этих случаях построение формализованных критериев на “открытый текст” – задача значительно более сложная.

Помимо криптографии частотные характеристики открытых сообщений существенно используются и в других сферах. Например, клавиатура компьютера, пишущей машинки или линотипа – это замечательное воплощение идеи ускорения набора текста, связанное с оптимизацией расположения букв алфавита относительно друг друга в зависимости от частоты их применения.

Классификация шифров

В качестве первичного признака, по которому производится классификация шифров, используется тип преобразования, осуществляемого с открытым текстом при шифровании. Если фрагменты открытого текста (отдельные буквы или группы букв) заменяются некоторыми их эквивалентами в шифртексте, то соответствующий шифр относится к классу

шифров замены. Если буквы открытого текста при шифровании лишь меняются местами друг с другом, то мы имеем дело с *шифром перестановки*. С целью повышения надежности шифрования зашифрованный текст, полученный применением некоторого шифра, может быть еще раз зашифрован с помощью другого шифра. Всевозможные такие композиции различных шифров приводят к третьему классу шифров, которые обычно называют *композиционными шифрами*. Заметим, что композиционный шифр может не входить ни в класс шифров замены, ни в класс шифров перестановки. В результате получаем первый уровень классификации шифров.

Классификация шифров замены

Если ключ зашифрования совпадает с ключом расшифрования $k_z = k_p$, то такие шифры называют *симметричными* если же $k_z \neq k_p$, – *асимметричными*.

Исторически известный шифр – *пропорциональной замены* представляет собой пример шифра многозначной замены, *шифр гаммирования* – пример шифра однозначной замены.

Также шифры замены делятся на шифры однозначной и многозначной замены, поточные и блочные, одноалфавитные (шифры простой замены) и многоалфавитные.

Классификация шифров в виде схемы приведена на рисунке 1.5.

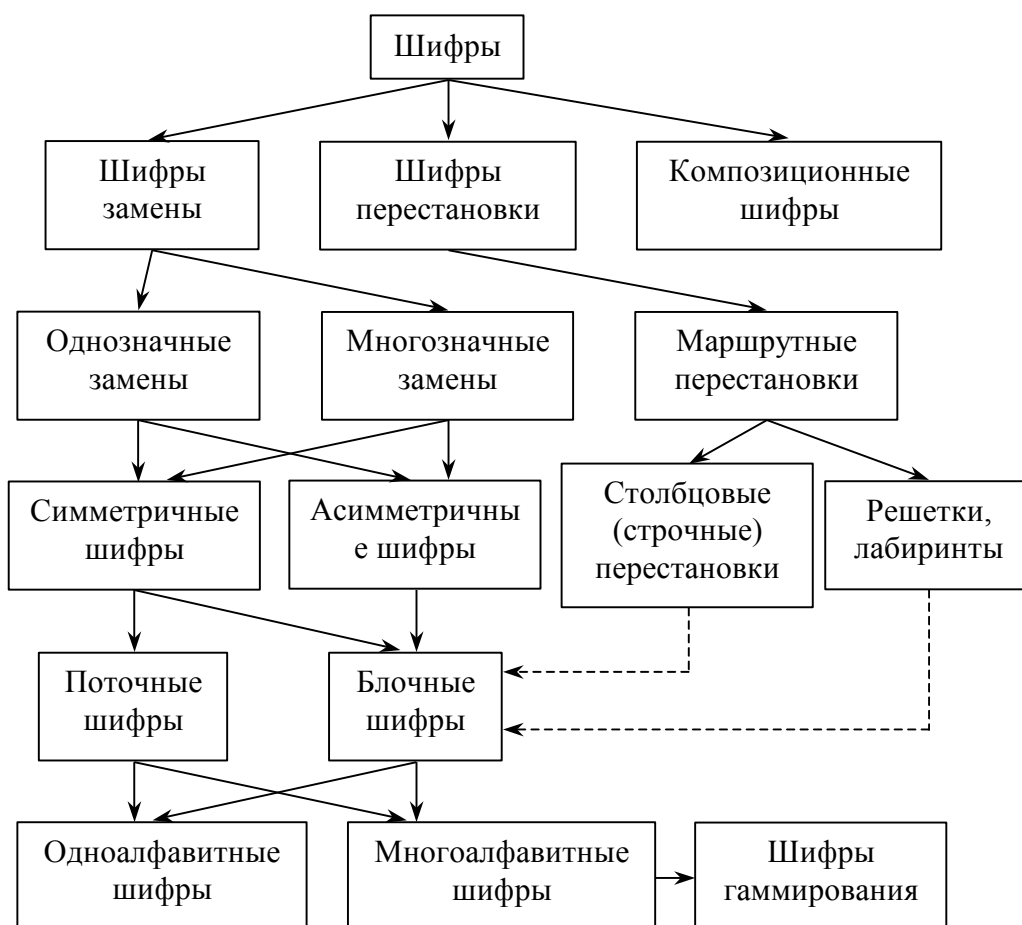


Рис. 6.5. Классификация шифров

Прокомментируем приведенную схему. Подчеркнем, что стрелки, выходящие из любого прямоугольника схемы, указывают лишь на наиболее значимые частные подклассы шифров. Пунктирные стрелки, ведущие из подклассов шифров перестановки, означают, что эти шифры можно рассматривать и как блочные шифры замены в соответствии с тем, что открытый текст делится при шифровании на блоки фиксированной длины, в каждом из которых производится некоторая перестановка букв. Одноалфавитные и многоалфавитные шифры могут быть как поточными, так и блочными. В то же время шифры гаммирования, образующие подкласс многоалфавитных шифров, относится к поточным, а не к блочным шифрам. Кроме того, они являются симметричными, а не асимметричными шифрами.

Шифры перестановки

При шифровании перестановкой символы шифруемого текста переставляются по определенному правилу в пределах блока этого текста. Шифры перестановки являются самыми простыми и, вероятно, самыми древними шифрами.

Широкое применение получили так называемые *маршрутные перестановки*, основанные на некоторой геометрической фигуре. Отрезок открытого текста записывается в такую фигуру по некоторой траектории.

В качестве ключа в шифрующих таблицах используются:

1. размер таблицы;
2. слово или фраза, задающие перестановку;

Широкое распространение получила разновидность маршрутной перестановки, называемая вертикальной перестановкой: сообщение записывается в таблицу по строкам слева направо. Например, зашифруем фразу

ВОТ ПРИМЕР ШИФРА ВЕРТИКАЛЬНОЙ ПЕРЕСТАНОВКИ

Запишем эту фразу в таблицу (таблица 6.5).

Таблица 6.5.

В	О	Т	П	Р	И	М
Е	Р	Ш	И	Ф	Р	А
В	Е	Р	Т	И	К	А
Л	Ь	Н	О	Й	П	Е
Р	Е	С	Т	А	Н	О
В	К	И				

После заполнения таблицы текстом сообщения по строкам для формирования шифртекста считывают содержимое таблицы по столбцам. Получается такое зашифрованное сообщение:

ВЕВЛРВОРЕЪЕКТШРНСИПИТОТРФИЙАИРКПНМААЕО

Естественно, отправитель и получатель сообщения должны заранее условиться об общем ключе в виде размера таблицы.

При расшифровании, в первую очередь, надо определить число длинных столбцов, то есть число букв в последней строке прямоугольника. Для этого нужно разделить число букв в сообщении на число столбцов таблицы. Ясно, что остаток от деления и будет искомым числом. В нашем примере $38 = 7 \times 5 + 3$, поэтому в заполненной таблице имеется 3 длинных и 4 коротких столбца. Когда это число определено, буквы криптограммы можно водворить на их собственные места, и сообщение будет прочитано естественным образом.

Несколько большей стойкостью к раскрытию обладает метод шифрования, называемый одиночной перестановкой по ключу. Этот метод отличается от предыдущего тем, что столбцы таблицы переставляются по ключевому слову, фразе или набору чисел длиной в строку таблицы.

Применим в качестве ключа, например, слово

ПЕЛИКАН,

а текст сообщения возьмем из предыдущего примера (таблица 6.6).

Таблица 6.6

Ключ →

П	Е	Л	И	К	А	Н
7	2	5	3	4	1	6
В	О	Т	П	Р	И	М
Е	Р	Ш	И	Ф	Р	А
В	Е	Р	Т	И	К	А
Л	Ь	Н	О	Й	П	Е
Р	Е	С	Т	А	Н	О
В	К	И				

В верхней строке таблицы записан ключ, а номера под буквами ключа определены в соответствии с естественным порядком соответствующих букв ключа в алфавите. Если бы в ключе встретились одинаковые буквы, они бы были пронумерованы слева направо.

Отметим, что нецелесообразно заполнять последнюю строку прямоугольника "нерабочими" буквами, так как это дало бы противнику, получившему в свое распоряжение

данную криптограмму, сведения о длине числового ключа. В самом деле, в этом случае длину ключа следовало бы искать среди делителей длины сообщения.

При считывании содержимого таблицы по столбцам в порядке, указанном числовым ключом, получим зашифрованное сообщение:

ИРКПН ОРЕЪЕК ПИТОТ РФИЙА ТШРНСИ МААЕО ВЕВЛРВ

При расшифровании, аналогично предыдущему примеру, в первую очередь, надо определить число длинных столбцов. В нашем примере $38 = 7 \times 5 + 3$, поэтому в заполненной таблице имеется 3 длинных и 4 коротких столбца. Согласно числовому ключу, начальные буквы криптограммы берутся из шестого (по счету слева) столбца, он – короткий (так как первые три столбца – длинные), поэтому первые пять букв образуют шестой столбец. Следующие шесть букв образуют второй столбец (он – длинный). И так далее.

Более сложные маршрутные перестановки могут использовать другие геометрические фигуры и более "хитрые" маршруты, как, например, при обходе шахматной доски "ходом коня", пути в некотором лабиринте и т.п. Возможные варианты зависят от фантазии составителя системы и, конечно, естественного требования простоты ее использования.

Для обеспечения дополнительной скрытности можно повторно зашифровать сообщение, которое уже прошло шифрование. Такой метод шифрования называется двойной перестановкой. В случае двойной перестановки столбцов и строк таблицы перестановки определяются отдельно для столбцов и отдельно для строк. Сначала в таблицу записывается текст сообщения, а потом поочередно переставляются столбцы, а затем строки. При расшифровании порядок перестановок должен быть обратным.

В средние века для шифрования перестановкой применялись и магические квадраты.

Магическими квадратами называют квадратные таблицы с вписанными в их клетки последовательными натуральными числами, начиная от 1, которые дают в сумме по каждому столбцу, каждой строке и каждой диагонали одно и то же число.

Шифруемый текст вписывали в магические квадраты в соответствии с нумерацией их клеток. Если затем выписать содержимое такой таблицы по строкам, то получится шифртекст, сформированный благодаря перестановке букв исходного сообщения. В те времена считалось, что созданные с помощью магических квадратов шифртексты охраняет не только ключ, но и магическая сила.

Пример магического квадрата и его заполнения сообщением

ПРИЛЕТАЮ ВОСЬМОГО

16	3	2	13		О	И	Р	М
5	10	11	8		Е	О	С	Ю
9	6	7	12		В	Т	А	Ъ

4	15	14	1		Л	Г	О	П
---	----	----	---	--	---	---	---	---

Рис. 1.6. Пример магического квадрата 4×4 и его заполнения сообщением ПРИЛЕТАЮ ВОСЬМОГО

Шифртекст, получаемый при считывании содержимого правой таблицы по строкам, имеет вполне загадочный вид:

ОИРМ ЕОСЮ ВТАЪ ЛГОП

Число магических квадратов быстро возрастает с увеличением размера квадрата. Существует только один магический квадрат размером 3×3 (если не учитывать его повороты). Количество магических квадратов 4×4 составляет уже 880, а количество магических квадратов 5×5 – около 250000.

Магические квадраты средних и больших размеров могли служить хорошей базой для обеспечения нужд шифрования того времени, поскольку практически нереально выполнить ручную перебор всех вариантов для такого шифра.

Шифры простой замены

При шифровании заменой (подстановкой) символы шифруемого текста заменяются символами того же или другого алфавита с заранее установленным правилом замены.

Будем рассматривать только одноалфавтные однозначные замены, т.е. каждый символ исходного текста заменяется символами того же алфавита одинаково на всем протяжении текста. Такие замены называют *шифрами простой замены*.

Система шифрования Цезаря

Свое название шифр Цезаря получил по имени римского императора Гая Юлия Цезаря, который использовал этот шифр при переписке с Цицероном (около 50 г. до н.э.).

При шифровании исходного текста каждая буква заменялась на другую букву того же алфавита по следующему правилу. Заменяющая буква определялась путем смещения по алфавиту от исходной буквы на k букв. При достижении конца алфавита выполнялся циклический переход к его началу. Цезарь использовал шифр замены при смещении $k = 3$. Такой шифр замены можно задать таблицей подстановок, содержащей соответствующие пары букв открытого текста и шифртекста. Совокупность возможных подстановок для $k = 3$ показана в таблице 1.7.

Таблица 1.7. Одноалфавтные подстановки ($k = 3, m = 26$)

A	→	D	J	→	M	S	→	V
B	→	E	K	→	N	T	→	W

C	→	F
D	→	G
E	→	H
F	→	I
G	→	J
H	→	K
I	→	L

L	→	O
M	→	P
N	→	Q
O	→	R
P	→	S
Q	→	T
R	→	U

U	→	X
V	→	Y
W	→	Z
X	→	A
Y	→	B
Z	→	C

Например, послание Цезаря

VENI VIDI VICI

(в переводе на русский означает "Пришел, Увидел, Победил"), направленное его другу Аминтию после победы над понтийским царем Фарнаком, сыном Митридата, выглядело бы в зашифрованном виде так:

YHQL YLGL YLFL

Формализуем шифр Цезаря. Введем следующие обозначения:

m – число букв в алфавите (мощность алфавита);

(x_0, x_1, \dots) – открытый текст;

(y_0, y_1, \dots) – шифртекст (закрытый текст);

k – ключ, $0 \leq k < m$.

Тогда процесс шифрования можно представить в виде

$y_i = E_k(x_i) = (x_i + k) \bmod m$ – зашифрование,

$x_i = D_k(y_i) = (y_i + 26 - k) \bmod m$ – расшифрование.

Для шифра Цезаря $k = 3$.

Достоинством системы шифрования Цезаря является простота шифрования и расшифрования. К недостаткам системы Цезаря следует отнести следующие:

- подстановки, выполняемые в соответствии с системой Цезаря, не маскируют частот появления различных букв исходного открытого текста. Шифр Цезаря легко вскрывается на основе анализа частот появления букв в шифртексте
- сохраняется алфавитный порядок в последовательности заменяющих букв; при изменении значения k изменяются только начальные позиции такой последовательности;
- число возможных ключей k мало;

Криптоаналитическая атака против системы одноалфавитной замены начинается с подсчета частот появления символов: определяется число появлений каждой буквы в шифртексте. Затем полученное распределение частот букв в шифртексте сравнивается с распределением частот букв в алфавите исходных сообщений, например в английском. Буква с наивысшей частотой появления в шифртексте заменяется на букву с наивысшей частотой

появления в английском языке и т.д. Вероятность успешного вскрытия системы шифрования повышается с увеличением длины шифртекста.

Концепция, заложенная в систему шифрования Цезаря, оказалась весьма плодотворной, о чем свидетельствуют ее многочисленные модификации. Несколько таких модификаций будут рассмотрены ниже.

Аффинная система подстановок Цезаря

В системе шифрования Цезаря использовалось только сложение по модулю m . Применяя одновременно операции сложения и умножения по модулю m над элементами множества $Z_m = \{0, 1, 2, \dots, m-1\}$, можно получить систему подстановок, которую называют аффинной системой подстановок Цезаря.

Введем следующие обозначения:

m – число букв в алфавите (мощность алфавита);

(x_0, x_1, \dots) – открытый текст;

(y_0, y_1, \dots) – шифртекст (закрытый текст);

В качестве ключа используется пара чисел $k = (a, b)$, где a, b – целые числа, $0 < a < m$, $0 \leq b < m$. Числа a и m должны быть взаимно простыми, т.е. наибольший общий делитель НОД $(a, m) = 1$.

Тогда процесс шифрования представляется в виде

$y_i = E_{a,b}(x_i) = (ax_i + b) \bmod m$ – зашифрование,

$x_i = D_{a,b}(y_i) = a^{-1}(y_i + 26 - b) \bmod m$ – расшифрование,

где a^{-1} – обратный a элемент.

Если числа a и m взаимно просты, то существует число a^{-1} , удовлетворяющее сравнению $a \cdot a^{-1} \equiv 1 \pmod m$.

Число a^{-1} называют обратным к a по модулю m . Например, $3^{-1} = 9$, т.е. $3 \cdot 9 \equiv 27 \pmod m \equiv 1 \pmod m$.

Зашифруем слово НОРЕ с помощью аффинного шифра, полагая $k = (3, 5)$.

Для удобства вычисления заменим буквы латинского алфавита на их порядковые номера (Таблица 1.8):

Таблица 6.8

A	B	C	D	E	F	G	H	I	J	K	L	M
0	1	2	3	4	5	6	7	8	9	10	11	12
N	O	P	Q	R	S	T	U	V	W	X	Y	Z
13	14	15	16	17	18	19	20	21	22	23	24	25

Используя аффинный шифр с ключом $k = (3,5)$, получаем следующее соответствие между числовыми кодами букв:

t	0	1	2	3	4	5	6	7	8	9	10	11	12
$3t + 5$	5	8	11	14	17	20	23	0	3	6	9	12	15

t	13	14	15	16	17	18	19	20	21	22	23	24	25
$3t + 5$	18	21	24	1	4	7	10	13	16	19	22	25	2

Преобразуя числа в буквы английского языка, получаем следующее соответствие для букв открытого текста и шифртекста:

A	B	C	D	E	F	G	H	I	J	K	L	M
F	I	L	O	R	U	X	A	D	G	J	M	P

N	O	P	Q	R	S	T	U	V	W	X	Y	Z
S	V	Y	B	E	H	K	N	Q	T	W	Z	C

Исходное сообщение NOPE преобразуется в шифртекст AVYR.

Достоинством аффинной системы является удобное управление ключами – ключи шифрования и расшифрования представляются в компактной форме в виде пары чисел (a, b) . Недостатки аффинной системы аналогичны недостаткам системы шифрования Цезаря.

Аффинная система использовалась на практике несколько веков назад, а сегодня ее применение ограничивается большей частью иллюстрациями основных криптологических положений.

Система Цезаря с ключевым словом

Система шифрования Цезаря с ключевым словом является одноалфавитной системой подстановки. Особенностью этой системы является использование ключевого слова для смещения и изменения порядка символов в алфавите подстановки.

Выберем некоторое число k , $0 \leq k < 26$, и слово или короткую фразу в качестве *ключевого слова*. Желательно, чтобы все буквы ключевого слова были различными. Пусть выбраны слово DIPLOMAT в качестве ключевого слова и число $k = 5$.

Ключевое слово записывается под буквами алфавита, начиная с буквы, числовой код которой совпадает с выбранным числом k :

0	1	2	3	4	5					10					15					20				25	
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
					D	I	P	L	O	M	A	T													

Оставшиеся буквы алфавита подстановки записываются после ключевого слова в алфавитном порядке:

Таблица 1.9. Шифрующая таблица с ключевым словом БАНДЕРОЛЬ

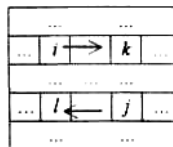
Б	А	Н	Д	Е	Р	О	Л
Ь	В	Г	Ж	З	И	Й	К
М	П	С	Т	У	Ф	Х	Ц
Ч	Ш	Щ	Ы	Ъ	Э	Ю	Я

Процедура шифрования включает следующие шаги:

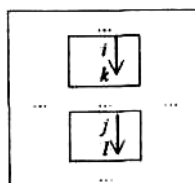
1. Открытый текст исходного сообщения разбивается на пары букв (биграммы). Текст должен иметь четное количество букв и в нем не должно быть биграмм, содержащих две одинаковые буквы. Если эти требования не выполнены, то в текст добавляются “пустышки”. “Пустышкой” является некоторая редкая для данного типа текста буква (или знак), которая вставляется между одинаковыми буквами биграммы или добавляется в текст для того, чтобы его длина стала четной.

2. Последовательность биграмм открытого текста преобразуется с помощью шифрующей таблицы в последовательность биграмм шифртекста по следующим правилам:

а) Если обе буквы биграммы открытого текста не попадают на одну строку или столбец (как, например, буквы А и Й в таблице 3.5), тогда находят буквы в углах прямоугольника, определяемого данной парой букв. (В нашем примере пара букв АЙ отображается в пару ОВ.)

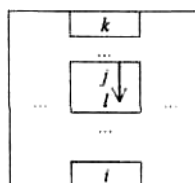


б) Если обе буквы биграммы открытого текста принадлежат одному столбцу таблицы, то буквами шифртекста считаются буквы, которые лежат под ними. (Например, биграмма НС дает биграмму шифртекста ГЩ.)

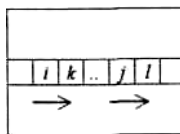


Если при этом буква открытого текста находится в нижней строке, то для шифртекста берется соответствующая буква из верхней строки того же столбца.

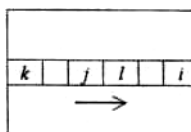
(Например, биграмма ВШ дает биграмму шифртекста ПА.)



с) Если обе буквы биграммы открытого текста принадлежат одной строке таблицы, то буквами шифртекста считаются буквы, которые лежат справа от них. (Например, биграмма НО дает биграмму шифртекста ДЛ.)



Если при этом буква открытого текста находится в крайнем правом столбце, то для шифра берут соответствующую букву из левого столбца в той же строке. (Например, биграмма ФЦ дает биграмму шифртекста ХМ.)



Зашифруем текст

ВСЕ ТАЙНОЕ СТАНЕТ ЯВНЫМ

Разбиение этого текста на биграммы дает

ВС ЕТ АЙ НО ЕС ТА НЕ ТЯ ВН ЫМ

Данная последовательность биграмм открытого текста преобразуется с помощью шифрующей таблицы (таблица 3.5) в следующую последовательность биграмм шифртекста:

ГП ДУ ОВ ДЛ НУ ПД ДР ЦЫ ГА ЧТ

При расшифровании применяется обратный порядок действий.

Криптоанализ шифра Плейфера опирается на частотный анализ биграмм, триграмм и четырехграмм шифртекста и особенности замены шифрвеличин на шифробозначения, связанные с расположением алфавита в прямоугольнике.

Шифр Хилла

Шифр был назван по имени своего создателя Лестора Хилла.

Шифрвеличинами здесь являются n -граммы открытого текста ($n \geq 2$), представленного некоторым числовым кодом.

Правило зашифрования:

Имеется алфавит открытого текста, состоящий из m букв.

Ключ K – некоторая обратимая матрица размером $[n \times n]$ над множеством $Z_m = \{0, 1, 2, \dots, m-1\}$.

Исходный текст разбивается на блоки по n букв (n -граммы), каждую из которых можно представить в виде вектора $\vec{x}_j = (x_{1j}, x_{2j}, \dots, x_{nj})$, где j – номер блока.

$\vec{y}_j = (y_{1j}, y_{2j}, \dots, y_{nj})$ – n -грамма шифртекста.

Процесс шифрования представляется в виде:

$$\vec{y}_j = K \cdot \vec{x}_j \pmod{m} - \text{зашифрование,}$$

$$\vec{x}_j = K^{-1} \cdot \vec{y}_j \pmod{m} - \text{расшифрование,}$$

Рассмотрим пример.

Алфавит: латинский, $m = 26$.

Открытый текст: PAY MORE MONEY.

$$\text{Ключ: } K = \begin{bmatrix} 3 & 3 \\ 2 & 5 \end{bmatrix}.$$

$n = 2$, т.е. разбиение открытого текста производится на биграммы.

Разобьем открытый текст на биграммы и заменим каждую букву ее числовым эквивалентом.

PA	→	15 0
YM	→	24 12
OR	→	14 17
EM	→	4 12
ON	→	14 13
EY	→	4 24

Преобразуем биграммы открытого текста \vec{x}_j в биграммы шифртекста \vec{y}_j .

$$\vec{y}_1 = \begin{bmatrix} 3 & 3 \\ 2 & 5 \end{bmatrix} \cdot \begin{bmatrix} 15 \\ 0 \end{bmatrix} \pmod{26} = \begin{bmatrix} 45 \\ 30 \end{bmatrix} \pmod{26} = \begin{bmatrix} 19 \\ 4 \end{bmatrix},$$

$$\vec{y}_2 = \begin{bmatrix} 3 & 3 \\ 2 & 5 \end{bmatrix} \cdot \begin{bmatrix} 24 \\ 12 \end{bmatrix} \pmod{26} = \begin{bmatrix} 4 \\ 4 \end{bmatrix},$$

$$\vec{y}_3 = \begin{bmatrix} 3 & 3 \\ 2 & 5 \end{bmatrix} \cdot \begin{bmatrix} 14 \\ 17 \end{bmatrix} \pmod{26} = \begin{bmatrix} 15 \\ 9 \end{bmatrix},$$

$$\vec{y}_4 = \begin{bmatrix} 3 & 3 \\ 2 & 5 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 12 \end{bmatrix} \pmod{26} = \begin{bmatrix} 22 \\ 16 \end{bmatrix},$$

$$\vec{y}_5 = \begin{bmatrix} 3 & 3 \\ 2 & 5 \end{bmatrix} \cdot \begin{bmatrix} 14 \\ 13 \end{bmatrix} \pmod{26} = \begin{bmatrix} 3 \\ 15 \end{bmatrix},$$

$$\vec{y}_6 = \begin{bmatrix} 3 & 3 \\ 2 & 5 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 24 \end{bmatrix} \pmod{26} = \begin{bmatrix} 6 \\ 24 \end{bmatrix}.$$

Заменяя в биграммах шифртекста числа на соответствующие буквы согласно таблице 3.4, получаем 12-грамму шифртекста

TE EE PJ WQ DP GY

Для расшифрования биграмм \vec{y}_j , шифртекста и восстановления биграмм \vec{x}_j открытого текста необходимо выполнить обратное преобразование K^{-1} согласно уравнению

$$\vec{x}_j = K^{-1} \cdot \vec{y}_j.$$

В рассмотренном примере матрицы преобразования имели размер 2×2 и шифровались биграммы (пары) букв. Хотя буква Е может быть зашифрована по-разному в различных парах исходного сообщения, одна и та же пара, например ЕМ, будет шифроваться всегда одинаково на протяжении всего исходного текста.

Увеличение значности шифрвеличин (n) резко усложняет попытки вскрытия открытого текста по известному тексту криптограммы.

Шифры сложной замены

Шифры сложной замены называют многоалфавитными, так как для шифрования каждого символа исходного сообщения применяют свой шифр простой замены. Многоалфавитная подстановка последовательно и циклически меняет используемые алфавиты.

Правила замены при r -алфавитной подстановке:

символ x_0 исходного сообщения заменяется символом y_0 из алфавита B_0 ,

символ x_1 заменяется символом y_1 из алфавита B_1 , и так далее,

символ x_{r-1} заменяется символом y_{r-1} из алфавита B_{r-1} ,

символ x_r заменяется символом y_r снова из алфавита B_0 , и т.д.

Общая схема многоалфавитной подстановки для случая $r = 4$ показана на в таблице 6.10.

Таблица 6.10. Схема r -алфавитной подстановки для случая $r = 4$

Входной символ:	X_0	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9
Алфавит подстановки:	B_0	B_1	B_2	B_3	B_0	B_1	B_2	B_3	B_0	B_1

Эффект использования многоалфавитной подстановки заключается в том, что обеспечивается маскировка естественной статистики исходного языка, так как конкретный символ из исходного алфавита A может быть преобразован в несколько различных символов шифровальных алфавитов B_j . Степень обеспечиваемой защиты теоретически пропорциональна длине периода r в последовательности используемых алфавитов B_j .

Одной из старейших и наиболее известных многоалфавитных систем является система шифрования Виженера.

Шифр "двойной квадрат" Уитстона

В 1854 г. англичанин Чарльз Уитстон разработал новый метод шифрования биграммами, который называют "двойным квадратом". Свое название этот шифр получил по аналогии с полибианским квадратом. Шифр Уитстона открыл новый этап в истории развития криптографии. В отличие от полибианского шифра "двойной квадрат" использует сразу две таблицы, размещенные по одной горизонтали, а шифрование идет биграммами, как в шифре Плейфейра. Эти не столь сложные модификации привели к появлению на свет качественно новой криптографической системы ручного шифрования. Шифр "двойной квадрат" оказался очень надежным и удобным и применялся Германией даже в годы второй мировой войны.

Поясним процедуру шифрования этим шифром на примере. Пусть имеются две таблицы со случайно расположенными в них русскими алфавитами (рисунок 1.6).

Ж	Щ	Н	Ю	Р
И	Т	Ь	Ц	Б
Я	М	Е	.	С
В	Ы	П	Ч	
:	Д	У	О	К
З	Э	Ф	Г	Ш
Х	А	,	Л	Ъ

И	Ч	Г	Я	Т
,	Ж	Ь	М	О
З	Ю	Р	В	Щ
Ц	:	П	Е	Л
Ъ	А	Н	.	Х
Э	К	С	Ш	Д
Б	Ф	У	Ы	

Рис. 6.6. Две таблицы со случайно расположенными символами русского алфавита для шифра "двойной квадрат"

Перед шифрованием исходное сообщение разбивают на биграммы. Каждая биграмма шифруется отдельно. Первую букву биграммы находят в левой таблице, а вторую букву – в правой таблице. Затем мысленно строят прямоугольник так, чтобы буквы биграммы лежали в его противоположных вершинах. Другие две вершины этого прямоугольника дают буквы биграммы шифртекста.

Предположим, что шифруется биграмма исходного текста ИЛ. Буква И находится в столбце 1 и строке 2 левой таблицы. Буква Л находится в столбце 5 и строке 4 правой таблицы. Это означает, что прямоугольник образован строками 2 и 4, а также столбцами 1 левой таблицы и 5 правой таблицы. Следовательно, в биграмму шифртекста входят буква О, расположенная в столбце 5 и строке 2 правой таблицы, и буква В, расположенная в столбце 1 и строке 4 левой таблицы, т.е. получаем биграмму шифртекста ОВ.

Ж	Щ	Н	Ю	Р
И	Т	Ь	Ц	Б
Я	М	Е	.	С
В	Ы	П	Ч	
:	Д	У	О	К

И	Ч	Г	Я	Т
,	Ж	Ь	М	О
З	Ю	Р	В	Щ
Ц	:	П	Е	Л
Ъ	А	Н	.	Х

З	Э	Ф	Г	Ш	Э	К	С	Ш	Д
Х	А	,	Л	Ъ	Б	Ф	У	Ы	

Если обе буквы биграммы сообщения лежат в одной строке, то и буквы шифртекста берут из этой же строки. Первую букву биграммы шифртекста берут из левой таблицы в столбце, соответствующем второй букве биграммы сообщения. Вторая же буква биграммы шифртекста берется из правой таблицы в столбце, соответствующем первой букве биграммы сообщения. Поэтому биграмма сообщения ТО превращается в биграмму шифртекста ЖБ. Аналогичным образом шифруются все биграммы сообщения:

Сообщение ПР ИЛ ЕТ АЮ _Ш ЕС ТО ГО

Шифртекст ПЕ ОВ ЦН ФМ ЕШ РФ БЖ ДЦ

Шифрование методом "двойного квадрата" дает весьма устойчивый к вскрытию и простой в применении шифр. Взламывание шифртекста "двойного квадрата" требует больших усилий.

Одноразовая система шифрования

Почти все применяемые на практике шифры характеризуются как условно надежные, поскольку они могут быть в принципе раскрыты при наличии неограниченных вычислительных возможностей. Абсолютно надежные шифры нельзя разрушить даже при использовании неограниченных вычислительных возможностей. Существует единственный такой шифр, применяемый на практике, -одноразовая система шифрования. Характерной особенностью одноразовой системы шифрования является одноразовое использование ключевой последовательности.

Одноразовая система шифрует исходный открытый текст $(x_0, x_1, \dots, x_{n-1})$ в шифртекст $(y_0, y_1, \dots, y_{n-1})$ посредством подстановки Цезаря

$$y_i = (x_i + k_i) \bmod m, 0 \leq i < n,$$

где k_i – i -й элемент случайной ключевой последовательности.

Ключевое пространство одноразовой системы представляет собой набор дискретных случайных величин из Z_m и содержит m^n значений.

Процедура расшифрования описывается соотношением

$$x_i = (y_i + k_i) \bmod m, 0 \leq i < n,$$

где k_i – i -й элемент той же самой случайной ключевой последовательности.

Одноразовая система изобретена в 1917 г. американцами Дж. Моборном и Г. Вернамом. Для реализации этой системы подстановки иногда используют одноразовый блокнот. Этот блокнот составлен из отрывных страниц, на каждой из которых напечатана таблица со случайными числами (ключами) k_i . Блокнот выполняется в двух экземплярах: один используется отправителем, а другой - получателем. Для каждого символа x_i сообщения

используется свой ключ k_i из таблицы только один раз. После того как таблица использована, она должна быть удалена из блокнота и уничтожена. Шифрование нового сообщения начинается с новой страницы.

Этот шифр абсолютно надежен, если набор ключей k_i действительно случаен и непредсказуем. Если криптоаналитик попытается использовать для заданного шифртекста все возможные наборы ключей и восстановить все возможные варианты исходного текста, то они все окажутся равновероятными. Не существует способа выбрать исходный текст, который был действительно послан. Теоретически доказано, что одноразовые системы являются нераскрываемыми системами, поскольку их шифртекст не содержит достаточной информации для восстановления открытого текста.

Казалось бы, что благодаря данному достоинству одноразовые системы следует применять во всех случаях, требующих абсолютной информационной безопасности. Однако возможности применения одноразовой системы ограничены чисто практическими аспектами. Существенным моментом является требование одноразового использования случайной ключевой последовательности. Ключевая последовательность с длиной, не меньшей длины сообщения, должна передаваться получателю сообщения заранее или отдельно по некоторому секретному каналу. Это требование не будет слишком обременительным для передачи действительно важных одноразовых сообщений, например, по горячей линии Вашингтон-Москва. Однако такое требование практически неосуществимо для современных систем обработки информации, где требуется шифровать многие миллионы символов.

В некоторых вариантах одноразового блокнота прибегают к более простому управлению ключевой последовательностью, но это приводит к некоторому снижению надежности шифра. Например, ключ определяется указанием места в книге, известной отправителю и получателю сообщения. Ключевая последовательность начинается с указанного места этой книги и используется таким же образом, как в системе Вижинера. Иногда такой шифр называют шифром с бегущим ключом. Управление ключевой последовательностью в таком варианте шифра намного проще, так как длинная ключевая последовательность может быть представлена в компактной форме. Но с другой стороны, эти ключи не будут случайными. Поэтому у криптоаналитика появляется возможность использовать информацию о частотах букв исходного естественного языка.

Шифрование методом гаммирования

Под *гаммированием* понимают процесс наложения по определенному закону гаммы шифра на открытые данные. *Гамма шифра* – это псевдослучайная последовательность,

выработанная по заданному алгоритму для зашифрования открытых данных и расшифрования зашифрованных данных.

Процесс зашифрования заключается в генерации гаммы шифра и наложении полученной гаммы на исходный открытый текст обратимым образом, например с использованием операции сложения по модулю 2.

Следует отметить, что перед зашифрованием открытые данные разбивают на блоки T_i^o одинаковой длины, обычно по 64 бита. Гамма шифра вырабатывается в виде последовательности блоков Γ_i^u аналогичной длины.

Уравнение зашифрования можно записать в виде

$$T_i^u = \Gamma_i^u \oplus T_i^o, i = 1 \dots M,$$

где T_i^u – i -й блок шифртекста;

Γ_i^u – i -й блок гаммы шифра;

T_i^o – i -й блок открытого текста;

M – количество блоков открытого текста.

Процесс расшифрования сводится к повторной генерации гаммы шифра и наложению этой гаммы на зашифрованные данные. Уравнение расшифрования имеет вид

$$T_i^o = \Gamma_i^u \oplus T_i^u.$$

Получаемый этим методом шифртекст достаточно труден для раскрытия, поскольку теперь ключ является переменным. По сути дела гамма шифра должна изменяться случайным образом для каждого шифруемого блока. Если период гаммы превышает длину всего шифруемого текста и злоумышленнику неизвестна никакая часть исходного текста, то такой шифр можно раскрыть только прямым перебором всех вариантов ключа. В этом случае криптостойкость шифра определяется длиной ключа.

Методы генерации псевдослучайных последовательностей чисел

При шифровании методом гаммирования в качестве ключа используется случайная строка битов, которая объединяется с открытым текстом, также представленным в двоичном виде (например, A = 00000, B = 00001, C = 00010 и т.д.), с помощью побитового сложения по модулю 2, и в результате получается шифрованный текст. Генерирование непредсказуемых двоичных последовательностей большой длины является одной из важных проблем классической криптографии. Для решения этой проблемы широко используются генераторы двоичных псевдослучайных последовательностей.

Генерируемые псевдослучайные ряды чисел часто называют гаммой шифра или просто гаммой (по названию буквы γ греческого алфавита, часто используемой в математических формулах для обозначения случайных величин).

Обычно для генерации последовательности псевдослучайных чисел применяют компьютерные программы, которые, хотя и называются генераторами случайных чисел, на самом деле выдают детерминированные числовые последовательности, которые по своим свойствам очень похожи на случайные.

К криптографически стойкому генератору псевдослучайной последовательности чисел (гаммы шифра) предъявляются три основных требования:

- период гаммы должен быть достаточно большим для шифрования сообщений различной длины;
- гамма должна быть практически непредсказуемой, что означает невозможность предсказать следующий бит гаммы, даже если известны тип генератора и предшествующий кусок гаммы;
- генерирование гаммы не должно вызывать больших технических сложностей.

Длина периода гаммы является самой важной характеристикой генератора псевдослучайных чисел. По окончании периода числа начнут повторяться, и их можно будет предсказать. Требуемая длина периода гаммы определяется степенью закрытости данных. Чем длиннее ключ, тем труднее его подобрать. Длина периода гаммы зависит от выбранного алгоритма получения псевдослучайных чисел.

Второе требование связано со следующей проблемой: как можно достоверно убедиться, что псевдослучайная гамма конкретного генератора является действительно непредсказуемой? Пока не существуют такие универсальные и практически проверяемые критерии и методики. Чтобы гамма считалась непредсказуемой, т.е. истинно случайной, необходимо, чтобы ее период был очень большим, а различные комбинации битов определенной длины были равномерно распределены по всей ее длине.

Третье требование обуславливает возможность практической реализации генератора программным или аппаратным путем с обеспечением необходимого быстродействия.

Один из первых способов генерации псевдослучайных чисел на ЭВМ предложил в 1946 г. Джон фон Нейман. Суть этого способа состоит в том, что каждое последующее случайное число образуется возведением в квадрат предыдущего числа с отбрасыванием цифр младших и старших разрядов. Однако этот способ оказался ненадежным и от него вскоре отказались.

Линейный конгруэнтный генератор

Из известных процедур генерации последовательности псевдослучайных целых чисел наиболее часто применяется так называемый линейный конгруэнтный генератор. Этот генератор вырабатывает последовательность псевдослучайных чисел $y_1, y_2, \dots, y_{i-1}, y_i, \dots$, используя соотношение

$$y_i = (a \cdot y_{i-1} + b) \bmod m,$$

где $y_i - i$ -е (текущее) число последовательности;

y_{i-1} – предыдущее число последовательности;

a – множитель (коэффициент);

b – приращение;

m – модуль;

y_0 – порождающее число (исходное значение).

Данное уравнение генерирует псевдослучайные числа с периодом повторения, который зависит от выбираемых значений параметров a , b и m и может достигать значения m .

Например, при $y_0 = a = b = 7$, $m = 10$ последовательность выглядит так: 7, 6, 9, 0, 7, 6, 9, 0, 7, ... Длина периода полученной последовательности равна 4.

Значение модуля m должно быть велико, обычно m берется равным 2^n . Приращение b должно быть взаимно простым с m , т.е. НОД (b, m) = 1. Коэффициент a должен быть нечетным числом.

Конгруэнтные генераторы, работающие по алгоритму, предложенному Национальным бюро стандартов США, используются, в частности, в системах программирования. Эти генераторы имеют длину периода 2^{24} и обладают хорошими статистическими свойствами. Однако такая длина периода мала для криптографических применений. Кроме того, доказано, что последовательности, генерируемые конгруэнтными генераторами, не являются криптографически стойкими.

Регистр сдвига с линейной обратной связью

Существует способ генерации последовательностей псевдослучайных чисел на основе линейных рекуррентных соотношений.

Рассмотрим рекуррентные соотношения и их разностные уравнения:

$$\sum_{j=0}^k h_j a_{i+j} = 0,$$

$$a_{i+k} = -\sum_{j=0}^{k-1} h_j a_{i+j}, \quad (3)$$

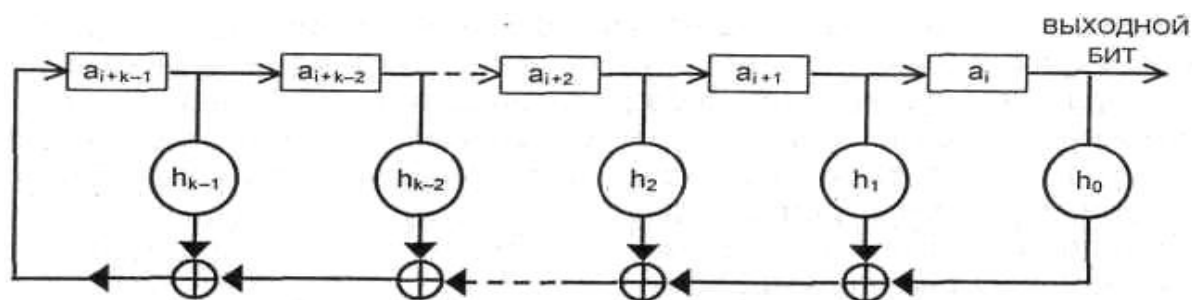
где $h_0 \neq 0$, $h_k = 1$ и каждое h_i принадлежит полю $\text{GF}(q)$.

Поле называется множество F , на котором определены операции сложения и умножения, удовлетворяющие ассоциативному, коммутативному и дистрибутивному законам, причём имеются как аддитивная (0), так и мультипликативная (1) единицы, каждый элемент имеет обратный элемент по сложению, кроме того, каждый элемент, кроме аддитивной единицы 0 имеет и обратный элемент по умножению.

Конечное поле $F(p)$ – это поле с конечным числом p элементов. В общем случае число элементов $p = q^n$, где q – некоторое простое число и $n \geq 1$. Такие конечные поля называют полями Галуа и обозначают $GF(q^n)$ или $GF(q)$ при $n = 1$.

Решением этих уравнений является последовательность элементов a_0, a_1, a_2, \dots поля $GF(q)$. Соотношение (3.1) определяет правило вычисления a_k по известным значениям величин $a_0, a_1, a_2, \dots, a_{k-1}$. Затем по известным значениям $a_0, a_1, a_2, \dots, a_k$ находят a_{k+1} и т.д. В результате по начальным значениям $a_0, a_1, a_2, \dots, a_{k-1}$ можно построить бесконечную последовательность, причем каждый ее последующий член определяется из k предыдущих. Последовательности такого вида легко реализуются на компьютере, при этом реализация получается особенно простой, если все h_i и a_i принимают значения 0 и 1 из поля $GF(2)$.

На рисунке 1.6 показана линейная последовательная переключательная схема, которая может быть использована для вычисления суммы (3) и, следовательно, для вычисления значения a_k по значениям k предыдущих членов последовательности.



Обозначения:



Сумматор по модулю 2



Цепь (отвод) с коэффициентом передачи h , $h = 0$ или 1



Запоминающая ячейка, хранящая a , т.е. на выходе ячейки $a = 0$ или $a = 1$

Рис. 6.6. Генератор с регистром сдвига

Исходные величины $a_0, a_1, a_2, \dots, a_{k-1}$ помещаются в разряды сдвигового регистра, последовательные сдвиги содержимого которого соответствуют вычислению последовательных символов, при этом выход после i -го сдвига равен a_i . Данное устройство

называют генератором последовательности чисел, построенным на базе сдвигового регистра с линейной обратной связью.

Конфигурацию обратных связей (отводов) h_i в генераторе со сдвиговым регистром определяет вид многочлена

$$h(x) = \sum_{j=0}^k h_j x^j,$$

где x – формальная переменная;

h_j – коэффициент при x^j , принимающий значение 0 или 1;

Другими словами, если у многочлена $h(x)$ коэффициент $h_j = 1$, это означает, что отвод h_j в схеме генератора присутствует, если же у многочлена $h(x)$ коэффициент $h_j = 0$, то отвод h_j в схеме генератора отсутствует.

В качестве $h(x)$ необходимо выбирать неприводимый примитивный многочлен

Многочлен $h(x)$ называется *неприводимым многочленом* степени m , если $h(x)$ нельзя разложить на сомножители – многочлены степени меньше m .

При таком выборе многочлена $h(x)$ со старшей степенью m генератор обеспечивает выдачу псевдослучайной последовательности двоичных чисел с максимально возможным периодом $2^m - 1$.

Рассмотрим в качестве примера трехразрядный сдвиговый регистр с линейной обратной связью (рисунок 1.7), построенный в соответствии с неприводимым примитивным многочленом

$$h(x) = x^3 + x^2 + 1,$$

где коэффициенты $h_3 = 1, h_2 = 1, h_1 = 0, h_0 = 1$.

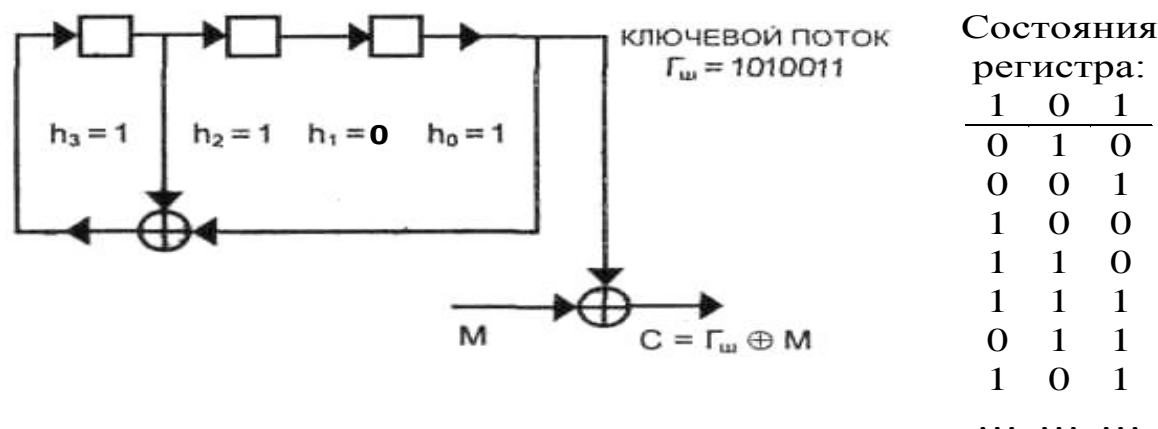


Рис. 6.7. Трехразрядный регистр сдвига с обратными связями
(генератор гаммы шифра Γ_w)

Пусть ключом является 101. Регистр начинает работать с этого состояния; последовательность состояний регистра приведена на рисунке 3.5. Регистр проходит через все семь ненулевых состояний и снова возвращается в свое исходное состояние 101. Это – наиболее длинный период данного регистра с линейной обратной связью. Такая последовательность называется *последовательностью максимальной длины* для сдвигового регистра (Maximal Length Shift Register Sequence – MLSRS). Питерсон и Уэлдон показали, что при любом целом m существует m -битовая последовательность MLSRS с периодом $2^m - 1$. В частности, при $m = 100$ последовательность будет иметь период $2^{100} - 1$ и не повторится 10^{16} лет при передаче ее по линии связи со скоростью 1 Мбит/с.

В нашем примере выходной последовательностью (гаммой шифра) $\Gamma_{\text{ш}}$ сдвигового регистра с обратной связью является последовательность 1010011, которая циклически повторяется. В этой последовательности имеется четыре единицы и три нуля, и их распределение настолько близко к равномерному, насколько это возможно в последовательности, имеющей длину 7. Если рассмотреть пары последовательных битов, то пары 10 и 01 появляются по два раза, а пары 00 и 11 – один раз, что опять оказывается настолько близким к равномерному распределению, насколько это возможно. В случае последовательности максимальной длины для m -разрядного регистра это свойство равномерности распространяется на тройки, четверки и т.д. битов, вплоть до m -битовых групп. Благодаря такой близости к равномерному распределению последовательности максимальной длины часто используются в качестве псевдослучайных последовательностей в криптографических системах, которые имитируют работу криптостойкой системы одноразового шифрования.

Хотя такая криптографическая система осуществляет имитацию заведомо криптостойкой системы одноразового шифрования, сама она не отличается стойкостью и может быть раскрыта за несколько секунд работы компьютера при условии наличия известного открытого текста.

Если отводы регистра с обратной связью зафиксированы, то для нахождения начального состояния регистра достаточно знать m битов открытого текста. Чтобы найти m битов ключевого потока, m битов известного открытого текста складывают по модулю 2 с соответствующими m битами шифртекста. Полученные m битов дают состояние сдвигового регистра с обратной связью в обратном направлении на некоторый момент времени. Затем, моделируя работу регистра назад, можно определить его исходное состояние.

1.2. Компьютерный практикум по классическим шифрам

Шифры замены

1. Основы шифрования.
2. Шифры однозначной замены.
3. Полиграммные шифры.
4. Омфонические шифры.
5. Полиалфавитные шифры.
6. Нерегулярные шифры.

1. Основы шифрования

Сущность шифрования методом замены заключается в следующем [9]. Пусть шифруются сообщения на русском языке и замене подлежит каждая буква этих сообщений. Тогда, букве **A** исходного алфавита сопоставляется некоторое множество символов (шифрозамен) **M_A**, **B** – **M_B**, ..., **Я** – **M_Я**. Шифрозамены выбираются таким образом, чтобы любые два множества (**M_I** и **M_J**, **i** ≠ **j**) не содержали одинаковых элементов (**M_I ∩ M_J = ∅**).

Таблица, приведенная на рис.1.8, является ключом шифра замены. Зная ее, можно осуществить как шифрование, так и расшифрование.

А	Б	...	Я
M _A	M _B	...	M _Я

Рис.6.8. Таблица шифрозамен

При шифровании каждая буква **A** открытого сообщения заменяется любым символом из множества **M_A**. Если в сообщении содержится несколько букв **A**, то каждая из них заменяется на любой символ из **M_A**. За счет этого с помощью одного ключа можно получить различные варианты шифрограммы для одного и того же открытого сообщения.

Так как множества **M_A**, **M_B**, ..., **M_Я** попарно не пересекаются, то по каждому символу шифрограммы можно однозначно определить, какому множеству он принадлежит, и,

следовательно, какую букву открытого сообщения он заменяет. Поэтому расшифрование возможно и открытое сообщение определяется единственным образом.

Приведенное выше описание сущности шифров замены относится ко всем их разновидностям за исключением полиалфавитных шифров, в которых для зашифрования разных символов исходного алфавита могут использоваться одинаковые шифрозамены (т.е. $M_i \cap M_j \neq \emptyset, i \neq j$).

Метод замены часто реализуется многими пользователями при работе на компьютере. Если по забывчивости не переключить на клавиатуре набор символов с латиницы на кириллицу, то вместо букв русского алфавита при вводе текста будут печататься буквы латинского алфавита («шифрозамены»).

Шифры замены можно разделить на следующие **подклассы** (разновидности):

- шифры однозначной замены (моноалфавитные, простые подстановочные). Количество шифрозамен для каждого символа исходного алфавита равно 1 ($|M_i| = 1$ для одного символа);



- полиграммные шифры. Аналогичен предыдущему за исключением того, что шифрозамене соответствует сразу блок символов исходного сообщения ($|M_i| = 1$ для блока символов);

- омофонические шифры (однозвучные, многозначной замены). Количество шифрозамен для отдельных символов исходного алфавита больше 1 ($|M_i| \geq 1$ для одного символа);

- полиалфавитные шифры (многоалфавитные). Состоит из нескольких шифров однозначной замены. Выбор варианта алфавита для зашифрования одного символа зависит от особенностей метода шифрования ($|M_i| > 1$ для одного символа);

- нерегулярные шифры. Шифрозамены состоят из разного количество символов.

Для записи исходных и зашифрованных сообщений используются строго определенные алфавиты. Алфавиты для записи исходных и зашифрованных сообщений могут отличаться. Символы обоих алфавитов могут быть представлены буквами, их сочетаниями, числами, рисунками, звуками, жестами и т.п. В качестве примера можно привести пляшущих

человечков из рассказа А. Конан Дойла () и рукопись рунического письма () из романа Ж. Верна «Путешествие к центру Земли».

2. Шифры однозначной замены

Максимальное количество ключей для любого шифра этого вида не превышает $n!$, где n – количество символов в алфавите. С увеличением числа n значение $n!$ растет очень быстро ($1! = 1$, $5! = 120$, $10! = 3628800$, $15! = 1307674368000$). При больших n для приближенного вычисления $n!$ можно воспользоваться формулой Стирлинга

$$n! \approx \sqrt{2\pi n} * \left(\frac{n}{e}\right)^n. \quad (4)$$

Шифр Цезаря. Данный шифр был придуман Гаем Юлием Цезарем и использовался им в своей переписке (1 век до н.э.). Применительно к русскому языку суть его состоит в следующем. Выписывается исходный алфавит (**А, Б, ..., Я**), затем под ним выписывается тот же алфавит, но с циклическим сдвигом на 3 буквы влево.

А	Б	В	Г	Д	Е	Ё	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ы	Ь	Ъ	Э	Ю	Я
Г	Д	Е	Ё	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ы	Ь	Ъ	Э	Ю	Я	А	Б	В

Рис.1.9. Таблица шифрозамен для шифра Цезаря

При зашифровке буква **А** заменяется буквой **Г**, **Б** - на **Д** и т. д. Так, например, исходное сообщение «АБРАМОВ» после шифрования будет выглядеть «ГДУГПСЕ». Получатель сообщения «ГДУГПСЕ» ищет эти буквы в нижней строке и по буквам над ними восстанавливает исходное сообщение «АБРАМОВ».

Ключом в шифре Цезаря является величина сдвига нижней строки алфавита. Количество ключей для всех модификаций данного шифра применительно к алфавиту русского языка равно 33. Возможны различные модификации шифра Цезаря, в частности лозунговый шифр.

Лозунговый шифр. Для данного шифра построение таблицы шифрозамен основано на лозунге (ключе) – легко запоминаемом слове. Вторая строка таблицы шифрозамен заполняется сначала словом-лозунгом (причем повторяющиеся буквы отбрасываются), а затем остальными буквами, не вошедшие в слово-лозунг, в алфавитном порядке. Например, если выбрано слово-лозунг «ДЯДИНА», то таблица имеет следующий вид.

А	Б	В	Г	Д	Е	Ё	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ы	Ь	Ъ	Э	Ю	Я
Д	Я	И	Н	А	Б	В	Г	Е	Ё	Ж	З	Й	К	Л	М	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ы	Ь	Ъ	Э	Ю

Рис. 1.9. Таблица шифрозамен для лозунгового шифра

При шифровании исходного сообщения «АБРАМОВ» по приведенному выше ключу шифрограмма будет выглядеть «ДЯПДКМИ».

В качестве лозунга рекомендуется выбирать фразу, в которой содержатся конечные буквы алфавита. В общем случае, количество вариантов нижней строки (применительно к русскому языку) составляет $33! (\geq 10^{35})$.

Полибианский квадрат. Шифр изобретен греческим государственным деятелем, полководцем и историком Полибием (III век до н.э.). Применительно к русскому алфавиту суть шифрования заключалась в следующем. В квадрат 6x6 выписываются буквы (необязательно в алфавитном порядке).

	1	2	3	4	5	6
1	А	Б	В	Г	Д	Е
2	Ё	Ж	З	И	Й	К
3	Л	М	Н	О	П	Р
4	С	Т	У	Ф	Х	Ц
5	Ч	Ш	Щ	Ъ	Ы	Ь
6	Э	Ю	Я	-	-	-

Рис. 1.10. Таблица шифрозамен для полибианского квадрата

Шифруемая буква заменяется на координаты квадрата (строка-столбец), в котором она записана. Например, если исходное сообщение «АБРАМОВ», то шифрограмма – «11 12 36 11 32 34 13». В Древней Греции сообщения передавались с помощью оптического телеграфа (с помощью факелов). Для каждой буквы сообщения вначале поднималось количество факелов, соответствующее номеру строки буквы, а затем номеру столбца.

Тюремный шифр. Эта звуковая разновидность полибианского квадрата была разработана заключенными. Система состояла из нескольких ударов, обозначающих строки и столбцы в таблице с буквами алфавита. Один удар, а потом еще два соответствовали строке 1 и столбцу 2, т.е. букве **Б**. Пауза служила разделителем между строками и

столбцами. Таким образом, зашифровать исходное сообщение «АБРАМОВ» можно следующим образом.

А	тук ___ тук
Б	тук ___ тук, тук
Р	тук, тук, тук ___ тук, тук, тук, тук, тук, тук
А	тук ___ тук
М	тук, тук, тук ___ тук, тук
О	тук, тук, тук ___ тук, тук, тук, тук
В	тук ___ тук, тук, тук

Рис. 6.11. Пример использования тюремного шифра

Шифрующая система Трисемуса (Тритемия). В 1508 г. аббат из Германии Иоганн Трисемус написал печатную работу по криптологии под названием «Полиграфия». В этой книге он впервые систематически описал применение шифрующих таблиц, заполненных алфавитом в случайном порядке. Для получения такого шифра замены обычно использовались таблица для записи букв алфавита и ключевое слово (или фраза). В таблицу сначала вписывалось по строкам ключевое слово, причем повторяющиеся буквы отбрасывались. Затем эта таблица дополнялась не вошедшими в нее буквами алфавита по порядку. На рис. 6.12 изображена таблица с ключевым словом «ДЯДИНА».

Д	Я	И	Н	А	Б
В	Г	Е	Ё	Ж	З
Й	К	Л	М	О	П
Р	С	Т	У	Ф	Х
Ц	Ч	Ш	Щ	Ы	Ь
Ъ	Э	Ю	-	-	-

Рис.6.12. Таблица шифрозамен для шифра Трисемуса

Каждая буква открытого сообщения заменяется буквой, расположенной под ней в том же столбце. Если буква находится в последней строке таблицы, то для ее шифрования берут самую верхнюю букву столбца. Например, исходное сообщение «АБРАМОВ», зашифрованное – «ЖЗЦЖУФЙ».

Шифр масонов. В XVIII в. масоны создали шифр, чтобы скрыть от общественности свои коммерческие сделки. Как поведали те, кто прежде состоял в рядах этого общества,

масоны пользовались способом засекречивания, весьма похожим на шифр розенкрейцеров. В «решетке» и в углах находятся точки, которыми заменяются буквы:

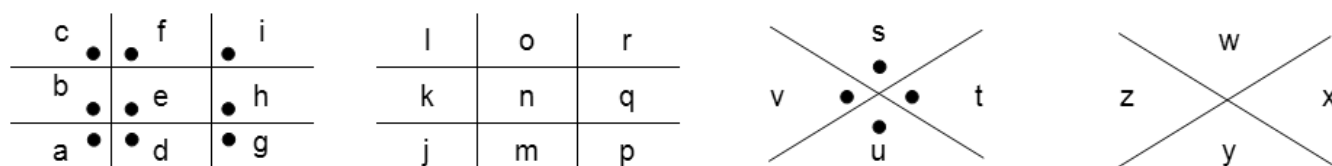


Рис. 1.13. Шифр масонов

Так как клятвы хранить тайну нарушались не раз, большинство Великих лож масонов в США больше не пользуются письменными шифрами, предпочитая передавать устные инструкции во время закрытых ритуалов.

С помощью шифра масонов можно легко расшифровать следующую фразу.

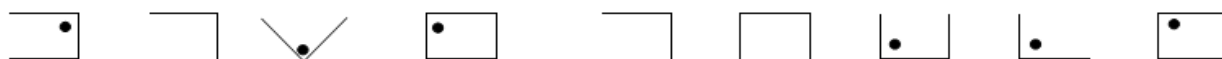


Рис. 1.14. Пример использования шифра масонов

Это первый уровень, на котором находятся все впервые вступившие в общество члены: Blue Lodge (рус. «Голубая (Синяя) ложа»).

Одним из существенных **недостатков шифров однозначной замены** является их легкая вскрываемость. При вскрытии шифрограмм используются различные приемы, которые даже при отсутствии мощных вычислительных средств позволяют добиться положительного результата. Один из таких приемов базируется на том, что в шифрограммах остается информация о частоте встречаемости букв исходного текста. Если в открытом сообщении часто встречается какая-либо буква, то в зашифрованном сообщении также часто будет встречаться соответствующий ей символ. Еще в 1412 году Шихаба ал-Калкашанди в своем труде «Субх ал-Ааша» привел таблицу частоты появления арабских букв в тексте на основе анализа текста Корана. Для разных языков мира существуют подобные таблицы. Так, например, для русского языка такая таблица выглядит следующим образом.

Таблица 6.11. Вероятности появления букв русского языка в текстах*

Буква (символ)	Вероятность	Буква	Вероятность	Буква	Вероятность	Буква	Вероятность
Пробел	0.146	Р	0.042	Я	0.017	Ж	0.007
О	0.094	Л	0.039	З	0.016	Ш	0.006
Е	0.071	В	0.038	Ы	0.015	Ц, Ю	0.005
А	0.069	К	0.029	Г	0.014	Щ	0.004
И	0.064	М	0.027	Ь, Б	0.013	Ф	0.003
Н	0.057	П	0.026	Ч	0.012	Э	0.002
Т	0.054	Д	0.024	Й	0.010	Ъ	0.001
С	0.046	У	0.023	Х	0.008		

*) В таблице приведены оценки вероятностей появления букв русского языка и пробела, полученные на основе анализа научно-технических и художественных текстов общим объемом более 1000000 символов.

Существуют подобные таблицы для пар букв (биграмм). Например, часто встречаемыми биграммами являются «то», «но», «ст», «по», «ен» и т.д. Другой прием вскрытия шифрограмм основан на исключении возможных сочетаний букв. Например, в текстах (если они написаны без орфографических ошибок) нельзя встретить сочетания «чя», «щы», «ъь» и т.п.

Для усложнения задачи вскрытия шифров однозначной замены еще в древности перед шифрованием из исходных сообщений исключали пробелы и/или гласные буквы. Другим способом, затрудняющим вскрытие, является шифрование **биграммами** (парами букв).

Полиграммные шифры

Полиграммные шифры замены - это шифры, в которых одна шифрозамена соответствует сразу нескольким символам исходного текста.

Биграммный шифр Порты. Шифр Порты, представленный им в виде таблицы, является первым известным биграммным шифром. Размер его таблицы составлял 20 x 20 ячеек; наверху горизонтально и слева вертикально записывался стандартный алфавит (в нем не было букв J, K, U, W, X и Z). В ячейках таблицы могли быть записаны любые числа, буквы или символы - сам Джованни Порты пользовался символами - при условии, что содержимое ни одной из ячеек не повторялось. Применительно к русскому языку таблица шифрозамен может выглядеть следующим образом.

	А	Б	В	Г	Д	Е (Ё)	Ж	З	И (Й)	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я
А	001	002	003	004	005	006	007	008	009	010	011	012	013	014	015	016	017	018	019	020	021	022	023	024	025	026	027	028	029	030	031
Б	032	033	034	035	036	037	038	039	040	041	042	043	044	045	046	047	048	049	050	051	052	053	054	055	056	057	058	059	060	061	062
В	063	064	065	066	067	068	069	070	071	072	073	074	075	076	077	078	079	080	081	082	083	084	085	086	087	088	089	090	091	092	093
Г	094	095	096	097	098	099	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124
Д	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155
Е (Ё)	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186
Ж	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217
З	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248
И (Й)	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279
К	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310
Л	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341
М	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372
Н	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403
О	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434
П	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465
Р	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496
С	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525	526	527
Т	528	529	530	531	532	533	534	535	536	537	538	539	540	541	542	543	544	545	546	547	548	549	550	551	552	553	554	555	556	557	558
У	559	560	561	562	563	564	565	566	567	568	569	570	571	572	573	574	575	576	577	578	579	580	581	582	583	584	585	586	587	588	589
Ф	590	591	592	593	594	595	596	597	598	599	600	601	602	603	604	605	606	607	608	609	610	611	612	613	614	615	616	617	618	619	620
Х	621	622	623	624	625	626	627	628	629	630	631	632	633	634	635	636	637	638	639	640	641	642	643	644	645	646	647	648	649	650	651
Ц	652	653	654	655	656	657	658	659	660	661	662	663	664	665	666	667	668	669	670	671	672	673	674	675	676	677	678	679	680	681	682
Ч	683	684	685	686	687	688	689	690	691	692	693	694	695	696	697	698	699	700	701	702	703	704	705	706	707	708	709	710	711	712	713
Ш	714	715	716	717	718	719	720	721	722	723	724	725	726	727	728	729	730	731	732	733	734	735	736	737	738	739	740	741	742	743	744
Щ	745	746	747	748	749	750	751	752	753	754	755	756	757	758	759	760	761	762	763	764	765	766	767	768	769	770	771	772	773	774	775
Ъ	776	777	778	779	780	781	782	783	784	785	786	787	788	789	790	791	792	793	794	795	796	797	798	799	800	801	802	803	804	805	806
Ы	807	808	809	810	811	812	813	814	815	816	817	818	819	820	821	822	823	824	825	826	827	828	829	830	831	832	833	834	835	836	837
Ь	838	839	840	841	842	843	844	845	846	847	848	849	850	851	852	853	854	855	856	857	858	859	860	861	862	863	864	865	866	867	868
Э	869	870	871	872	873	874	875	876	877	878	879	880	881	882	883	884	885	886	887	888	889	890	891	892	893	894	895	896	897	898	899
Ю	900	901	902	903	904	905	906	907	908	909	910	911	912	913	914	915	916	917	918	919	920	921	922	923	924	925	926	927	928	929	930
Я	931	932	933	934	935	936	937	938	939	940	941	942	943	944	945	946	947	948	949	950	951	952	953	954	955	956	957	958	959	960	961

Рис. 6.15. Таблица шифрозамен для шифра Порты

Шифрование выполняется парами букв исходного сообщения. Первая буква пары указывает на строку шифрозамены, вторая - на столбец. В случае нечетного количества букв в исходном сообщении к нему добавляется вспомогательный символ («пустой знак»). Например, исходное сообщение «АБ РА МО В», зашифрованное – «002 466 355 093». В качестве вспомогательного символа использована буква «Я».

Шифр Playfair (англ. «Честная игра»). В начале 1850-х гг. Чарлз Уитстон придумал так называемый «прямоугольный шифр». Леон Плейфер, близкий друг Уитстона, рассказал об этом шифре во время официального обеда в 1854 г. министру внутренних дел лорду Пальмерстону и принцу Альберту. А поскольку Плейфер был хорошо известен в военных и

дипломатических кругах, то за творением Уитстона навечно закрепилось название «шифр Плейфера».

Данный шифр стал первым буквенным биграммным шифром (в биграммной таблице Порты использовались символы, а не буквы). Он был предназначен для обеспечения секретности телеграфной связи и применялся британскими войсками в Англо-бурской и Первой мировой войнах. Им пользовалась также австралийская служба береговой охраны островов во время Второй мировой войны.

Шифр предусматривает шифрование пар символов (биграмм). Таким образом, этот шифр более устойчив к взлому по сравнению с шифром простой замены, так как затрудняется частотный анализ. Он может быть проведен, но не для 26 возможных символов (латинский алфавит), а для $26 \times 26 = 676$ возможных биграмм. Анализ частоты биграмм возможен, но является значительно более трудным и требует намного большего объема зашифрованного текста.

Для шифрования сообщения необходимо разбить его на биграммы (группы из двух символов), при этом, если в биграмме встретятся два одинаковых символа, то между ними добавляется заранее оговоренный вспомогательный символ (в оригинале – **X**, для русского алфавита - **Я**). Например, «зашифрованное сообщение» становится «за ши фр ов ан но ес о**Я** об ще ни е**Я**». Для формирования ключевой таблицы выбирается лозунг и далее она заполняется по правилам шифрующей системы Трисемуса. Например, лозунг «ДЯДИНА»

Д	Я	И	Н	А	Б
В	Г	Е	Ё	Ж	З
Й	К	Л	М	О	П
Р	С	Т	У	Ф	Х
Ц	Ч	Ш	Щ	Ы	Ь
Ъ	Э	Ю	-	1	2

Рис.6.16. Ключевая таблица для шифра Playfair

Затем, руководствуясь следующими правилами, выполняется зашифровывание пар символов исходного текста:

1. Если символы биграммы исходного текста встречаются в одной строке, то эти символы замещаются на символы, расположенные в ближайших столбцах справа от

соответствующих символов. Если символ является последним в строке, то он заменяется на первый символ этой же строки.

2. Если символы биграммы исходного текста встречаются в одном столбце, то они преобразуются в символы того же столбца, находящимися непосредственно под ними. Если символ является нижним в столбце, то он заменяется на первый символ этого же столбца.

3. Если символы биграммы исходного текста находятся в разных столбцах и разных строках, то они заменяются на символы, находящиеся в тех же строках, но соответствующие другим углам прямоугольника.

Пример шифрования.

- биграмма «за» формирует прямоугольник – заменяется на «жб»;
- биграмма «ши» находятся в одном столбце – заменяется на «юе»;
- биграмма «фр» находятся в одной строке – заменяется на «хс»;
- биграмма «ов» формирует прямоугольник – заменяется на «йж»;
- биграмма «ан» находятся в одной строке – заменяется на «ба»;
- биграмма «но» формирует прямоугольник – заменяется на «ам»;
- биграмма «ес» формирует прямоугольник – заменяется на «гт»;
- биграмма «оя» формирует прямоугольник – заменяется на «ка»;
- биграмма «об» формирует прямоугольник – заменяется на «па»;
- биграмма «ще» формирует прямоугольник – заменяется на «шё»;
- биграмма «ни» формирует прямоугольник – заменяется на «ан»;
- биграмма «ея» формирует прямоугольник – заменяется на «ги».

Шифрограмма – «жб юе хс йж ба ам гт ка па шё ан ги».

$$(18 * 0 + 17 * 1 + 15 * 17) \bmod 33 = 8 \quad (3)$$

АМО - 0 13 15

$$(6 * 0 + 27 * 13 + 1 * 15) \bmod 33 = 3 \quad (\Gamma)$$

$$(13 * 0 + 16 * 13 + 32 * 15) \bmod 33 = 28 \quad (\text{Б})$$

$$(28 * 0 + 17 * 13 + 15 * 15) \bmod 33 = 17 \quad (\text{Р})$$

Вяя - 2 32 32

$$(6 * 2 + 27 * 32 + 1 * 32) \bmod 33 = 17 \quad (\text{Р})$$

$$(13 * 2 + 16 * 32 + 32 * 32) \bmod 33 = 11 \quad (\text{К})$$

$$(28 * 2 + 17 * 32 + 15 * 32) \bmod 33 = 24 \quad (\text{Ч})$$

Для расшифрования тройки чисел шифрограммы необходимо умножить на матрицу расшифрования.

КЯЗ - 11 32 8

$$(2 * 11 + 26 * 32 + 17 * 8) \bmod 33 = 0 \quad (\text{А})$$

$$(26 * 11 + 20 * 32 + 4 * 8) \bmod 33 = 1 \quad (\text{Б})$$

$$(13 * 11 + 30 * 32 + 21 * 8) \bmod 33 = 17 \quad (\text{Р})$$

ГБР - 3 28 17

$$(2 * 3 + 26 * 28 + 17 * 17) \bmod 33 = 0 \quad (\text{А})$$

$$(26 * 3 + 20 * 28 + 4 * 17) \bmod 33 = 13 \quad (\text{М})$$

$$(13 * 3 + 30 * 28 + 21 * 17) \bmod 33 =$$

$$\begin{aligned}
 &15 && (O) \\
 &PKЧ - 17\ 11\ 24 \\
 &(2 * 17 + 26 * 11 + 17 * 24) \bmod 33 = \\
 &2 && (B) \\
 &(26 * 17 + 20 * 11 + 4 * 24) \bmod 33 = \\
 &32 && (я) \\
 &(13 * 17 + 30 * 11 + 21 * 24) \bmod 33 \\
 &= 32 && (я)
 \end{aligned}$$

В результате будет получен набор чисел «0 1 17 0 13 15 2 32 32», соответствующий исходному сообщению со вспомогательными символами «АБРАМОВяя».

¹**Обратная матрица** - матрица A^{-1} , при умножении на которую, исходная матрица A дает в результате единичную матрицу E .

Омофонические шифры

Другое направление повышения стойкости шифров замены состоит в том, чтобы каждое множество шифробозначений M_i содержало более одного элемента. При использовании такого шифра одну и ту же букву (если она встречается несколько раз в сообщении) заменяют на разные шифрозамены из M_i . Это позволяет скрыть истинную частоту встречаемости букв открытого сообщения.

Система омофонов. В 1401 г. Симеоне де Крема стал использовать таблицы омофонов для сокрытия частоты появления гласных букв в тексте при помощи более чем одной шифрозамены. Такие шифры позже стали называться **шифрами многозначной замены** или **омофонами**². Они получили развитие в XV веке. В книге «Трактат о шифрах» Леона Баттисты Альберти (итальянский ученый, архитектор, теоретик искусства, секретарь папы Климентия XII), опубликованной в 1466 г., приводится описание шифра замены, в котором каждой букве ставится в соответствие несколько эквивалентов, число которых пропорционально частоте встречаемости буквы в открытом тексте. Так, если ориентироваться на то число шифрозамен для буквы **О** должно составлять 94, для буквы **Е** – 71 и т.д. При этом каждая шифрозамена должна состоять из 3 цифр и их общее количество равно 1000. На рис.1.18 представлен фрагмент таблицы шифрозамен.

№ п/п	Пробел	А	Б	В	...	М	...	О	...	Р	...	Я
1	012	311	128	175	...	037	...	248	...	064	...	266
2	042	357	950	194	...	149	...	267	...	189	...	333
...
13	278	495	990	199	...	349	...	303	...	374	...	749
...
17	342	519		427	...	760	...	306	...	469	...	845
...
27	437	637		524	...	777	...	432	...	554		
...		
38	457	678		644				824	...	721		
...		
42	628	776						828	...	954		
...		
69	681	901						886				
...				
94	974							903				
...	...											
146	976											

Рис.6.18. Фрагмент таблицы шифрозамен для системы омофонов

При шифровании символ исходного сообщения заменяется на любую шифрозамену из своего столбца. Если символ встречается повторно, то, как правило, используют разные шифрозамены. Например, исходное сообщение «АБРАМОВ» после шифрования может выглядеть «357 990 374 678 037 828 175».

Книжный шифр. Заметным вкладом греческого ученого Энея Тактика в криптографию является предложенный им так называемый книжный шифр, описанный в сочинении «Об обороне укрепленных мест». Эней предложил прокалывать малозаметные дырки в книге или в другом документе над буквами секретного сообщения. Интересно отметить, что в первой мировой войне германские шпионы использовали аналогичный шифр, заменив дырки на точки, наносимые симпатическими чернилами³ на буквы газетного текста.

После первой мировой войны книжный шифр приобрел иной вид. Шифрозамена для каждой буквы определялась набором цифр, которые указывали на номер страницы, строки и позиции в строке. Количество книг, изданных за всю историю человечества, является величиной ограниченной (по крайней мере, явно меньше, чем 15!). Однако отсутствие полной электронной базы по изданиям делает процедуру вскрытия шифрограмм почти не выполнимой. В связи с этим книжный шифр относят к категории совершенных.

Вариантные шифры. Вариантные шифры напоминают полибианский квадрат, но для каждой строки и столбца используется по два буквенных идентификатора. В квадрат (прямоугольник) шифрозамен вначале записывается ключевое слово без повторяющихся букв, а затем дополняется не вошедшими в него буквами по порядку следования в алфавите.

Каждой строке и столбцу квадрата ставится в соответствие по две буквы алфавита. Буквы для идентификации строк и столбцов не должны повторяться.

	Й	У	Е	Г	Щ	Х
	Ц	К	Н	Ш	З	Ъ
ФЫ	Д	Я	И	Н	А	Б
ВА	В	Г	Е	Ё	Ж	З
ПР	Й	К	Л	М	О	П
ОЛ	Р	С	Т	У	Ф	Х
ДЖ	Ц	Ч	Ш	Щ	Ы	Ь
ЭЯ	Ъ	Э	Ю	-	-	-

Рис. 6.19. Пример таблицы шифрозамен вариантного шифра с ключевым словом «ДЯДИНА»

Комбинации букв-идентификаторов строки и столбца дают по восемь шифрозамен для каждой буквы исходного текста. Например, для буквы Д возможны шифрозамены: **ФЙ, ЙФ, ФЦ, ЦФ, ЫЙ, ЙЫ, ЫЦ** и **ЦЫ**. Для таблицы шифрозамен, приведенной исходное сообщение «АБРАМОВ» может быть зашифровано как «ЫЗ ЫХ ОЦ ЗФ ГР РЩ АЙ».

²**Омофоны** (греч. homos - одинаковый и phone - звук) - слова, которые звучат одинаково, но пишутся по-разному и имеют разное значение.

³**Симпатические (невидимые) чернила** — чернила, записи которыми являются изначально невидимыми и становятся видимыми только при определенных условиях (нагрев, освещение, химический проявитель и т. д.).

Полиалфавитные шифры

Напомним, что полиалфавитные шифры состоят из нескольких шифров однозначной замены и отличаются друг от друга способом выбор варианта алфавита для зашифрования одного символа.

Диск Альберти. В «Трактате о шифрах» Альберти приводит первое точное описание многоалфавитного шифра на основе шифровального диска.

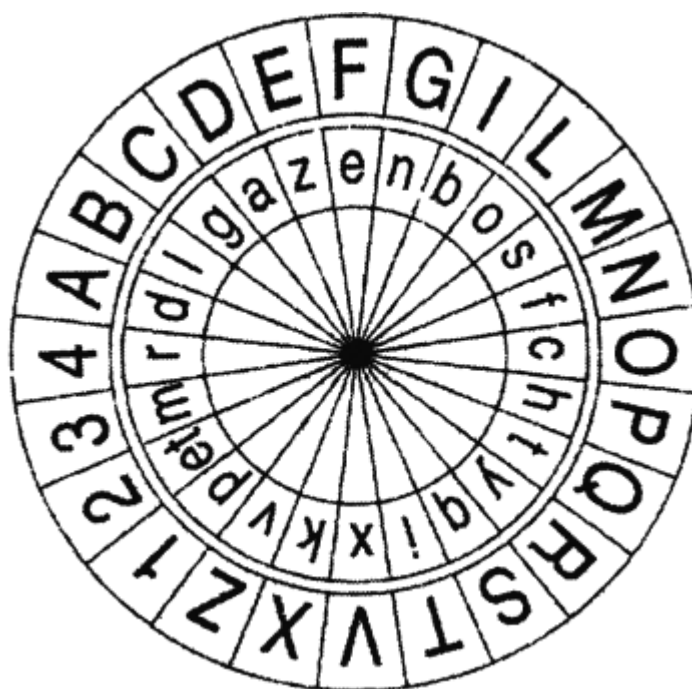


Рис.6.20. Диск Альберти

Он состоял из двух дисков – внешнего неподвижного (на нем были нанесены буквы в алфавитном порядке и цифры 1, 2, 3, 4) и подвижного внутреннего диска на котором буквы были переставлены. Процесс шифрования заключался в нахождении буквы открытого текста на внешнем диске и замене ее на букву с внутреннего диска, стоящую под ней. После этого внутренний диск сдвигался на одну позицию и шифрование второй буквы производилось уже по новому шифралфавиту. Ключом данного шифра являлся порядок расположения букв на внутреннем диске и его начальное положение относительно внешнего диска.

Таблица Трисемуса. Одним из шифров, придуманных немецким аббатом Трисемусом, стал многоалфавитный шифр, основанный на так называемой «таблице Трисемуса» - таблице со стороной равной n , где n – количество символов в алфавите. В первой строке матрицы записываются буквы в порядке их очередности в алфавите, во второй – та же последовательность букв, но с циклическим сдвигом на одну позицию влево, в третьей – с циклическим сдвигом на две позиции влево и т.д.

А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я
Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я	А
В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я	А	Б
Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я	А	Б	В
Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я	А	Б	В	Г
Е	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я	А	Б	В	Г	Д
Ж	З	И	Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я	А	Б	В	Г	Д	Е
З	И	Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я	А	Б	В	Г	Д	Е	Ж
И	Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я	А	Б	В	Г	Д	Е	Ж	З
Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я	А	Б	В	Г	Д	Е	Ж	З	И
К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я	А	Б	В	Г	Д	Е	Ж	З	И	Й
Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я	А	Б	В	Г	Д	Е	Ж	З	И	Й	К
М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я	А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л
Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я	А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М
О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я	А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н
П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я	А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О
Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я	А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П
С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я	А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П	Р
Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я	А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С
У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я	А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С	Т
Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я	А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С	Т	У
Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я	А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф
Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я	А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х
Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я	А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц
Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я	А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч
Щ	Ъ	Ы	Ь	Э	Ю	Я	А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш
Ъ	Ы	Ь	Э	Ю	Я	А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ
Ы	Ь	Э	Ю	Я	А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ
Ь	Э	Ю	Я	А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы
Э	Ю	Я	А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь
Ю	Я	А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э
Я	А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю

Рис. 6.21. Таблица Трисемуса

Здесь первая строка является одновременно и строкой букв открытого текста. Первая буква текста шифруется по первой строке, вторая буква по второй и так далее после использования последней строки вновь возвращаются к первой. Так сообщение «АБРАМОВ» приобретет вид «АВТГРУИ».

Система шифрования Виженера. В 1586 г. французский дипломат Блез Виженер представил перед комиссией Генриха III описание простого, но довольно стойкого шифра, в основе которого лежит таблица Трисемуса.

Перед шифрованием выбирается ключ из символов алфавита. Сама процедура шифрования заключается в следующем. По i -ому символу открытого сообщения в первой строке определяется столбец, а по i -ому символу ключа в крайнем левом столбце – строка. На пересечении строки и столбца будет находиться i -ый символ, помещаемый в шифрограмму. Если длина ключа меньше сообщения, то он используется повторно.

Например, исходное сообщение «АБРАМОВ», ключ – «ДЯДИНА», шифрограмма – «ДАФИЩОЖ».

Справедливости ради, следует отметить, что авторство данного шифра принадлежит итальянцу Джованни Батиста Беллазо, который описал его в 1553 г. История «проигнорировала важный факт и назвала шифр именем Виженера, несмотря на то, что он ничего не сделал для его создания». Беллазо предложил называть секретное слово или фразу **паролем** (ит. password; фр. parole - слово).

В 1863 г. Фридрих Касиски опубликовал алгоритм атаки на этот шифр, хотя известны случаи его взлома шифра некоторыми опытными криптоаналитиками ещё в XVI веке. Несмотря на это шифр Виженера имел репутацию исключительно стойкого к «ручному» взлому еще долгое время. Так, известный писатель и математик Чарльз Лютвидж Доджсон (Льюис Кэрролл) в своей статье «Алфавитный шифр», опубликованной в детском журнале в 1868 г., назвал шифр Виженера невзламываемым. В 1917 году научно-популярный журнал «Scientific American» также отозвался о шифре Виженера, как о неподдающемся взлому.

Роторные машины. Идеи Альберти и Беллазо использовались при создании электромеханических роторных машин первой половины XX века. Некоторые из них использовались в разных странах вплоть до 1980-х годов. Большинство использовало понятие ротора - механического колеса, используемого для выполнения подстановки. Наиболее известной из роторных машин является немецкая машина времен Второй мировой войны «Энигма».

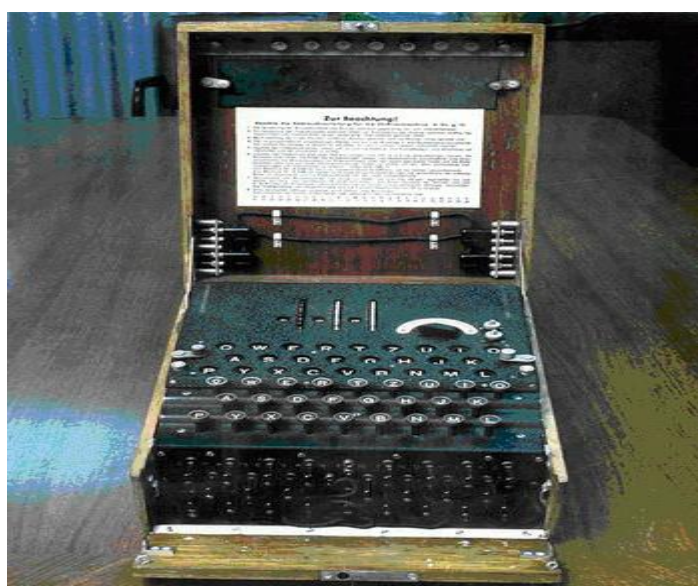


Рис. 6.22. Шифровальная машина Энигма

Роторная машина, включающая клавиатуру и набор роторов, реализует вариант шифра Виженера. Каждый ротор представляет собой произвольное размещение алфавита, имеет 26 позиций (применительно к латинскому алфавиту) и выполняет простую подстановку. Например, ротор может быть использован для замены **A** на **F**, **B** на **U**, **C** на **I** и так далее.

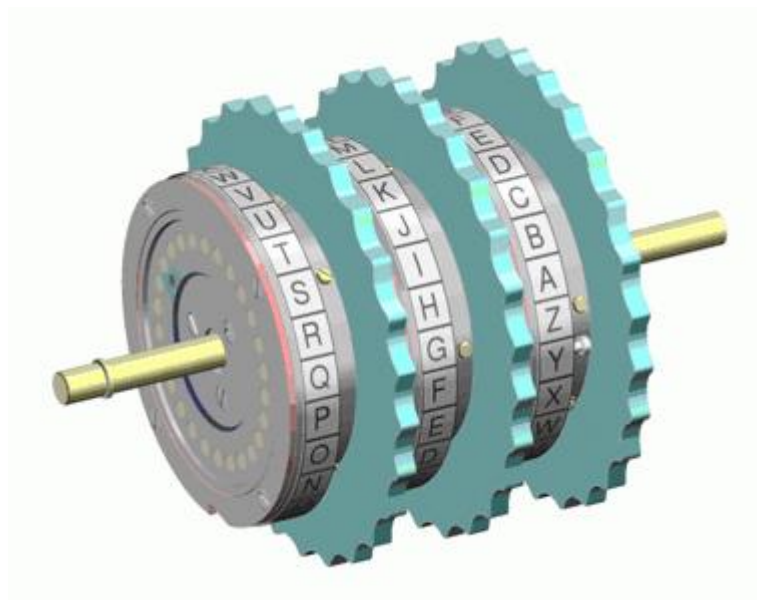


Рис. 6.23. Три последовательно соединённых ротора

Выходные штыри одного ротора соединены с входными штырями следующего ротора и при нажатии символа исходного сообщения на клавиатуре замыкали электрическую цепь, в результате чего загоралась лампочка с символом шифрозамены.

Шифрующее действие «Энигмы» показано для двух последовательно нажатых клавиш - ток течёт через роторы, «отражается» от рефлектора, затем снова через роторы.

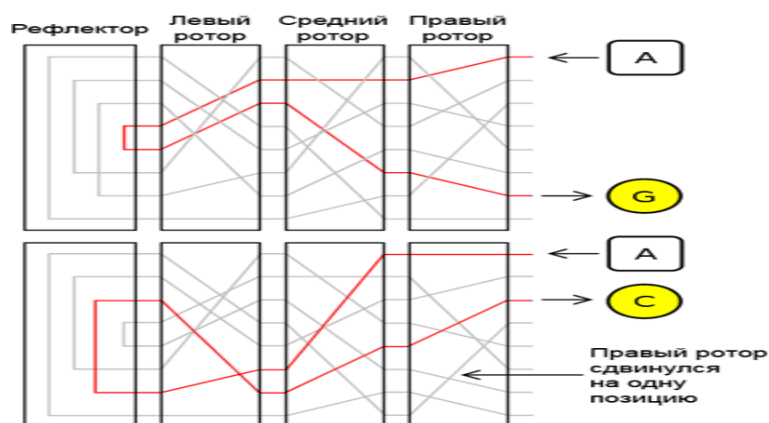


Рис. 6.24. Схема шифрования

Примечание. Серыми линиями показаны другие возможные электрические цепи внутри каждого ротора. Буква А шифруется по-разному при последовательных нажатиях одной клавиши, сначала в G, затем в С. Сигнал идет по другому маршруту за счёт поворота ротора после нажатия предыдущей буквы исходного сообщения.

В некоторых шифрах в самой шифрограмме могут содержаться символы, предписывающие использование того или иного алфавита.

Шифры Тени. Главными развлечениями для американцев тридцатых годов XX века были бульварное чтение и радио. Для раскрутки своих книжек издательство Street & Smith проспонсировало радиопередачу, ведущим в которой был Тень (англ. Shadow), загадочный рассказчик со зловещим голосом, который в начале каждого выпуска заявлял: «Кто знает, что за зло прячется в сердцах людей? Тень знает!». Успех радиопередачи подтолкнул издательство к решению начать выпускать серию книг, в которой главным героем был бы Тень. Свои услуги предложил Уолтер Гибсон, большой любитель фокусов и головоломок. Под псевдонимом Максвелл Грант он принялся писать роман за романом, да с такой скоростью, что за свою жизнь написал почти 300 книжек о грозе тех, кто нечист помыслами. В новелле «Цепочка смерти» супергерой воспользовался так называемым кодом направления, хотя на самом деле он действует скорее как шифр, чем как код:

a	b	c	d	e	f	g	h	i	j	k	l	m	
n	o	p	q	r	s	t	u	v	w	x	y	z	
пробел		1	2	3	4								

Рис. 6.25. Таблица шифрозамен и управляющих символов

Управляющие символы в последней строке таблицы служат для изменения кода (выбора шифралфавита) для зашифрования/дешифрования. Линии внутри каждого кружка фактически являются стрелками, подсказывающими адресату, как держать лист бумаги. Символ 1 означает, что лист надо держать как обычно: верх и низ расположены на своих местах, а сообщение читается слева направо. Символ 2 требует поворота на 90° вправо, а символ 3 указывает, что лист бумаги следует перевернуть вверх ногами. Символ 4 обозначает поворот на 90° влево.

Эти дополнительные символы могут появляться перед любой строчкой текста, а также в ее середине.

Из нижеприведенного примера можно узнать настоящие имя и фамилию супергероя.

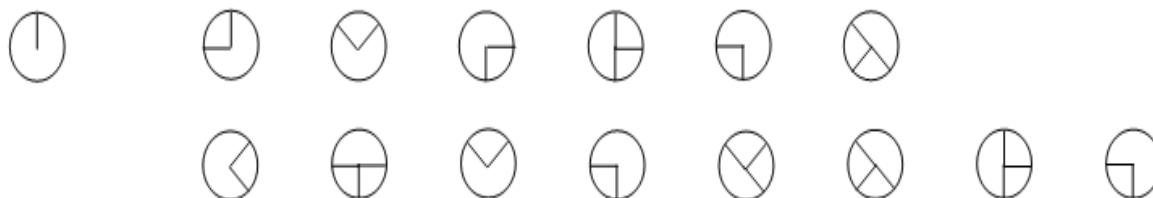


Рис. 6.26. Настоящие имя и фамилия Тени

Согласно первому управляющему символу, лист следует держать обычным образом, не поворачивая, и после замены буквы образуют «Lamont Cranston» (Ламонт Крэнстон).

Нерегулярные шифры

Еще одним направлением повышения стойкости шифров замены заключается в использовании нерегулярных шифров. В приведенных выше шифрах (**регулярных**) шифрозамены состоят из строго определенного количества символов (букв, цифр, графических элементов и т.д.) или в шифрограмме они отделяются друг от друга специальными символами (пробелом, точкой, запятой и т.д.). В нерегулярных шифрах шифрозамены состоят из разного количества символов и записываются в шифрограмме в подряд (без выделения друг от друга), что значительно затрудняет криптоанализ.

Совмещенный шифр (совмещенная таблица). Данный шифр применялся еще семейством Ардженти - криптологами, разрабатывавшими шифры для Папы Римского в XVI в. В XX столетии этим способом пользовались коммунисты в ходе гражданской войны в Испании. В начале войны противники фашизма в Испании контролировали большинство крупных городов и защищали свою связь, включая радиопередачи, с помощью различных методов шифрования, в том числе совмещенных шифров.

Вариант коммунистов получил название «совмещенный» из-за необычного использования одно- и двухцифровых шифрозамен, благодаря чему сообщение приобретало дополнительную защиту от потенциального дешифровальщика. Некоторые буквы шифровались одной цифрой, другие же - парой цифр. При этом криптоаналитик

противника совершенно не представлял, где в перехваченных сообщениях находятся одноцифровые, а где двухцифровые шифрозамены.

Таблица шифрозамен состоит из 10 столбцов с нумерацией 0, 9, 8, 7, 6, 5, 4, 3, 2 и 1. В начальную строку вписывается ключевое слово без повторяющихся букв. В последующие строки вписываются по десять не вошедших в него букв по порядку следования в алфавите. Строки, за исключением начальной, нумеруются по порядку, начиная с 1.

	0	9	8	7	6	5	4	3	2	1
	Д	Я	И	Н	А					
1	Б	В	Г	Е	Ё	Ж	З	Й	К	Л
2	М	О	П	Р	С	Т	У	Ф	Х	Ц
3	Ч	Ш	Щ	Ы	Ь	Ъ	Э	Ю	-	-

Рис. 6.27. Пример таблицы шифрозамен совмещенного шифра с ключевым словом «ДЯДИНА»

При шифровании буквы исходного сообщения, входящие в ключевое слово, заменяются на одну цифру (номер столбца), остальные – двумя (номера строки и столбца). Например, для приведенной выше таблицы шифрозамен исходное сообщение «АБРАМОВ» будет зашифровано как «610276202919».

При получении шифрограммы адресат знает, что когда появляются цифры 1, 2 или 3, с ними обязательно связана еще одна цифра, поскольку они представляют собой цифровую пару. Так что 35 - это, несомненно, пара, а 53 - нет, ведь в таблице нет строки с номером 5. Перехват такого сообщения третьей стороной даст ей всего лишь ряд цифр, потому что криптоаналитик противника не имеет ни малейшего представления, какие цифры одиночные, а какие входят в состав пар.

Компьютерный практикум

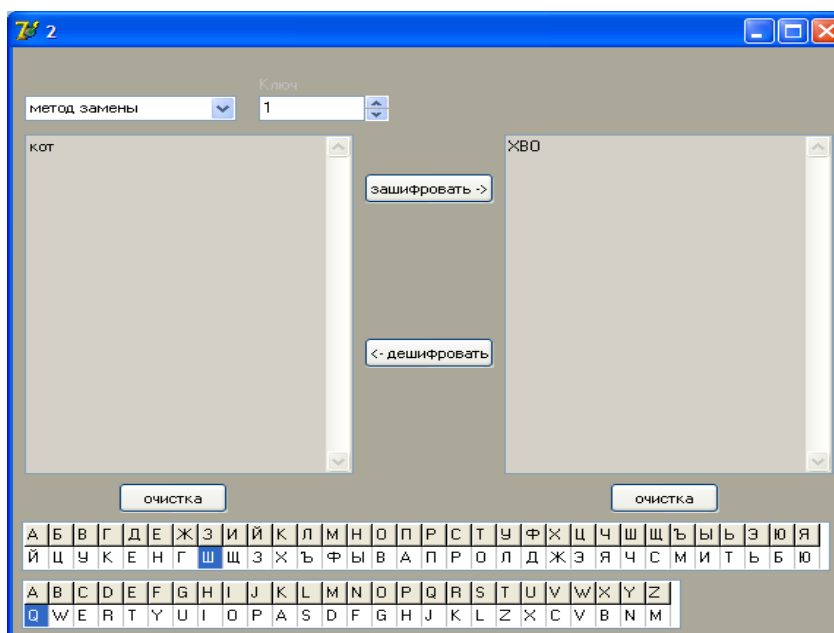
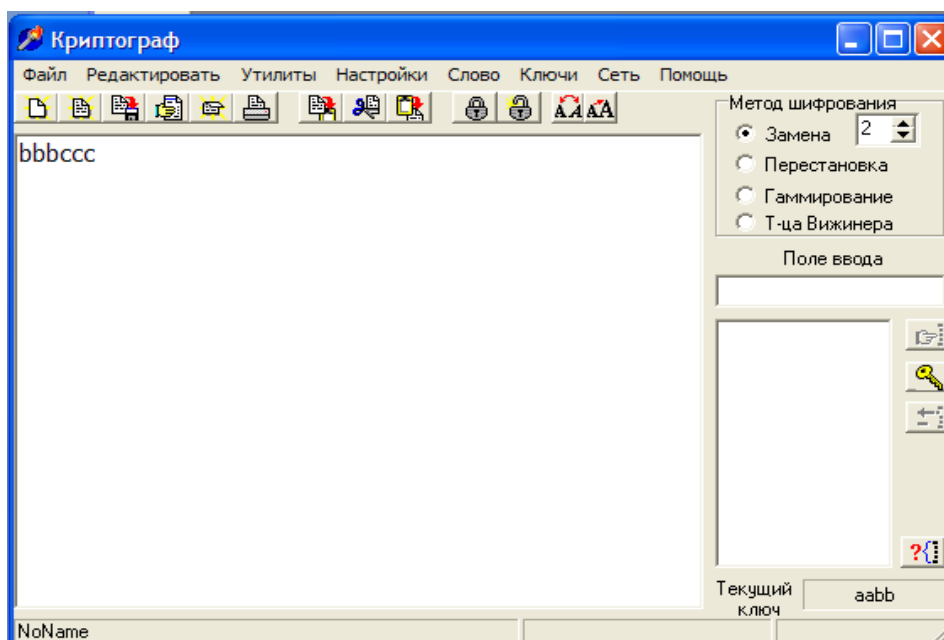
Задание 1

В данной работе изучаются три программы:

- Криптограф
- Шифр Цезаря

Криптограф

В программе реализуется шифр замены, перестановки, гаммирования и шифрование на основе таблицы Вижинера. В поле метод шифрования можно выбрать ключ, на который сдвигается шифруемое сообщение. Специальными кнопками можно выбрать шифрование или дешифрование.



В данной программе шифр замены имеет специальную таблицу для шифрования, поэтому ключом является таблица.

Задание 2

Перестановка
 Главное Дешифровать Помощь О программе

Впишите D: 4 Ключ подходит:

1	2	3	4
1	4	2	3

Начальная матрица перестановки

Я		у	ч
у	с	ь	
в		Т	У
С	У	Р	е
.	Я		у

Открытый текст 17 символов

Я учусь в ТУСУРе.

Шифрованный текст 20 символов

Яуч ьсвТУ СРеУ. уЯ

Результирующая матрица перестановки

Я	у	ч	
у	ь		с
в	Т	У	
С	Р	е	У
.		у	Я

Перестановка
 Главное Дешифровать Помощь О программе

Впишите D: 4 Ключ подходит:

1	2	3	4
1	4	2	3

Начальная матрица перестановки

Я	у	ч	
у	ь		с
в	Т	У	
С	Р	е	У
.		у	Я

Открытый текст 20 символов

Я учусь в ТУСУРе.Я у

Шифрованный текст 20 символов

Яуч ьсвТУ СРеУ. уЯ

Результирующая матрица перестановки

Я		у	ч
у	с	ь	
в		Т	У
С	У	Р	е
.	Я		у

6.3. Задания на самостоятельную работу по классическим шифрам

Задания на криптоанализ классических шифров

Шифр столбцовой перестановки

При решении заданий на криптоанализ шифров перестановки необходимо восстановить начальный порядок следования букв текста. Для этого используется анализ совместимости символов, в чем может помочь таблица сочетаемости.

Таблица 1. Сочетаемость букв русского языка

Г	С	Слева		Справа	Г	С
3	97	л, д, к, т, в, р, н	А	л, н, с, т, р, в, к, м	12	88
80	20	я, е, у, и, а, о	Б	о, ы, е, а, р, у	81	19
68	32	я, т, а, е, и, о	В	о, а, и, ы, с, н, л, р	60	40
78	22	р, у, а, и, е, о	Г	о, а, р, л, и, в	69	31
72	28	р, я, у, а, и, е, о	Д	е, а, и, о, н, у, р, в	68	32
19	81	м, и, л, д, т, р, н	Е	н, т, р, с, л, в, м, и	12	88
83	17	р, е, и, а, у, о	Ж	е, и, д, а, н	71	29
89	11	о, е, а, и	З	а, н, в, о, м, д	51	49
27	73	р, т, м, и, о, л, н	И	с, н, в, и, е, м, к, з	25	75
55	45	ь, в, е, о, а, и, с	К	о, а, и, р, у, т, л, е	73	27
77	23	г, в, ы, и, е, о, а	Л	и, е, о, а, ь, я, ю, у	75	25
80	20	я, ы, а, и, е, о	М	и, е, о, у, а, н, п, ы	73	27
55	45	д, ь, н, о, а, и, е	Н	о, а, и, е, ы, н, у	80	20

11	89	р, п, к, в, т, н	О	в, с, т, р, и, д, н, м	15	85
65	35	в, с, у, а, и, е, о	П	о, р, е, а, у, и, л	68	32
55	45	и, к, т, а, п, о, е	Р	а, е, о, и, у, я, ы, н	80	20
69	31	с, т, в, а, е, и, о	С	т, к, о, я, е, ь, с, н	32	68
57	43	ч, у, и, а, е, о, с	Т	о, а, е, и, ь, в, р, с	63	37
15	85	п, т, к, д, н, м, р	У	т, п, с, д, н, ю, ж	16	84
70	30	н, а, е, о, и	Ф	и, е, о, а, е, о, а	81	19
90	10	у, е, о, а, ы, и	Х	о, и, с, н, в, п, р	43	57
69	31	е, ю, н, а, и	Ц	и, е, а, ы	93	7
82	18	е, а, у, и, о	Ч	е, и, т, н	66	34
67	33	ь, у, ы, е, о, а, и, в	Ш	е, и, н, а, о, л	68	32
84	16	е, б, а, я, ю	Щ	е, и, а	97	3
0	100	м, р, т, с, б, в, н	Ы	л, х, е, м, и, в, с, н	56	44
0	100	н, с, т, л	Ь	н, к, в, п, с, е, о, и	24	76
14	86	с, ы, м, л, д, т, р, н	Э	н, т, р, с, к	0	100
58	42	ь, о, а, и, л, у	Ю	д, т, щ, ц, н, п	11	89
43	57	о, н, р, л, а, и, с	Я	в, с, т, п, д, к, м, л	16	84

Таблица 2. Сочетаемость букв английского языка

Г	С	Слева		Справа	Г	С
19	81	l,c,d,m,n,s,w,t,r,e,h	A	n,t,s,r,l,d,c,m	6	94
55	45	y,b,n,t,u,d,o,s,a,e	B	e,l,u,o,a,y,b,r	70	30
61	39	u,o,s,n,a,i,l,e	C	h,o,e,a,i,t,r,l,k	59	41
52	48	r,i,l,a,n,e	D	e,i,t,a,o,u	54	46
8	92	c,b,e,m,v,d,s,l,n,t,r,h	E	r,d,s,n,a,t,m,e,c,o	21	79
69	31	s,n,f,d,a,i,e,o	F	t,o,e,i,a,r,f,u	52	48
36	64	o,d,u,r,i,e,a,n	G	e.h.o.r.a.t.f.w.i.s	42	58
7	93	g,e,w,s,c,t	H	e,a,i,o	90	10
13	87	f,m,w,e,n,l,d,s,r,h,t	I	n,t,s,o,c,r,e,m,a,l	17	83
28	72	y,w,t,s,n,e,c,b,a,c	J	u,o,a,e,m,w	88	12
53	47	y,u,i,n,a,r,o,c	K	e,i,n,a,t,s	68	32
52	48	m,p,t,i,b,u,o,e,l,a	L	e,i,y,o,a,d,u	65	35
69	31	s,d,m,r,i,a,o,e	M	e,a,o,i,p,m	71	29
89	11	u,e,o,a,i	N	d,t,g,e,a,s,o,i,c	32	68
21	79	o,d,l,p,h,n,e,c,f,s,i,r,t	O	n,f,r,u,t,m,l,s,w,o	18	82
47	53	r,l,t,n,i,p,m,a,o,u,e,s	P	o,e,a,r,l,u,p,t,i,s	59	41
20	80	o,n,l,e,d,r,s	Q	u	10	0
70	30	p,i,u,t,a,o,e	R	e,o,a,t,i,s,y	61	39
48	52	d,t,o,u,r,n,s,i,a,e	S	t,e,o,i,s,a,h,p,u	41	59
43	57	u,o,d,t,f,e,i,n,s,a	T	h,i,o,e,a,t,r	38	62
35	65	p,f,t,l,b,d,s,o	И	n,s,t,r,l,p,b,c	8	92
88	12	r,u,o,a,i,e	V	e,i,o,a	99	1
48	52	g,d,y,n,s,t,o,e	W	a,h,i,e,o,n	80	20
95	5	u,n,i,e	X	p,t,i,a,u,c,k,o	38	62
24	76	b,n,a,t,e,r,l	Y	a,o,s,t,w,h,i,e,d,m	38	62
88	12	o,n,a,i	Z	e,i,w	86	14

При анализе сочетаемости букв друг с другом следует иметь в виду зависимость появления букв в открытом тексте от значительного числа предшествующих букв. Для анализа этих закономерностей используют понятие условной вероятности.

Систематически вопрос о зависимости букв алфавита в открытом тексте от предыдущих букв исследовался известным русским математиком А.А.Марковым (1856 — 1922). Он доказал, что появления букв в открытом тексте нельзя считать независимыми друг от друга. В связи с этим А. А. Марковым отмечена еще одна устойчивая закономерность открытых текстов, связанная с чередованием гласных и согласных букв. Им были подсчитаны частоты встречаемости биграмм вида гласная-гласная (g,g), гласная-согласная (g,c), согласная-гласная (c,g), согласная-согласная (c,c) в русском тексте длиной в 10^5 знаков. Результаты подсчета отражены в следующей таблице:

Таблица 1.12. Чередование гласных и согласных

	Г	С	Всего
Г	6588	38310	44898
С	38296	16806	55102

Пример решения:

Дан шифр-текст: СВПООЗЛУЙЬСТЬ_ЕДПСОКОКАЙЗО

Текст содержит 25 символов, что позволяет записать его в квадратную матрицу 5×5 . Известно, что шифрование производилось по столбцам, следовательно, расшифрование следует проводить, меняя порядок столбцов.

С	В	П	О	О
З	Л	У	Й	Ь
С	Т	Ь	–	Е
Д	П	С	О	К
К	А	Й	З	О

Необходимо произвести анализ совместимости символов (Таблица сочетаемости букв русского и английского алфавита, а также таблицы частот биграмм представлена выше). В первом и третьем столбце сочетание СП является крайне маловероятным для русского языка, следовательно, такая последовательность столбцов быть не может. Рассмотрим другие запрещенные и маловероятные сочетания букв: ВП (2,3 столбцы), ПС (3,1 столбцы), ПВ (3,2 столбцы). Перебрав их все, получаем наиболее вероятные сочетания биграмм по столбцам:

В	О	С	П	О
Л	Ь	З	У	Й

Т	Е	С	Ь	–
П	О	Д	С	К
А	З	К	О	Й

Получаем осмысленный текст: ВОСПОЛЬЗУЙТЕСЬ_ПОДСКАЗКОЙ

Задание: Расшифровать фразу, зашифрованную столбцовой перестановкой.

1. ОКЕСНВРП_ЫРЕАДЕЫН_В_РСИКО
2. ДСЛИЕЗТЕА_Ь_ЛЫЮВМИ_АОЧХК
3. НМВИАИ_НЕВЕ_СМСТУОРДИАНКМ
4. ЕДСЗЫНДЕ_МУБД_УЭ_КРЗЕМНАЫ
5. СОНРЧОУО_ХДТ_ИЕИ_ВЗКАТРРИ
6. _ОНКА_БНЫЕЦВЛЕ_К_ТГОАНЕИР
7. НЗМАЕЕАА_Г_НОТВОССОТЬЯАЛС
8. РППОЕААДТВЛ_ЕБЬЛНЫЕ_ПА_ВР
9. ОПЗДЕП_ИХРДОТ_И_ВРИТЧ_САА
10. ВКБЮСИРЙУ_ОБВНЕ_СОАПНИОТС
11. ПКТИРАОЛНАОИЧ_З_ЕСЬНЕЛНЖО
12. ИПКСОЕ_ТСМНАЧИ_ОЕН_ГДЕЛА_
13. АМВИННЬТЛЕАНЕ_ЙОВ_ОПХАРТО
14. АРЫКЗЫ_КЙТНЛ_ААЫ_ОЛБКЫТРТ
15. _ПАРИИВИАРЗ_БРА_ИСТЬЛТОЕК
16. П_ЛНАЭУВКАА_ЦИИВР_ОКЧЕДРО
17. ЖВНОАН_АТЗОЬСН_БЮ_ФВИИКИЗ
18. ОТВГОСЕЬТАДВ_С_ЬЗАТТЕЫАЧ
19. ЯАМРИТ_ДЖЕХ_СВЕД_ТСУВЕТНО
20. УЬБДТ_ОЕГТВ_ОЫКЭА_ВКАИУЦИ
21. ЛТБЕЧЛЖЫЕ_ОАПТЖРДУ_ЛМНОА
22. ИТПРКРФАГО_АВЯИА_ЯНЖУАКАН
23. ПКЕЕРПО_ЙУСТ_ИТПСУТЛЯЕИН
24. ИЬЖЗНСД_ТДН_ЕТ_НУВЕУРЫГОЫ
25. ЕОУРВА_НЬРИАДИЦЕПИ_РНШВЫЕ

Шифр двойной перестановки

Пример решения:

Дан шифр-текст: ЫОЕЧТТОУ_СНСОРЧТРИАИДЬН_Е

Текст содержит 25 символов, что позволяет записать его в квадратную матрицу 5x5. Известно, что шифрование производилось сначала по столбцам, а затем по строкам, следовательно, расшифрование следует проводить тем же способом.

Ы	О	Е	Ч	Т
Т	О	У	–	С
Н	С	О	Р	Ч
Т	Р	Н	А	И
Д	Ь	Н	–	Е

Производим анализ совместимости символов. Если в примере столбцовой перестановки можно было легко подобрать нужную комбинацию путем перебора, то здесь лучше воспользоваться таблицей частот букв русского языка (см. приложение). Для оптимизации скорости выполнения задания можно проверить все комбинации букв только в первой строке. Получаем ОЕ-15, ОЧ-12, ЕТ-33, ТЕ-31, ЧО-х, ЕО-7, ЧЫ-х, ОЫ-х, ТЫ-11, ТЧ-1, ЧЕ-23 (где х-запрещенная комбинация).

Из полученных результатов можно предположить следующую комбинацию замены столбцов **2 4 3 5 1**:

О	Ч	Е	Т	Ы
О	–	У	С	Т
С	Р	О	Ч	Н
Р	А	Н	И	Т
Ь	–	Н	Е	Д

Теперь необходимо переставить строки в нужном порядке. **3 2 4 5 1**:

С	Р	О	Ч	Н
О	–	У	С	Т
Р	А	Н	И	Т
Ь	–	Н	Е	Д
О	Ч	Е	Т	Ы

Получаем осмысленный текст: СРОЧНО_УСТРАНИТЬ_НЕДОЧЕТЫ

Задание: Расшифровать фразу, зашифрованную двойной перестановкой (сначала были переставлены столбцы, затем строки)

1. СЯСЕ_ _ЛУНЫИАККННОГЯДУЧАТН
2. МСЕЫ_ЛЫВЕНТОСАНТУЕИ_РЛПОБ
3. АМНРИД_УЕБСЫ_ЕЙРСООКОТНВ_
4. ОПЧУЛС_БОУНЕВ_ОЖАЕОНЕЩЕИН
5. ЕШИАНИРЛПГЕЧАВРВ_СЫНА_ЛО
6. АРАВНРСВЕЕОАВ_ЗАНЯА_КМРЕИ
7. А_ЛТАВЙООЛСО_ТВ_ШЕЕНЕСТ_Ь
8. ФИ_ЗИММУЫНУУБК_Е_ДЬШЫИВЧУ
9. ВР_ЕСДЕИ_ТПХРОИ_ЗБУАДНУА_
- 10.ЦТААЙПЕЕ_ТБГУРРСВЬЕ_ОРЗВВ
- 11.АВАРНСЧАА_НЕДВЕДЕРПЕОЙ_ИС
- 12.ДОПК_СОПАЛЕЧНЛ_ГИНЙОИЖЕ_Т
- 13.ЛУАЗИЯНСА_ДТДЕАИ_ШРФЕОНГ_
- 14.С_ОЯНВ_СЬСЛААВРЧЕАРТОГДЕС
- 15.ЗШАФИПРАЛОЕНЖ_ОЪН_ДАРВОНА
- 16.КЭЕ_ТДУМБ_ЬСЗЕДНЕЗМАОР_ТУ
- 17._ЕАЛЯРАНВЯАЧДА_ЕРПЕСАНВ_Ч
- 18._И_ЕНТРЗИ_ОКЕВНОДЛЕША_ИМП
- 19.РОБДОЕВПС_МСХЪА_ _ИВПСНИОТ
- 20.ЕСДНОГТЕАНН_НЕОВМР_ЕУНПТЕ
- 21._ЙЕСТОВО_НИИНЛАЕТИЖДСОПВ_
- 22.НДИАЕОЫЛПНЕ_ _НВЕАНГТ_ИЗЛА
- 23.П_БИРДЛЬНЕВ_ОП_ОПЗДЕВЫГЕА
- 24.МДООИТЕЬ_СМТ_НАДТЕСУБЕХНО
- 25.АИНАЛЖНОЛЕШФ_ЗИ_УАРОЪСНЕ_

Шифр простой замены

Криптоанализ шифра простой замены основан на использовании статистических закономерностей языка. Так, например, известно, что в русском языке частоты букв распределены следующим образом:

Таблица 1.13. Частоты букв русского языка
(в 32-буквенном алфавите со знаком пробела)

- 0,175	О 0,090	Е,Ё 0,072	А 0,062
И 0,062	Т 0,053	Н 0,053	С 0,045
Р 0,040	В 0,038	Л 0,035	К 0,028
М 0,026	Д 0,025	П 0,023	У 0,021
Я 0,018	Ы 0,016	З 0,016	Ь,Ъ 0,014
Б 0,014	Г 0,013	Ч 0,012	Й 0,010
Х 0,009	Ж 0,007	Ю 0,006	Ш 0,006
Ц 0,004	Щ 0,003	Э 0,003	Ф 0,002

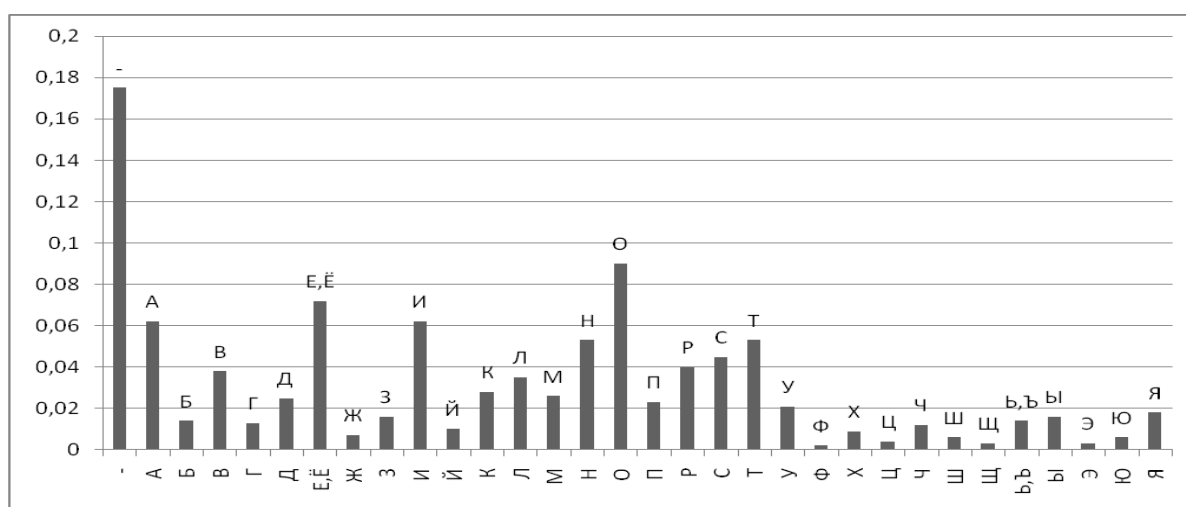


Рис. 6.28. Диаграмма частот букв русского языка

Для получения более точных сведений об открытых текстах можно строить и анализировать таблицы k -грамм при $k > 2$, однако для учебных целей вполне достаточно ограничиться биграммами. Неравновероятность k -грамм (и даже слов) тесно связана с характерной особенностью открытого текста – наличием в нем большого числа повторений отдельных фрагментов текста: корней, окончаний, суффиксов, слов и фраз. Так, для русского языка такими привычными фрагментами являются наиболее частые биграммы и триграммы:

СТ, НО, ЕН, ТО, НА, ОВ, НИ, РА, ВО, КО,

	А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П
А	2	1	3	8	1	7	6	1	7	7	1	2	1	4	3	1
	2	5		4			5			9	7	9	5			
Б	5					9	1		6			6		2	2	
														1		
В	3	1	5	3	3	3		2	1		7	1	3	9	5	6
	5					2		7			0			8		
Г	7				3	3			5		1	5		1	5	
														0		
Д	2		3	1	1	2	1	1	1		1	5	1	1	2	3
	5					9		3					3	2		
Е	2	9	1	1	2	7	5	1	6	1	1	3	2	6	7	1
		8	1	7			0		5	3	5	4	3		6	
Ж	5	1			6	1			5					6		
					2											
З	3	1	7	1	5	3			4		2	1	2	9	9	1
	5															
И	4	6	2	5	1	2	2	2	1	1	1	2	2	3	8	1
		2		0	1		2	3	9	1	9	1	0	2	3	
Й	1	1	4	1	3		1	2	4		5	1	2	7	9	7
К	2	1	4	1		4	1	1	2		1	4	1	2	6	2
	4							6						6		
Л	2	1	1	1	1	3	2	1	3		1	2	1	8	3	2
	5					3		6						0		
М	1	2	4	1	1	2	1	2	2		3	1	3	7	1	5
	8					1		3						9		
Н	5	1	2	3	3	3			5		3			2	6	2
	4					4		8					1	4	7	
О	1	2	8	3	4	1		1	1	2	1	4	3	3	9	1
	8	4	2	7	5		7	8	2	9	9	1	8	0	8	
П	7					1			4			9		1	4	
					5									6		

СТО, ЕНО, НОВ, ТОВ, ОВО, ОВА

Полезной является информация о сочетаемости букв, то есть о предпочтительных связях букв друг с другом, которую легко извлечь из таблиц частот биграмм.

Имеется в виду таблица, в которой слева и справа от каждой буквы расположены наиболее предпочтительные "соседи" (в порядке убывания частоты соответствующих биграмм). В таких таблицах обычно указывается также доля гласных и согласных букв (в процентах), предшествующих (или следующих за) данной букве.

Пример криптоанализа шифра замены

Известно, что зашифровано стихотворение Р. Киплинга в переводе С.Я. Маршака. Шифрование заключалось в замене каждой буквы на двузначное число. Отдельные слова разделены несколькими пробелами, знаки препинания сохранены. Таблица частот букв русского языка приведена выше.

29 15 10 17 29 22 25 31 15 33 35 41 43 45 35 57 45 25 17 59 15 10 25 41 25 69, 59 78 29 82
25 78 25 17 15 10 88 90 78 25 62 25 22 10 57 73 79 35 67 78 90 88 29 45 35 29, 54 57 90 31 90 73
22 88 15 88 29 15 17 69 41 25 15, 70 17 90 57 43 59 15 78 15 62 22 25 17 57 25 69 88 15 82 17 25
88 29 45 35...

Подсчитаем частоты шифрообразований:

Обозначение	9	5	0	7	2	5	1	3	5	1	3	5	4	7
Количество		0				2							4	

Обозначение	9	9	8	2	8	0	2	3	9	7	4	0	7
Количество													1

Из таблица частот букв русского языка видно, что чаще всего встречается буква О, на втором месте Е. В нашем шифр-тексте чаще всего встречается обозначение 25 (12 раз), на втором месте идет обозначение 15 (10 раз), остальные обозначения им существенно уступают. Поэтому можем выдвинуть гипотезу: 25=О, 15=Е. Однако, текст у нас не очень большой, поэтому закономерности русского языка проявляются в нем не обязательно в строгом соответствии с таблицей частот букв русского языка. Поэтому возможен и вариант:

25=E, 15=O. Но тогда последнее слово в третьей строке имеет окончание EO, что возможно, но все же более вероятный вариант OE. Итак, будем работать с текстом, считая, что 25=O, 15=E.

Теперь нам поможет знак препинания: «29, ...». Крайне маловероятно, чтобы запятая стояла после согласной. Итак, 29 – гласная, причем вероятнее всего 29=И или 29=А, т.к. гласные Я, Ю, Э, У встречаются в осмысленных текстах на русском языке намного реже, чем И и А, что не противоречит таблице частот шифр-текста.

В последней строке: 88 15, но 15=E, следовательно, 88 – согласная, причем наиболее вероятные значения – это Н и Т. Итак, 25=O, 15=E, 29=A $\begin{pmatrix} A \\ И \end{pmatrix}$, 88= $\begin{pmatrix} H \\ T \end{pmatrix}$. Теперь третье слово в третьей строке имеет 4 варианта:

- 29=И, 88=Н: 22 Н Е Н И Е
- 29=И, 88=Т: 22 Т Е Т И Е
- 29=А, 88=Н: 22 Н Е Н А Е
- 29=А, 88=Т: 22 Т Е Т А Е

Из рассмотренных вариантов лишь один является осмысленным, и он позволяет найти значение 22. Имеем: 22=М и третье слово в третьей строке М Н Е Н И Е.

Теперь рассмотрим второе слово в первой строке. Е 10 17 И, причем 10 и 17 – согласные, и это не М и не Н. Наиболее вероятное слово Е С Л И, т.е. 10=С, 17=Л. Конечно, если мы, продолжая работать с текстом, вдруг получим «нечитаемое» слово, то придется вернуться к этому этапу и рассмотреть другие варианты. Однако, это маловероятно, поскольку вряд ли в стихотворении были слова наподобие Е Р Т И, Е В Л И и т.п.

Далее, первое слово второй строки: 59 78 И, причем 59 и 78 – согласные, и это не С, не Л, не М и не Н. Так что это слово П Р И, т.е. 59=П, 78=Р. Тогда шестое слово первой строки 45 О Л П Е, что дает значение 45=Т и тогда при 57=В получаем фрагмент «...В Т О Л П Е...». Также второе слово последней строки П Е Р Е 62 дает нам значение 62=Д.

Далее рассмотрим начало второй строки: «П Р И 82 О Р О Л Е С Н 90 Р О Д О М ...». Из него следует, что 82=К и 90=А.

Зная, что 82=К, посмотрим на самое последнее слово К Л О Н И Т 35, откуда станет ясно, что 35=Ь.

Перед последней атакой выпишем текст, заменяя известные обозначения буквами.

ИЕСЛИМОЗІЕЗЗЬ4143ТЬВТОЛПЕСО41О69,
 ПРИКОРОЛЕСНАРОДОМСВ7379Ь67РАНИТЬ
 И,54ВА31А73МНЕНИЕЛ6941ОЕ,
 70ЛАВ43ПЕРЕДМОЛВО69НЕКЛОНИТЬ...

Из последней строки: 69=Ю, тогда слова Л Ю 41 О Е и С О 41 О Ю определяют 41: 41=Б. Теперь из четвертого слова первой строки Б 43 Т Ъ получаем, что 43=Ы. А первое слово из последней строки 70 Л А В Ы – это Г Л А В Ы. Слово в первой строке М О З І Е З З Ъ угадывается из контекста: М О Ж Е Ш Ъ, т.е. 31=Ж, 33=Ш. Теперь второе слово в третьей строке запишется как 54 В А Ж А 73, откуда, с учетом контекста: 54=У, 73=Я. После этого окончание второй строки имеет вид «... С В Я 79 Ъ 67 Р А Н И Т Ъ». Легко определяются буквы 79=З, 67=Х.

Ответ: ИЕСЛИМОЖЕШЬБЫТЬВТОЛПЕСОБОЮ,
 ПРИКОРОЛЕСНАРОДОМСВЯЗЬХРАНИТЬ
 И,УВАЖАЯМНЕНИЕЛЮБОЕ,
 ГЛАВЫПЕРЕДМОЛВОЮНЕКЛОНИТЬ...

Задания: Расшифровать текст. Каждой букве алфавита соответствует двузначное число.

1.

58 62 32 39 99 31 29 58 72 62 99 58 13 54 15 56 31 63 39 72 84 15 13 56 77 15 82 56 56 56 58
 54 29 77 56 – 39 99 56 31 56 7732 12 15 54 31 48 7663 15 52 13 39 72 39 5416 72 39 32 72 62 58
 58 15,37 62 7752 39 13 39 72 39 32 3931 62 54 39 77 84 39 21 31 3916 72 62 99 58 13 15 54 56
 13 4616 39 58 13 95 16 15 13 62 12 46 31 39 6272 15 77 54 56 13 56 6284 31 39 32 56 7658 63 62
 7233 62 12 39 54 62 33 62 58 52 39 9199 62 29 13 62 12 46 31 39 58 13 56.56 31 63 39 72 84 15
 82 56 39 31 31 48 6213 62 76 31 39 12 39 32 56 5616 72 39 33 31 3954 39 53 12 565437 56 77 31
 62 58,39 37 72 15 77 39 54 15 31 56 62,16 72 39 56 77 54 39 99 58 13 54 39,39 13 52 72 48 5433
 62 12 39 54 62 52 9531 62 37 48 54 15 12 48 6254 39 77 84 39 21 31 39 58 13 5616 3958 52 39 72
 39 58 13 5616 39 12 95 33 62 31 56 295639 37 72 15 37 39 13 52 6256 31 63 39 72 84 15 82 56
 56,1513 15 52 21 6216 3915 54 13 39 84 15 13 56 77 15 82 56 5616 72 39 56 77 54 39 99 58 13 54
 62 31 31 48 76,95 16 72 15 54 12 62 31 33 62 58 52 56 765656 31 48 76 16 72 39 82 62 58 58 39
 54.

2.

3925 20 34 82 6366 46 35 20 25 828639 51 74 35 51 66 20 4437 25 27 51 35 44 20 90 37 51 25
 25 51 6391 20 11 37 46 4825 20 37 61 51 14 82 8266 82 35 29 82 91 25 5174 51 24 78 51 24 59 46
 86 51 44 74 20 25 37 37,37 44 82 31 11 37 82 51 46 25 51 34 82 25 37 828637 25 27 51 35 44 20
 90 37 51 25 25 48 4446 82 78 25 51 14 51 18 37 59 44,51 74 82 35 20 90 37 59 446690 82 25 25
 48 44 3761 10 44 20 18 20 44 37,8661 20 25 86 51 39 66 86 51 44 1066 82 86 46 51 35 10 3766 51
 46 51 39 51 6366 39 59 91 37.56 46 51 86 20 66 20 82 46 66 5924 35 10 18 37 7851 35 18 20 25
 37 91 20 90 37 63,4651,66 51 18 14 20 66 25 5135 82 91 10 14 29 46 20 46 20 4435 20 91 14 37
 56 25 48 7837 66 66 14 82 24 51 39 20 25 37 63, 35 10 86 51 39 51 24 37 46 82 14 3744 25 51 18
 37 7837 9125 37 7891 25 20 31 4651 61 51 66 25 51 39 25 48 7839 37 24 20 78 10 18 35 51 91,25
 5125 82 10 24 82 14 59 31 4624 51 14 42 25 51 18 5139 25 37 44 20 25 37 5924 20 25 25 48 4439
 51 74 35 5166 20 44,66 56 37 46 20 59,56 46 5151 61 82 66 74 82 56 82 25 37 8237 25 27 51 35
 44 20 90 37 51 25 25 51 6361 82 91 51 74 20 66 25 51 66 46 3725 8237 44 82 82 4666 44 48 66 14
 20,82 66 14 3751 46 66 10 46 66 46 39 10 82 4639 37 24 37 44 20 5910 18 35 51 91 20.

3.

74 29 23 27 17 99 71 254932 29 34 27 63 32 25 17 99 60 62 25 34 95 29 53 59 82 27 71 29 77
 99 34 27 91 17 99 71 49 99 27 15 60 32 25 50 27 17 62 27 95 27 50 25 91 32 59 77 95 29 50 25 99
 59,25 99 74 29 53 25 59 17 99 25 91 23 49 71 25 17 99 604925 34 32 25 71 95 27 82 27 32 32
 2529 50 17 25 15 77 99 32 59 7762 95 25 53 95 29 23 32 25 17 99 60 34 15 35 17 27 99 27 71 25
 12 2599 95 29 45 49 74 29. 62 95 27 63 34 2771 17 27 12 25,50 27 17 62 27 95 27 50 25 91 32 29
 3595 29 50 25 99 29 17 29 82 49 8362 2517 27 50 2762 95 25 34 59 74 99 25 7150 27 53 25 62 29
 17 32 25 17 99 4917 71 35 53 29 32 2917 32 29 15 49 23 49 27 8232 29 34 27 63 32 2595 29 50 25
 99 29 77 10 27 12 2525 50 25 95 59 34 25 71 29 32 49 3549 95 27 53 27 95 71 49 95 25 71 29 32
 49 27 8274 95 49 99 49 23 32 89 837425 99 74 29 53 5950 15 25 74 25 7162 49 99 29 32 49
 354953 29 62 25 82 49 32 29 77 10 49 8359 17 99 95 25 91 17 99 71.34 15 3562 25 17 15 27 34 32
 49 8325 62 99 49 82 29 15 60 32 2562 95 49 82 27 32 27 32 49 2734 49 17 74 25 71 89 8382 29 17
 17 49 71 25 7112 25 95 35 23 27 9153 29 82 27 32 89.74 29 23 27 17 99 71 25 49 32 29 34 27 63
 32 25 17 99 60 95 29 50 25 99 8934 25 17 99 49 12 29 27 99 17 3525 62 99 49 82 49 53 29 67 49
 27 9162 95 25 12 95 29 82 82 32 25 12 2525 50 27 17 62 27 23 27 32 49 35.

4.

48 2318 40 94 35 62 53 94 25 53 15 3591 35 40 35, 52 23 5253 40 3594 35 40 2394 23 91 52
 94 49 24 23 84 8994 23 64 55 53 15 18 53 91, 24 53 88 23 62 12 25 7694 2364 35 24 49, 35 9449
 88 5348 94 23 24,41 91 3591 23 5231 49 15 53 91. 47 91 3541 49 62 84 91 62 3535 91 41 23 84 91
 2531 29 24 3564 35 27 35 88 5394 2391 35,52 35 91 35 55 35 5335 9425 84 64 29 91 23 24,52 35

40 15 2348 23 62 53 55 94 49 2448 2349 40 35 242541 49 91 8994 5394 23 24 53 91 53 24 94
 2315 53 62 49 12 52 49,12 53 15 12 49 6053 18 4994 23 62 84 91 55 53 41 49.53 40 3594 35 40
 23,62 29 48 62 23 6284 62 35 25 1815 62 25 88 53 94 25 53 18 52 35 24 53 31 23 94 25 53 62 35
 48 15 49 27 23,64 35 24 49 41 25 24 23 35 91 55 23 88 53 94 94 29 7684 25 40 94 23 243564 55
 53 64 38 91 84 91 62 25 2594 2364 49 91 25 2564 35 41 91 256291 4988 5384 53 52 49 94 15
 4949 15 23 55 25 24 23 84 8935 31 3541 91 35 – 91 35.52 23 52 35 76-91 3564 55 53 15 18 53
 918440 24 49 27 25 1884 91 49 52 35 1835 91 24 53 91 53 246291 53 18 94 35 91 49.

5.

79 6131 96 28 35 85 5226 30 24 21 52 85 59 49 79 30 88 7949 30 52 79 59 85 26 30 24 21 59
 85 42 79 88 61 28 35 86 5096 28 52 30 50,24 30 96 74 21 59 9059 30 96 30 24 85 61 8626 96 85
 88 79 96 79 24 61 79 1128 52 79 78 31 85, - 21 50 30 96 85 31 21 61 59 31 85 1126 79 24 96 79 59
 35 79 31 5996 30 31 52 21 50 61 79 1131 21 96 35 85 61 31 85,2126 79 78 30 50 2867 868561 30
 35:35 79 24 2467 79 28 24 30 61,35 96 85 61 21 24 69 21 35 9052 30 35,61 79 96 50 21 52 90 61
 86 1196 79 59 35,42 24 79 96 79 49 86 1149 30 59,49 79 52 79 59 8669 49 30 35 2159 26 30 52 79
 1126 46 30 61 85 69 86,88 79 52 28 67 86 3088 52 21 42 21,96 79 49 61 86 3067 30 52 86 3042 28
 67 86,42 21 88 79 96 30 52 79 3052 85 69 79,61 3085 59 26 79 96 78 30 61 61 79 3024 21 74 3061
 21 50 30 31 79 5061 2149 79 42 96 21 59 35 61 86 30 26 96 86 29 85 31 85..

6.

56 27 54 54 27 56 51 32 82 16 63 49 27 63 11 30 73 35 23 54 89 70 27 63 27 493270 35 16 97
 82 16 67 73 27 51 30 56 32 6370 29 63 27 49 32 73 29 5473 2748 29 13 29 82 56 82 27 9554 27 35
 27 18 51 29,97 56 2770 29 63 305151 35 15 63 89 48 16.16 63 15 11 51 3082 2949 65 27 54 32 63
 304929 61 2763 32 48 30-27 56 51 35 15 56 30 233227 11 70 27 35 27 18 32 56 29 63 89 82 30
 23,27 82 3051 30 5111 1573 35 29 54 70 27 49 65 32 38 30 63 3073 35 32 23 56 82 16 6770 49 56
 35 29 97 16.82 27 49 51 27 1351 29 54 3027 8227 73 16 49 56 32 6370 29 63 27 49 32 73 29 54 82
 15 9516 73 27 353270 15 56 30 38 32 6332 92-73 27 5411 30 61 30 18 82 32 51 3049 63 27 18 29
 82 82 16 67 61 30 92 29 56 16.27 8249 16 82 16 6361 30 92 29 56 1673 27 5413 15 24 51 163270
 92 27 24 29 6373 2749 56 16 73 29 82 89 51 30 13.

7.

3428 68 91 1383 10 65 27 6849 10 26 65 27 68 75 26 39 785375 83 53 18 26 36 62 91.26 10 74
 53 1349 10 83 10 65 5353 36 68 72 28 1028 13 18 86 10 27 53 75 3983 6857 26 18 10 91535736
 53 6528 68 91 10,83 68 75 27 1334 13 24 13 18 53 36 74 5336 10 74 10 36 57 36 13,83 68 74 1091
 10 91 1036 1368 26 74 18 62 34 10 27 1036 10 75 26 13 86 3968 74 36 10.83 18 10 34 28 10,26 57

2650 62 27 6883 68 65 57 86 13.26 57 2649 10 83 10 65 5334 19 13 27 53 75 395334 75 1375 68
 50 68 1583 18 68 83 53 26 10 27 53.49 10 83 10 65 5310 27 74 68 72 68 27 44,83 68 28 72 68 18
 13 34 80 13 72 6891 10 75 27 10,83 68 26 10,75 26 10 18 68 1568 28 13 86 28 625313 96 1327 13
 74 10 18 75 26 34-91 13 36 26 68 27 1053,74 10 86 13 26 75 44,34 10 27 13 18 39 44 36 74
 53.3483 18 53 65 68 86 13 15 26 13 91 36 68 26 53 96 10,5318 44 28 68 9123 26 68 2628 78 75 75
 10 36 28 13 18-34 26 44 36 57 2772 68 27 68 34 573434 68 18 68 26,23 26 10 74 53 1572 18 53 47
 – 75 26 13 18 34 44 26 36 53 74,86 28 57 96 53 15,74 68 72 28 1018 10 36 13 36 68 1386 53 34 68
 26 36 68 1353 75 83 57 75 26 53 2628 57 65.

8.

45 34 26 34 9777 34 47 49 67 14 22 49 6747 34 49 39 77 6953 89 26 1097 10 49 10 77 45 53
 31 10 14 10 47 22.17 90 56 14 34 77 67 49,49 67 75 49 1053 14 5349 26 90 47 10,77 3439 47 56
 34 3156 26 67 52 34 13 10 84 22 5377 34 47 49 67 14 22 49 67 28 34 84 26 67 31,67 49 10 97 90
 31 10 14 53 47 223128 70 89 49 53 9314 10 56 10 9356 47 10,5345 34 84 90 26 34 93 69 58 37 28
 67 31 10 7047 84 10 14 22 77 10 7053 89 14 10,31 90 47 39 77 39 31 75 53 47 22,47 14 67 31 77
 6713 10 14 67,53 9734 89 6728 67 26 69 90,31 56 26 90 47 49 53 31 10 14 1013 34 26 84 31 3453
 97 26 70 69 77 39 5869 67 97 39 28 67 26 24 53 70,53 14 5356 26 67 49 10 53 77 10.97 10 84 34
 2839 52 53 84 67 89 6797 31 34 26 22 49 1052 26 67 47 10 14 533156 34 45 2269 14 7047 13 53
 89 10 77 53 7028 39 47 67 26 10,5353 89 26 1077 10 45 53 77 10 14 10 47 2247 77 67 31 10.

9.

81 49 86 49 1273 92 5081 50 15 5062 47 4915 56 50 51 7673 33 94 7615 94 65 81 47 76.94 76
 47 49 81 47 76,15 7662 47 76 2628 16 5162 76 2628 76 51 70 58 76 2673 86 65 84 76 94,47 7615
 94 65 81 47 7615 56 50 51 76.24 16 51 7062 76 49 2694 76 86 76 28 94 3362 49 47 1765 84 4915
 76 92 15 49 6247 4924 86 49 51 70 96 50 51 50.56 76 31 73 5047 49 62 47 76 31 7624 76 73 65 62
 50 513386 49 58 33 5115 56 50 567 065 62 47 16 62.47 65,47 50 73 7684 4943 76 56 7081 56 76-
 56 7673 49 51 50 56 70...1724 76 58 49 519294 76 51 51 49 73 84.76 94 50 12 50 92 58 33 15
 709294 50 28 33 47 49 56 496586 49 94 56 76 86 50,1773 49 86 84 50 51 15 1765 92 49 86 49 47
 47 76.86 49 94 56 76 86 76 6228 16 51 5062 76 51 76 73 50 1784 49 47 96 33 47 5028 50 51 70 12
 50 94 76 92 15 94 76 31 7692 76 12 86 50 15 56 50.94 76 31 73 501792 76 58 49 51,76 47 5081 56
 76-56 7615 76 15 86 49 73 76 56 76 81 49 47 47 7624 33 15 50 51 50,62 76 84 49 5647 76 92 16
 2665 94 50 12.

10.

2043 40 13 15 91 31 5475 31 91 12.88 56,88 40 29 1571 3113 15 91 1249 91 15 – 91 1529 31
 54 40 91 12...1715 61 69 31 44,2075 15 36 31 546275 25 15 29 84 65 31 25 56.90 4415 62 40 43
 40 54 65 2088 31 17 58 65 15 62 90 2690,75 15-17 90 29 90 44 15 44 56,88 31 29 40 54 31 62 90
 2649 31 54 15 17 31 621791 31 44 88 58 1315 49 62 40 13901725 15 43 15 17 15 4436 40 25 34
 90 62 3188 4036 31 31.15 8862 56 25 90 5449 91 15-91 1515 49 31 88 1275 25 15 91 90 17 88
 1575 40 13 88 56 69 31 31.29 40 71 3117 15 88 20 84 69 31 31.56 17 90 29 31 1744 31 88 20,75
 25 15 29 84 65 31 2588 31 65 62 15 54 12 62 1544 90 88 56 9175 15 44 56 49 40 54 65 20,17 65
 91 40 17 54 20 2015 91 17 90 65 36 56 8449 31 54 84 65 91 1288 4044 31 65 91 15,88 1517 65
 3171 3117 43 20 5465 31 61 201725 56 62 90,43 40 91 56 36 90 5465 90 52 40 25 31 91 569043
 40 52 15 17 15 25 90 54.

11.

65 27,67 40 58 34 11 4727 4227 45 82 34 11 14 4914 89 95 47.65 14 90 36 89 3434 67 36 90
 36 45 67 11 36 65 65 34 89 34,11 17 82 34 67 1924 3495 40 45 17 34 45 82 36 24 65 14 7025 36
 82 34 90 36 73.70 34 67 4945 67 95 40 65 40,17 34 45 95 36 24 1458 34 67 34 95 34 7334 65 1445
 36 73 90 40 4517 95 36 59 47 11 40 82 14,24 40 11 65 341465 40 24 36 42 65 3417 34 24 25 49 67
 4040 25 36 95 14 58 34 45 40 25 14,69 67 3411 45 3642 3645 27 11 36 95 36 65 65 40 4924 36 95
 42 40 11 40,90 82 36 6534 34 65,4558 34 36 7345 34 11 36 67 45 58 14 7345 34 31 6317 34 24 24
 36 95 42 14 11 40 36 6765 34 95 25 40 82 19 65 47 3624 14 17 82 34 25 40 67 14 90 36 45 58 14
 3634 67 65 34 32 36 65 14 49,17 34 65 36 25 65 34 89 2765 40 82 40 42 14 11 40 36 6767 34 95
 89 34 11 82 31,17 95 14 45 47 82 40 36 6765 4089 40 45 67 95 34 82 1459 40 82 36 67 65 47 3667
 95 27 17 17 471434 59 25 36 65 14 11 40 36 67 45 4917 95 34 18 45 34 31 63 65 47 25 1424 36 82
 36 89 40 56 14 49 25 14.4017 34 67 34 25 2763 24 36 45 1965 14 58 40 5865 3617 34 82 40 89 40
 36 67 45 4965 36 82 36 89 40 82 19 65 3417 95 36 59 47 11 40 67 1945 34 11 36 67 45 58 14 2559
 34 36 11 47 2517 82 34 11 56 40 25,“25 34 95 45 58 14 2524 19 49 11 34 82 40 25”.36 42 36 82
 1490 67 34-45 58 40 65 24 40 8295 40 63 89 34 95 14 67 45 4917 3417 34 82 65 34 73...

12.

14 701465 3659 47 82 34,4058 40 5842 36.17 95 34 45 67 34-65 40 17 95 34 45 67 3432 36 45
 67 36 95 3425 27 42 14 58 34 11,65 4011 14 24-45 67 40 65 24 40 95 67 65 47 3636 11 95 34 17
 36 34 14 24 47,4563 40 17 40 24 65 34 89 36 95 25 40 65 45 58 14 25 1440 11 67 34 25 40 67 40
 25 14,14 67 40 82 19 49 65 45 58 14 25 1440 58 11 40 82 40 65 89 40 25 14,32 11 36 24 45 58 14
 25 1459 40 63 27 58 40 25 14,59 36 82 19 89 14 73 45 58 14 25 1425 14 65 40 25 14,18 95 40 65
 56 27 63 45 58 14 25 1445 14 89 40 95 36 67 40 25 141432 11 36 73 56 40 95 45 58 14 25 1490 40
 45 40 25 14.17 95 36 24 25 36 67 4745 65 40 95 49 42 36 65 14 49,11 63 49 67 47 3617 3434 67

24 36 82 19 65 34 45 67 14,25 34 42 65 3459 36 6334 45 34 59 47 7070 82 34 17 34 6717 95 14 34
 59 95 36 45 67 141195 40 63 65 47 7058 34 65 56 40 7036 11 95 34 17 4758 40 5882 36 89 40 82
 19 65 34,67 40 581465 4090 36 95 65 34 2595 47 65 58 36-58 40 58,45 34 59 45 67 11 36 65 65
 34,1417 95 34 14 63 34 32 82 3467 95 27 24 40 25 1465 36 11 36 24 34 25 47 7025 40 63 27 95
 27“14 65 67 36 65 24 40 65 67 34 11”.

13.

60 46 5746 52,28 15 57 3912 32 60 32 3246 5752 55 30 12 61 11 55 57 32 12 41,37 46 60 37
 32 9152 32 11 55 12 32 75 4646 5730 32 20 15 75 46 25 99 20 52 32 52 52 4667 55 25 55 12 12 32
 12 39 52 19 63“52 99 57 32 36”75 46 12 61 28 75 99(18 32 37 57 3952 99 57 32 3667 46 60 32 25
 63 159991 32 57 25 46 60 46 3660 19 37 46 57 19“37 67 99 25 55 12 3930 25 15 52 46 ”67 4620
 32 91 12 32),57 5537 55 91 55 4167 57 99 28 75 55.75 25 55 37 55 60 32 74,37 57 46 99 5767 25
 99 20 52 55 57 39,99 20 41 45 52 19 36,11 12 99 52 52 46 75 25 19 12 19 36,37 15 67 32 25 55 29
 25 46 11 99 52 55 91 99 28 32 37 75 99 36,60 19 37 46 57 52 19 36.“11 48 99 – 29 25 – 11 60 32
 52 55 11 74 55 57 39”,52 46 60 32 36 18 99 3637 55 91 46 12 32 5729 12 32 75 57 25 46 52 52 46
 3625 55 20 60 32 11 75 99,46 37 52 55 45 32 52 52 19 3655 67 67 55 25 55 57 15 25 46 36,78 46
 25 11 4699 91 32 52 15 32 91 46 36“57 32 63 52 99 75 46 3611 60 55 11 74 55 57 3967 32 25 60
 46 78 4660 32 75 55”(63 46 57 4111 4675 46 52 74 5511 60 55 11 74 55 57 46 78 4637 57 46 12
 32 57 99 41,37 46 78 12 55 37 52 4663 25 46 52 46 12 46 78 99 99,46 37 57 55 12 46 37 3932 45
 3267 41 57 52 55 11 74 55 57 393712 99 18 52 99 9112 32 57)...

14.

15 48 3252 326067 32 25 60 19 3625 55 2091 55 20 15 25 1567 25 99 63 46 11 99 12 466078
 46 12 46 60 15,28 57 4628 99 52,46 57 60 32 28 55 60 18 99 3620 5530 32 20 46 67 55 37 52 46
 37 57 3930 55 20 19,30 19 12 75 12 55 37 37 99 28 32 37 75 99 9137 15 63 46 67 15 57 28 99 75
 46 91. 60 37 60 46 3260 25 32 91 4146 52 67 46 25 55 30 46 57 55 12 52 55 37 46 60 32 37 57
 39,46 30 46 25 15 11 46 60 55 60 37 15 63 46 67 15 57 52 19 3267 46 11 37 57 15 67 197530 55
 20 3232 91 75 46 37 57 52 19 91 9911 55 57 28 99 75 55 91 99,37 99 78 52 55 12 39 52 19 91
 9925 55 75 32 57 55 91 99,67 25 9991 55 12 32 36 18 32 9167 25 99 75 46 37 52 46 60 32 52 99
 997557 46 52 61 37 32 52 39 75 46 3652 99 57 9960 20 12 32 57 55 60 18 99 91 996052 32 30 32
 37 5537 4637 60 99 37 57 46 91,25 55 37 37 19 67 55 4160 46 25 46 63 5525 55 20 52 46 74 60 32
 57 52 19 6346 37 12 32 67 99 57 32 12 39 52 19 6399 37 75 25-9911 55 48 3267 46 12 46 37 55 91
 9967 25 46 57 99 60 46 67 32 63 46 57 52 19 6391 99 52.28 57 46 75 55 37 55 32 57 37 4167 46
 11 37 57 15 67 46 6060 46 11 52 19 63,28 99 5230 19 1252 3257 55 7525 32 57 99 60. 46 11 52
 9957 46 12 39 75 4637 57 46 12 30 193775 46 12 61 28 75 46 369967 25 32 37 12 46 60 15 57 19
 32“37 67 99 25 55 12 39 75 99”-75 46 57 46 25 19 3252 32 20 60 55 52 19 3278 46 37 57 99,6046

57 12 99 28 99 3246 5720 11 32 18 52 99 6367 55 25 57 99 20 55 52,15 91 32 12 9967 25 32 46 11
46 12 32 60 55 57 3930 19 37 57 25 469930 32 2091 55 12 32 36 18 32 78 4660 25 32 11 5511 12
4137 46 30 37 57 60 32 52 52 46 78 4646 25 78 55 52 99 20 91 55.9960 37 32.

15.

45 74 5431 10 26 38 23 74,86 74 5425 89 26 38 16 74 7475 1645 56 90 25 86 90 75 90 10 2616
74 23 56 86 75 45 16 75 7495 10 13 31 95 10 51 74 16 89 74,36 75 95 75 5936 74 95 74 91 75 31
89 90 23 74 749036 95 89 26 89 90 8313 26 75 25 86 89-75 86 86 75 47 75,45 86 7575 16 8945 74
86 90 74 95 7525 56 86 75 33,75 29 95 10 86 89 90 23 89 25 389013 95 74 16 89 748925 26 56
91,86 75 95 45 10 26 899045 10 19 75 29 74,33 10 3331 89 33 89 7475 29 74 13 38 42 16 8389
1329 95 10 13 89 26 89 89,75 86 86 75 47 75,45 86 7536 75 31 90 74 95 16 56 26 25 4286 56 36
75 4633 10 46 54 10 16,2575 31 89 16 10 33 75 90 83 5456 25 74 95 31 89 74 5416 10 36 10 31 10
90 23 89 468916 1026 74 25 16 56 5925 90 89 16 38 59,8916 1075 86 26 89 45 16 75 47 7536 10
95 16 422531 95 56 47 75 47 7533 75 16 86 89 16 74 16 86 10.109021 86 7590 95 74 54 4286
74,16 1029 10 13 74,51 89 26 899025 90 75 7456 31 75 90 75 26 38 25 86 90 89 74,25 36 10 26
8916 1045 89 25 86 74 16 38 33 89 9136 95 75 25 86 83 16 33 10 919033 75 16 31 89 17 89 75 16
89 95 75 90 10 16 16 75 4636 95 75 91 26 10 31 74,36 95 89 16 89 54 10 26 8931 56 23,51 95 10
26 8916 1013 10 90 86 95 10 3367 95 56 33 86 83,31 51 74 548929 89 67 23 86 74 33 25 839086
95 8936 10 26 38 17 1086 75 26 19 89 16 75 46-8975 33 16 1086 10 3356 59 86 16 7525 90 74 86
89 26 89 25 38,8954 56 13 83 33 1089 47 95 10 26 10,8967 56 86 29 75 26 36 75 86 74 26 74 90
89 13 75 95 56...

16 89 45 74 47 75 9021 86 75 4 613 26 75 25 86 8916 7429 83 26 7536 26 75 91 75 47 75,16
10 75 29 75 95 75 86-86 10 33 75 4616 10 25 86 95 75 4633 10 3395 10 138936 95 89 31 10 74
8629 75 74 90 75 47 7533 56 95 10 51 10...

1036 75 86 75 5436 95 89 23 74 2633 75 16 74 178936 75 25 86 75 95 75 16 16 89 5454 83 25
26 42 548929 74 13 31 74 26 38 59.54 75 95 25 33 75 46 13 54 74 4616 10 33 75 16 74 17-86 7536
75 31 10 2613 16 10 33,33 75 86 75 95 75 47 7575 16 8951 31 10 26 8945 74 86 90 74 95 7525 56
86 75 33,8921 86 7529 83 26 7525 26 75 90 16 7554 74 31 16 83 4695 74 9029 75 74 90 75 4686
95 56 29 83,21 86 7575 13 16 10 45 10 26 75,45 86 7516 10 45 10 26 10 25 3895 10 29 75 86
83,8916 89 45 74 47 7556 51 7416 7489 13 54 74 16 89 86 38,16 7475 25 86 10 16 75 90 89 86
38,16 7436 74 95 74 89 47 95 10 86 38...

16.

15 22 67 30 93 4922 94 65 94 44 49,4939 51 22 75 49 411115 22 4911 53 51 75 51 78 94,44
4927 51 22 67 44 86 51,26 49 39 51 75“78 45 94 – 62 75 – 78 11 51 44 49 78 91 49 22 72 14”,9411

67 26 93 5 144 51 90 6793 51 44 94 11 6753 75 67 41 49 45 94 11 49 93 15 3035 49 15 67 11 67
 14,44 5145 78 49 11 65 94 1444 94 86 49 86 94 4115 20 75 53 75 94 26 67 11,44 5153 67 78 67 26
 75 51 11 49 11 65 94 14,35 22 6751 90 6715 39 51 75 22 5853 75 51 27 72 11 49 51 2215 67 11 15
 51 3944 51 53 67 78 49 93 51 86 881167 27 75 49 26 5127 51 15 53 93 67 22 44 67 90 6735 51 75
 44 67 90 6753 75 94 26 75 49 86 49,44 5126 44 49 20 18 51 90 6745 49 93 67 15 22 94.

67 35 51 75 51 78 44 67 1445 51 15 2286 67 39 49 44 78 94 75 49-9439 49 26 88 751511 94
 86 94 44 90 67 399415 22 75 49 65 94 93 67 1453 51 75 51 27 51 45 86 49 39 9478 11 94 44 88 93
 94 15 5811 53 51 75 51 78.26 78 51 15 5841 11 49 22 49 93 6753 75 67 45 51 86 22 67 75 67
 11,36 67 44 49 75 51 149486 75 67 44 65 22 51 14 44 67 111590 94 75 93 30 44 78 49 39 9493 49
 39 53,44 6744 51 75 51 49 93 58 44 67 1426 49 78 49 35 51 1427 72 93 6727 7267 15 11 51 22 94
 22 5811 15 2027 49 26 88.67 15 22 49 11 49 93 67 15 5844 51 39 49 93 6753 67 93 67 159453 30
 22 51 4422 51 39 44 67 22 72,86 67 22 67 75 88 20 44 51 26 11 49 44 72 5190 67 15 22 9494 15
 53 67 93 58 26 67 11 49 93 9439 49 15 22 51 75 15 86 94.11 15 5127 93 94 45 518615 49 39 67 93
 51 22 88,27 93 94 45 51,27 93 94 45 51,67 4411 72 75 49 15 22 49 51 2244 4990 93 49 26 49
 41,44 49 11 94 15 49 51 2244 49 7890 67 93 67 11 67 14,88 45 5153 75 51 86 75 49 15 44 6715 93
 72 65 44 67,86 49 8635 49 15 67 11 67 1467 2215 86 88 86 9444 88 78 94 2253 67 7844 67 1544
 51 26 44 49 86 67 39 88 2039 51 93 67 78 94 20,53 67 15 93 51 78 44 20 201115 11 67 51 1445 94
 26 44 94...

22 94 41 67 44 58 86 6718 51 93 86 44 88 9327 51 15 65 88 39 44 72 1453 94 15 22 67 93 51
 22-9439 51 93 67 78 94 3067 27 67 75 11 49 93 49 15 58,35 49 15 67 11 67 1453 67 78 93 67 39
 94 93 15 301186 67 93 51 44 86 49 41,44 6788 53 49 15 22 5844 5188 15 53 51 93,9415 11 67
 2049 11 22 67 39 49 22 94 35 51 15 86 88 2011 94 44 22 67 11 86 8844 5111 72 75 67 44 94 93.78
 11 5122 51 44 94,27 51 15 65 88 39 44 6711 72 44 72 75 44 88 1194 26 -53 67 7836 20 26 51 93
 30 45 49,53 67 78 41 11 49 22 94 93 9451 90 679488 11 67 93 67 86 93 94 44 4978 75 88 90 88
 2015 22 67 75 67 44 88,1122 51 39 44 67 22 88.

17.

56 67 9218 58 39 99 27 87 67 5625 56 80 67 10 17 92 39 6225 5627 24 95 56 3195 46 27 73 56
 3117 58 39 58 67 95 589256 95 40 24 40 17 92 39 626939 40 17 56 67 58-56 18 99 92 46 67 56
 87,69 5669 39 3680 17 92 67 2739 40 87 56 17 58 73 40.25 56 39 73 56 10 17 92,56 43 92 80 40
 10,95 56 23 80 4023 17 40 24 4025 46 92 69 14 95 67 27 739573 58 87 67 56 73 58.69 39 5869 56
 95 46 27 2325 46 92 67 10 17 5638 58 73 95 92 5856 38 58 46 73 40 67 92 10.25 46 92 18 56 46
 56 699225 27 17 62 73 56 6924 80 58 39 6218 14 17 5625 46 58 69 58 17 92 95 56 5887 67 56 43
 58 39 73 69 56,23 17 40 24 4046 40 24 18 58 23 40 17 92 39 62.56 80 67 40 95 5618 17 40 23 56
 80 40 46 1073 58 8743 58 80 69 27 8767 58 80 58 17 10 8773 46 58 67 92 46 56 69 56 9567 4087
 40 95 58 73 58 9273 14 39 10 38 58 95 46 40 73 67 56 25 56 69 73 56 46 58 67 67 14 8767 40 39

73 40 69 17 58 67 92 10 8792 67 39 73 46 27 95 73 56 46 4056 67 9239 56 69 58 46 99 58 67 67
5673 56 38 67 5624 67 40 17 92,24 4038 58 8725 46 92 99 17 92.25 56 67 10 73 92 1067 5892 87
58 17 92,80 17 1038 58 23 5695 56 67 95 46 58 73 67 5625 46 58 80 67 40 24 67 40 38 58 67 1469
39 5871 73 9299 73 27 95 92-67 5656 7367 92 8271 73 56 23 56 9267 5873 46 58 18 56 69 40 17
56 39 62.

67 5825 46 56 99 17 569287 92 67 27 73 14,95 40 9556 6727 69 92 80 58 1751 58 17 6292
8267 58 17 58 23 95 56 23 569271 95 24 56 73 92 38 58 39 95 56 23 5625 27 73 58 99 58 39 73 69
92 10-73 46 9225 27 17 62 73 4025 5625 46 40 69 56 87 2718 56 46 73 27,27 39 14 25 40 67 67 14
5838 58 46 73 56 69 56 3127 31 87 56 3173 27 87 18 17 58 46 56 69,17 40 87 25 56 38 58 95,25
58 46 58 95 17 36 38 40 73 58 17 58 319295 67 56 25 56 95.73 46 9269 14 25 27 95 17 14 8271 95
46 40 67 406969 92 80 5869 58 46 73 92 95 40 17 62 67 14 8225 46 10 87 56 27 23 56 17 62 67 92
95 56 69-56 67 9239 40 87 14 58,67 92 95 40 95 56 3156 99 92 18 95 92...

18 56 80 46 56 39 73 9246 40 80 92,56 6725 56 69 73 56 46 92 1725 46 5639 58 18 1025 56 17
36 18 92 69 99 27 36 39 1051 92 73 40 73 27:“38 73 5656 80 92 6738 58 17 56 69 58 9525 56 39
73 46 56 92 17,80 46 27 23 56 3124 40 69 39 58 23 80 4046 40 24 17 56 87 40 73 6239 87 56 43
58 73”.92,25 56 82 17 56 25 40 6925 5625 17 58 38 27 39 73 46 40 99 92 17 276924 67 40 9573
56 23 56,38 73 5667 40 25 40 46 67 92 9580 56 17 43 58 6718 80 92 73 58 17 62 67 5639 73 56 10
73 6267 4099 27 82 58 46 58,80 56 39 73 40 1795 92 67 43 40 1792 2425 46 92 99 92 73 14 8267
40 8095 56 17 58 67 56 8767 56 43 58 67.

18.

67 58 26 19 88 2332 37 15 23 90 63 7146 63 26-63 2658 2463 23 3732 956763 15 32 88 58 26-
6726 58 6741 16 24 90 63 52 30 2449 63 2688 26 37 23 38 23 16 6758 2390 26 41 90 63 68 24 58
58 26 7685 15 67 76 24 15 24.19 26 15 23 38 88 2663 15 32 88 58 24 2490 88 24 16 23 63 7163 23
37,46 63 26 41 5437 15 23 95 676758 2438 23 76 24 63 67 16 6768 26 68 90 24,67 58 23 46 2437
63 26 – 63 2658 24 19 16 32 85 54 4426 46 24 58 7141 54 90 63 15 2690 88 24 16 23 24 6390 26
26 63 68 24 63 90 63 68 32 11 30 67 2468 54 68 26 88 546724 30 24,46 24 19 2688 26 41 15 26 19
26,85 15 67 76 24 63 90 5237 16 24 68 24 63 23 63 71,68 15 23 95 67 58 2367 88 24 26 16 26 19
67 46 24 90 37 23 52,85 32 90 63 7188 23 95 243258 24 19 266758 2441 32 88 24 6388 26 37 23
38 23 63 24 16 71 90 63 68,58 263746 24 76 3258 23 7616 67 83 58 52 5237 16 24 68 24 63 23?63
26-63 26...49 63 26 6356 67 58 23 1685 26 38 68 26 16 52 1626 88 58 67 7676 23 73 26 7615 24
83 67 63 7158 24 90 37 26 16 71 37 2638 23 88 23 46.58 2441 54 16 2658 67 37 23 37 26 4437 15
23 95 67,90 26 68 24 15 83 24 58 58 26 4473 68 23 63 37 67 76 6763 15 24 58 67 15 26 68 23 58
58 54 76 6715 24 41 52 63 23 76 67-85 15 26 90 63 26-58 23 85 15 26 90 63 2626 37 15 24 90 63
58 54 2485 23 15 63 67 38 23 58 54,88 23 68 58 54 76-88 23 68 58 2619 15 26 38 67 68 83 67

2488 26 41 15 23 63 71 90 5268 90 2495 2488 2626 85 16 26 63 2367 76 85 24 15 67 23 16 67 38
 76 23,90 67 15 24 46 7188 23 58 58 26 4441 23 38 54,90 68 26 1132 19 15 26 38 326837 26 58 29
 2437 26 58 29 26 6868 54 85 26 16 58 67 16 67.58 23 19 15 52 58 32 16 6758 26 46 58 26 4485 26
 15 26 44,85 26 15 24 38 23 16 6737 26 16 11 46 37 32,85 15 26 58 67 37 16 6758 2341 23 38 3285
 26 8885 26 37 15 26 68 26 7676 15 23 37 23,38 23 16 26 95 67 16 679085 26 16 88 11 95 67 58
 5476 67 58,85 26 90 63 15 24 16 52 16 6767 3819 15 23 58 23 63 26 76 24 63 26 6867,90 85 15 23
 68 24 88 16 67 68 2615 24 83 67 68,46 63 2688 26 90 63 23 63 26 46 58 2658 23 85 23 37 26 90
 63 67 16 67,38 16 26 15 23 88 58 2685 26 16 11 41 26 68 23 16 67 90 7188 24 16 26 7615 32 3790
 68 26 67 736732 41 15 23 16 67 90 7168 26 90 68 26 52 90 6741 24 3876 23 16 24 44 83 24 19
 2688 16 5290 24 41 5232 15 26 58 23.

19.

3492 45 25 90 30 25 7116 62 37 7155 7189 18 96 6255 85 22 71 11 6262 24 62 89 71 55 55
 6285 55 16 71 92 71 24 55 62 11 62-90 30 49 30 24 55 18 7124 16 85 92 30 55 18 7152 37 85 55
 24 18,49 30 92 62 22 25 3022 85 24 16 18 7392 58 89 30 67 71 25,90 58 89 55 30 2086 71 16 25
 302416 45 89 85 25 62 1449 30 24 16 18-89 92 62 52 20 11 7168 16 6249 62 96 62 37 71 55 62,25
 62 96 8562 5530 34 24 16 92 30 96 85 71 4692 62 52 62 14,49 30 92 3089 30 55 62 2525 62 55 24
 71 92 34 62 34,49 62 22 30 16 18 9439 96 30 25 62 552462 89 71 90 90 30 92 30 37 85 34 30 45
 86 85 14 8534 62 52 5816 30 89 96 71 16 25 30 14 85,24 16 30 92 18 9425 62 14 49 30 24,62 89
 67 30 92 49 30 55 55 18 9439 62 55 30 92 85 258549 92 62 22 30 2052 92 71 89 71 52 71 55 19,85
 90 62 89 96 85 22 30 34 67 30 203452 37 62 55 7124 16 19 45 -25 30 25 -71 11 62-16 30 1452 62
 24 16 30 16 62 22 55 6262 49 18 16 55 62 11 6249 58 16 71 67 71 24 16 34 71 55 55 85 25 30,14
 30 16 92 62 24 302455 71 14 30 96 18 1424 16 30 37 71 14,3462 52 85 5549 92 71 25 92 30 24 55
 18 9452 71 55 1992 71 67 85 34 67 71 11 6249 62 85 24 25 30 16 1924 22 30 24 16 19 2055 3089
 71 92 71 11 58.49 92 71 52 71 96 19 55 6224 25 92 62 14 55 18 7149 62 37 85 16 25 85,55 7124
 49 62 24 62 89 55 18 7149 92 85 34 96 71 22 1934 55 85 14 30 55 85 7124 71 92 19 71 90 55 18
 7311 92 30 89 85 16 71 96 71 94.

85 14 71 96 62 24 198562 92 58 37 85 71,3025 30 2537 71.49 92 85 96 85 22 55 18 7392 30 90
 14 71 92 62 3462 73 62 16 55 85 22 85 9455 62 37,34 16 62 92 62 94,25 30 92 14 30 55 55 18
 9467 34 71 94 46 30 92 24 25 85 9449 71 92 62 22 85 55 55 85 252452 34 58 14 2052 71 24 20 16
 25 30 14 8549 92 85 22 85 55 52 30 96 62 34,3016 30 25 37 7149 62 16 71 92 16 18 9449 85 24 16
 62 96 71 16-25 62 96 19 1689 62 96 71 7122 71 1452 34 30 52 46 30 16 85 96 71 16 55 71 11 6234
 62 90 92 30 24 16 30,55 6258 73 62 37 71 55 55 18 948524 14 30 90 30 55 55 18 94-85 14 71 55
 55 6216 30 25 62 7162 92 58 37 85 7114 62 37 55 6289 71 9062 24 62 89 18 7349 92 62 89 96 71
 1449 92 85 62 89 92 71 24 16853449 62 92 16 62 34 18 7316 92 58 86 62 89 30 73.34 24 7149 92

62 52 58 14 30 55 62.90 52 71 67 55 85 7149 62 96 85 46 30 852489 62 96 19 67 85 1449 62 52 62
 90 92 71 55 85 71 1462 16 55 62 24 20 16 24 202524 58 89 10 71 25 16 30 142430 34 16 62 14 30
 16 85 22 71 24 25 85 1462 92 58 37 85 71 1455 3049 96 71 22 71,90 30 16 6255 7162 24 62 89
 6255 30 34 62 92 62 22 71 55 55 18 9425 30 92 30 89 85 5585 96 8549 92 62 24 16 71 55 19 25 85
 9449 85 24 16 62 96 71 163425 30 92 14 30 55 713490 52 71 67 55 85 7314 71 24 16 30 7324 22
 85 16 30 45 16 24 2055 71 49 92 71 14 71 55 55 18 1430 16 92 85 89 58 16 62 1458 34 30 37 30
 45 86 71 11 6224 71 89 2025 30 89 30 96 19 71 92 62,49 85 24 19 14 71 55 55 62 11 6292 30 90
 92 71 67 71 55 85 2055 7116 92 71 89 58 45 1685,3462 89 86 71 14,49 62 52 62 90 92 71 55 85
 9455 71 34 18 90 18 34 30 45 16,49 62 25 302485 7349 62 14 62 86 19 4555 7124 62 16 34 62 92
 20 1622 71 11 62-16 62 55 71 90 30 25 62 55 55 62 11 62.

20.

16 7453 74 47 47 8531 85 66 74 29 58 55 7416 96 74 66 85 55 11 66 5896 11 12 91 74 74 50
 96 11 12 91 85 49 53 58 8547 11 33 74 26 74 31 2329 47 85 2645 29 85 55 74 29,96 11 12 33 85
 96 74 29,33 11 96 74 285829 74 12 96 11 47 55 11-66 85 68 28 74 29 35 53 28 5847 35 16 85 96
 47 74 29 96 85 33 85 91 91 23 85,47 29 85 96 28 11 21 18 58 8591 74 29 85 91 61 28 58 3366 11
 28 74 33,66 85 68 28 74 29 35 53 28 5829 96 85 33 85 9188 35 55 6166 5891 8529 55 74 96 74
 4933 58 96 74 29 74 49,47 55 11 96 85 91 61 28 58 8511 29 55 74 50 35 47 23,68 96 74 33 11 31
 91 23 8568 96 35 12 74 29 58 28 58-55 96 11 28 58,91 85 29 85 47 55 6174 55 28 35 31 1129 12
 43 29 53 11 43 47 435891 85 29 85 31 74 33 7428 35 31 1147 16 85 53 58 29 53 11 4316 74 79 11
 96 91 11 4333 11 53 58 91 11...31 74 29 74 66 61 91 7447 28 74 96 7474 9174 55 33 85 55 58
 66,88 55 7447 96 85 31 5829 47 85 68 7438 55 74 68 7496 11 12 91 74 74 50 96 11 12 58 4391
 8516 74 16 11 31 11 85 55 47 4391 5829 74 85 91 91 23 26,91 5816 74 66 58 45 85 49 47 28 58
 2633 11 53 58 91,29 74 74 50 18 852974 28 96 85 47 55 91 74 47 55 43 26,91 11 47 28 74 66 61
 28 7433 74 79 91 7447 35 31 58 55 6116 7455 74 33 35,88 55 7474 9129 58 31 85 664729 85 96
 26 74 55 35 96 23,91 8591 11 50 66 21 31 11 85 55 47 4391 5833 11 66 85 49 53 58 2616 96 58 12
 91 11 28 74 2988 96 85 12 29 23 88 11 49 18 58 91 23,28 11 28-55 74:33 74 50 58 66 61 91 23
 2616 11 55 96 35 66 85 49,16 74 47 55 74 2991 1174 50 74 88 58 91 85,16 96 74 29 85 96 28 5831
 74 28 35 33 85 91 55 74 29,12 11 47 55 11 29,50 66 74 28 16 74 47 55 74 29...91 58 88 85 68
 7416 74 31 74 50 91 74 68 74.47 58 8545 85 91 91 74 8591 11 50 66 21 31 85 91 58 8591 8591
 1153 35 55 28 3516 96 58 50 11 29 66 43 66 7474 16 55 58 33 58 12 33 11.

74 9116 74 47 33 74 55 96 85 6666 85 29 85 85-55 11 3374 5516 11 91 11 33 85 96 58 28 11
 91 23 74 55 26 74 31 58 66 1111 47 62 11 66 61 55 58 96 74 29 11 91 91 11 4331 74 96 74 68
 11,91 852916 96 58 33 85 9635 79 85,31 11 66 85 28 7491 8555 11 28 11 4374 79 58 29 66 85 91
 91 11 43.5835 55 23 28 11 66 11 47 6174 91 1116 96 43 33 85 26 74 91 61 28 742955 74 55 47 11

33 23 4968 74 96 74 31 74 28,68 31 8558 2653 85 47 55 85 96 28 1131 74 66 79 91 1150 23 66
1129 23 49 55 58 91 1133 85 47 55 91 74 68 7491 85 66 85 68 11 66 11.

21.

40 77 40,29 75 5875 28 75!15 61 75 23 40 52 672929 54 52 1115 75 65 58 5415 84 40 29 54
61 67 28 75 77 7558 84 11 18 77 75 61 67 28 54 35 40,77 52 1115 75 37 11 84 11 52 54 28 11,28
4028 11 29 49 37 75 35 75 13,35 29 40 52 84 40 58 28 75 1335 54 84 15 54 65 28 75 1315 75 37
58 40 13 11 28 58 1129 75 90 29 49 72 40 11 58 37 8015 18 72 35 4029 84 11 13 11 2815 11 84 29
75 4113 54 84 75 29 75 41,5415 75 5211 1137 58 29 75 61 75 1337 61 75 82 11 28 4015 54 84 40
13 54 52 35 4054 9029 75 29 37 1118 8237 58 40 84 54 28 28 49 4680 52 11 84,35 40 35 54 13
5415 40 61 54 61 5461 11 5890 4037 58 7552 7515 75 80 29 61 11 28 54 8028 4035 75 28 29 11
41 11 84 1158 40 35 54 4629 75 5858 84 11 46 52 20 41 13 75 29 75 35-37 20 84 84 11 40 61 54
37 58 54 65 11 37 35 75 1137 75 65 11 58 40 28 54 11,11 37 61 5429 52 18 13 40 58 67 37 80,28
7513 11 37 58 28 49 46,28 40 52 7515 75 61 40 77 40 58 67,29 15 75 61 28 1118 37 58 84 40 54
29 40 11 58.54 33 7528 40 77 61 80 52 28 7515 75 35 40 90 49 29 40 11 5852 75 33 61 11 37 58
67,15 84 75 80 29 61 11 28 28 18 2054 4652 11 84 82 40 29 75 412915 11 84 29 18 2013 54 84 75
29 18 20:28 1835 40 3582 11,75 28 4075 58 15 84 40 29 54 61 4028 4011 29 84 75 15 11 41 37 35
54 4192 84 75 28 5826 11 61 49 4137 58 84 11 61 35 75 29 49 4133 40 58 40 61 67 75 285458 75
84 82 11 37 58 29 11 28 28 7515 75 84 29 40 61 4075 58 28 75 72 11 28 54 803777 11 84 13 40 28
37 35 75 4154 13 15 11 84 54 11 41,4029 52 75 33 40 29 75 3575 52 182972 11 37 58 28 40 52 26
40 58 75 1333 11 84 11 77 75 29 49 1133 40 58 40 84 11 5475 52 28 75 77 7554 9029 75 11 28 28
49 4615 75 84 58 75 2926 11 61 49 4652 29 4065 40 37 4015 40 61 54 61 5415 7558 75 4158 75
65 35 111877 75 84 54 90 75 28 58 40,77 52 1135 40 35 75 13 18 -58 7533 52 54 58 11 61 67 28
75 13 1829 75 80 35 1115 75 65 18 52 54 61 37 8077 11 84 13 40 28 37 35 54 4135 84 11 41 37 11
84...75 33 19 11 35 58 54 29 28 75 37 58 5484 40 52 5437 58 75 54 5818 58 75 65 28 54 58 67,65
58 7529 7529 58 75 84 18 2013 54 84 75 29 18 2090 52 11 72 28 54 41,15 18 37 58 675428 11 29
11 61 54 35 54 4129 75 11 28 28 75-13 75 84 37 35 75 4192 61 75 5829 13 11 37 58 113737 75 20
90 28 54 35 40 13 5415 40 58 84 18 61 54 84 75 29 40 6115 84 54 61 11 77 40 20 23 54 1129 75
52 495415 40 84 1884 40 9029 84 75 52 1133 4952 40 82 1137 58 84 11 61 80 6115 7528 40 37 58
75 80 23 54 13,4028 1115 84 54 29 54 52 11 29 72 54 13 37 8015 75 52 29 75 52 28 49 1361 75 52
35 40 1335 84 54 77 37-13 40 84 54 28 11.

22.

56 9631 57 87 3756 7584 77 87 24 96 73 68 75,56 7550 37 16 42 68 77,7720 73 3737 49 56 77
39 77 87 37,39 73 3712 84 9616 91 64 56 91 87 37.75 56 84 73 16 91 68 94 75 7531 57 87 7544 16

37 84 73 577556 96 49 77 73 96 14 87 75 12 57:96 84 87 7556 7744 37 28 37 68 37 56 56 75 68
 9656 96 7356 7550 37 16 42 68 77,56 7584 77 87 24 96 73 68 75,96 84 87 7573 77 2673 37 87 41
 68 3784 77 87 24 96 73 68 7731 96 4950 37 16 42 68 7775 87 7550 37 16 42 37 6831 96 4984 77
 87 24 96 73 68 75-56 9673 3739 73 3756 9612 64 37 28 75 73 41,56 3728 77 35 9656 9644 16 75
 31 87 75 35 77 73 41 84 61.12 84 9644 16 96 35 56 75 9616 77 84 68 87 77 28 5787 96 73 61
 736839 96 16 73 91,1235 75 49 56 4184 87 96 28 91 96 7356 96 26 96 28 87 96 56 56 3744 16 96
 73 12 37 16 61 73 4149 77 44 77 84 56 37 1412 77 16 75 77 56 73.56 91 35 56 3768 77 6826 37 35
 56 3731 57 84 73 16 96 9637 73 84 82 28 7784 26 77 73 57 12 77 73 41 84 61,91 31 75 16 77 73
 41 84 616839 96 16 73 37 12 37 1426 77 73 96 16 7575 4950 37 16 37 28 68 77,12 84 73 91 44 77
 96 731284 75 87 9149 77 44 77 84 56 37 1412 77 16 75 77 56 7337 73 64 37 28 77...

12 96 84 4137 68 16 91 35 77 82 22 75 1426 75 1612 56 96 49 77 44 56 3784 73 77 8756 9644
 16 37 84 73 3739 91 35 75 26-12 16 77 35 28 96 31 56 57 26.44 37 28 37 49 16 96 12 77 73 4184
 87 96 28 37 12 77 87 3712 84 96 647512 84 61.44 41 82 22 75 6444 75 12 3784 37 87 28 77 73 75
 68 37 12-1273 37 26,39 73 3737 56 7556 9684 37 87 28 77 73 75 68 7512 37 12 84 96,7750 16 91
 44 44 7749 77 64 12 77 73 7775 4912 37 96 56 56 37 1468 37 56 73 16 16 77 49 12 96 28 68 75,44
 16 96 84 73 77 16 96 87 37 50 3750 37 84 73 75 56 75 39 56 37 50 3764 26 57 16 61-1273 37
 26,39 73 3737 5612 37 12 84 9656 9644 37 16 73 41 9675 87 7512 87 77 28 96 87 96 94,75 87
 7573 377528 16 91 50 37 961237 28 56 37 2687 75 94 96,7744 37 87 68 37 12 56 75 6849 28 96
 42 56 96 1473 77 14 56 37 1444 37 87 75 94 75 75.12 84 9612 37 49 26 37 35 56 37,68 37 50 28
 7791 84 87 37 12 87 96 56 56 37 50 3784 75 50 56 77 87 7756 96 7356 7791 84 87 37 12 87 96 56
 56 37 2626 96 84 73 96.

23.

22 10 75 6247 1074 10 24 88 47 39 35 66 15 75 58 10 47 64 53 5385 66 35 10 69 62 28 10 24
 5366 49 53 47 47 10 49 64 10 58 3928 22 88 17 10 79 47 88 1547 66 22 53.4447 10 85 17 10 28 53
 24 75 443551 66 75 58 53 47 53 64 88.35 10 3572 62 28 10 24 6647 8817 10 69,4466 80 37 80 10
 2469 49 88 75 3937 74 53 17 66 58 28 66 17 88 47 53 885385 66 35 66 15,22 37 28 75 58 28 10,35
 66 58 66 17 62 8853 75 85 62 58 62 28 10 88 79 39,66 35 10 69 10 28 79 53 75 392849 10 28 47
 6669 47 10 35 66 74 62 4274 88 75 58 10 42.79 53 17 66 35 53 8828 66 17 66 58 1072 62 24 5317
 10 75 85 10 42 47 37 58 62,37 75 10 49 39 72 1037 58 66 47 37 24 102875 37 74 88 17 35 10
 42.4428 66 79 88 242842 66 24 24,51 49 8858 37 74 10 47 47 62 8869 88 17 35 10 24 1069 62 72
 35 6666 58 17 10 31 10 24 5364 28 88 58 625349 88 58 10 24 5353 47 58 88 17 39 88 17 10.37 49
 53 28 53 58 88 24 39 47 66,47 6642 66 69 44 53 4747 8837 69 47 10 2474 88 47 44.66 4785 17 66
 58 44 47 37 2417 88 51 53 75 58 17 10 64 53 66 47 47 37 9735 47 53 51 37.4428 69 44 2417 37 22
 35 37,66 72 7 41 03 54 73 7 2485 88 17 662872 17 66 47 69 66 28 37 9722 88 17 47 53 24 39 47

53 64 3753,75 35 24 66 47 53 28 79 53 75 3947 10 49 17 10 75 35 17 62 58 62 74 5375 58 17 10
 47 53 64 10 74 53,75 58 66 24 35 47 37 24 75 447585 88 17 28 66 1553 6974 47 66 31 88 75 58 28
 1047 88 66 31 53 49 10 47 47 66 75 58 88 15,35 66 58 66 17 62 8885 66 49 75 58 88 17 88 51 10
 24 5374 88 47 4425 58 66 1547 66 22 39 97.74 66 8853 74 44,42 66 17 42 8824 37 53 7572 66 17
 42 88 75,72 62 24 6647 10 22 88 17 58 10 47 662835 47 53 51 88,5322 88 17 47 53 24 1088 80
 8847 8837 75 85 88 24 5328 62 75 66 42 47 37 58 39.

24.

6152 16 36 26 14 5416 45 24 29 4595 1129 36 95 86 36 16 29 451452 49 75 36 4797 36 93 95
 61 54 26 6197 3626 86 45 97 49 95 41 29 11 47.93 49 30 61 86 95 11 93 56 11 86 83 8995 36 47
 49 1695 11 37 36 93 14 54 26 6195 1130 86 36 16 36 4721 86 11 33 49,2636 29 95 11 47 1495
 1130 95 45 86 16 49 95 95 14 8993 30 36 16 14 29,33 11 54 29 14 891471 11 52 16 36 19 49 95 95
 83 89,36 52 95 49 26 49 95 95 83 8952 11 54 98 26 86 16 11 93 36 89;75 93 49,29 11 2997 36 47
 95 14 54 36 26 4147 95 49,26 86 36 61 54 1197 54 61 33 95 11 6126 29 11 47 49 89 29 11.21 86
 3652 83 54 1126 11 47 11 6152 36 54 41 19 11 6129 36 47 95 11 86 1130 75 36 26 86 14 95 14 56
 49.6186 36 54 29 95 45 5493 30 49 16 41,36 95 1197 36 93 93 11 54 11 26 41.97 36 93 97 36 86
 36 54 29 36 4775 36 16 49 54 1154 98 26 86 16 11.3049 4952 49 71 33 11 54 36 26 86 95 36 4726
 30 49 86 496145 71 95 11 5426 49 52 61.95 1145 71 29 36 8933 49 54 49 71 95 36 8929 16 36 30
 11 86 1454 49 33 11 5461,97 36 26 86 11 16 49 30 19 14 891436 52 16 98 71 75 19 14 89,1416 11
 71 75 54 61 93 83 30 11 5454 49 97 95 14 95 4595 1197 36 86 36 54 29 49.6145 26 54 83 19 11
 5475 36 54 36 26.95 4926 36 30 26 49 4747 36 89-52 49 7136 52 49 16 86 36 95 36 30,95 49 97 16
 14 61 86 95 83 89,97 36 37 36 33 14 8995 1147 11 75 95 14 86 36 78 36 95 95 45 9871 11 97 14
 26 41.

-45 93 14 30 14 86 49 54 41 95 36,-26 29 11 71 11 5436 95,-95 11 2693 30 36 491447 8336 93
 95 36.30 97 16 36 24 49 47,30 3626 95 4995 14 24 86 3695 4926 97 36 26 36 52 95 3630 83 71 30
 11 86 4145 93 14 30 54 49 95 14 49.6116 36 52 29 3626 97 16 36 26 14 54:

-71 95 11 24 14 86,30 26 4921 86 3626 36 95?

-97 16 14 24 49 4797 36 26 54 49 93 95 14 8926 36 95.-33 49 26 86 36 4736 9597 36 29 11 71
 11 5495 1197 45 26 86 36 89 97 45 71 83 16 49 29,

26 86 36 61 30 19 14 8995 1147 16 11 47 36 16 95 36 8929 16 83 19 29 4995 36 24 95 36 75
 3626 86 36 54 14 29 11.-86 49 52 49 97 16 14 93 49 86 26 61,95 11 30 49 16 95 36 49,45 30 14 93
 49 86 4147 95 36 33 49 26 86 30 3626 95 36 30,97 16 49 33 93 4924 49 4793 36 52 49 16 49 19 41
 26 6193 3621 86 36 8995 36 24 14.29 11 29 36 4926 49 75 36 93 95 6124 14 26 54 3697 3686 30
 36 49 47 4529 11 54 49 95 93 11 16 98?

25.

48 84 13 3394 13 48 42 33 46 82,84 13 82 4894 82 46 84 33 4213 88 82 84 16 46 1625 8250 17
 481342 61 37 78 50 511682 42 13 82 84 16 46 1650 48 17 341376 82 25 82 1672 82 46 48 69 17
 82 28 82,28 84 4851 75 4875 84 33 46 1646 33 84 33 17,75 33 37 82 13 17 341638 48 37 17 16 46
 33.82 1713 58 94 25 33 69 58 13 33 4676 82 75 48 46 33 17 16 34,163476 82 25 33 69 58 13 33
 4648 50 5113 94 48,38 42 8217 1648 94 42 781350 16 37 48.1376 37 16 28 82 37 64 17 4817 48
 17 33 13 16 94 42 17 82 28 8272 58 46 8294 82 72 37 33 17 8213 94 48,38 42 8284 82 13 48 46 82
 94 78 76 82 13 16 84 33 42 7851 75 4851 94 82 76 64 16 501638 42 8269 37 34 4217 58 17 4869
 84 37 33 13 94 42 13 51 61 21 16 48:28 82 37 82 84 33,75 33 37 25 16 481688 82 46 82 84 17 58
 4894 42 37 33 17 58,94 82 25 37 82 13 16 21 33,94 25 37 58 42 58 481369 48 50 17 58 8828 46 51
 72 16 17 33 88,72 82 37 82 69 84 34 21 16 4850 82 37 3425 82 37 33 72 46 16,82 37 51 84 16
 3413 82 91 17 58,16 17 94 42 37 51 50 48 17 42 5813 37 33 38 48 13 33 17 16 341650 51 69 58 25
 16,76 46 48 17 16 42 48 46 78 17 58 8875 48 17 21 16 17,17 48 76 82 84 13 16 75 17 58 4869 13
 48 69 84 581676 46 33 17 48 42 58,25 37 33 94 25 16,25 82 42 82 37 58 50 1676 82 46 78 69 51
 61 42 94 3417 48 13 48 37 17 58 48,25 82 28 84 3376 16 64 51 4294 13 82 1650 48 37 69 25 16 48

25 33 37 42 16 17 58,37 33 94 42 48 17 16 341650 16 17 48 37 334 65 894 8213 94 48 50 1616
 8894 82 25 37 82 13 48 17 17 58 50 1669 33 50 48 38 33 42 48 46 78 17 58 50 1694 13 82 91 94
 42 13 33 50 16,94 48 37 48 72 37 34 17 58 8833 17 28 48 46 82 13,38 48 9188 46 48 72-88 13 33
 46 331676 37 48 13 82 69 17 48 94 48 17 16 4828 82 94 76 82 84 33,37 33 69 84 33 38 5117 33 28
 37 33 841364 25 82 46 33 88,19 16 28 51 37 5876 42 16 981698 33 37 48 91,88 37 33 17 34 21 16
 48 94 341394 33 50 82 5094 48 37 84 98 4876 16 37 33 50 16 84,42 48 17 7872 58 25 33,17 3325
 82 42 82 37 82 5076 82 25 82 16 42 94 3469 48 50 46 34,1637 58 72 58,17 3325 82 42 82 37 82
 9194 42 82 16 42

72 58 25,76 51 94 42 58 17 1613 94 48 50 16 46 82 94 42 16 13 82 28 8272 82 28 33.82 1751
 13 16 84 48 4613 48 21 1617 48 82 76 16 94 51 48 50 58 48,42 33 25 16 48,25 33 2551 46 16 98
 58,82 94 13 48 21 48 17 17 58 4828 33 69 82 13 58 50 1637 82 75 25 33 50 16,1625 16 42 33,25
 82 42 82 37 58 9151 50 16 37 33 48 4276 37 1669 13 51 25 33 8838 48 46 82 13 48 38 48 94 25 82
 28 8228 82 46 82 94 33.

Пример решения самостоятельной работы

15 вариант

Скольльзящая перестановка

Текст для расшифровки: ПАРИИВИАРЗ_БРА_ИСТЬЛТОЕК

Текст содержит 25 символов, т.е. записываем его в таблицу 5×5.

–	П	А	Р	И
И	В	И	А	Р
З	–	Б	<i>Р</i>	<i>А</i>
–	И	С	<i>Т</i>	<i>Ь</i>
Л	Т	О	<i>Е</i>	<i>К</i>

Расшифровку следует проводить меня порядок столбцов.

Воспользуемся таблицей сочетаемости букв. 4 и 5 столбцы идут друг за другом, т.к. биграммы РА, ТЬ и ЕК наиболее распространенные. 2 и 4 столбец идут друг за другом, т.к. биграммы ИТ и ТЕ тоже распространены. По выше перечисленным признакам не трудно догадаться, что столбцы будут располагаться в следующем порядке: 2,4,5,3,1. И зашифрованной фразой будет: *При аварии разбить стекло.*

В данном случае дешифровать текст можно было обычным методом перебора, не обращаясь к таблице сочетаемости букв.

Шифр двойной перестановки

Текст для расшифровки: ЗШАФИПРАЛОЕНЖ_ОЪН_ДАРВОНА.

Текст содержит 25 символов, т.е. записываем его в таблицу 5×5.

З	Ш	А	Ф	И
П	Р	А	Л	О
Е	Н	Ж	–	О
Ь	Н	–	Д	А
Р	В	О	Н	А

Расшифровку следует проводить меня порядок столбцов и строк.

Глядя на зашифрованный текст по первым пяти символам можно сразу предположить, что столбцы меняются следующим образом: 1,3,2,5,4.

З	А	Ш	И	Ф
---	---	---	---	---

П	А	Р	О	Л
Е	Ж	Н	О	–
Ь	–	Н	А	Д
Р	О	В	А	Н

Глядя на вторую таблицу можно, также, без труда определить порядок строк: 2,4,3,1,5.

Расшифрованный текст: Пароль_надежно_зашифрован.

Шифр простой замены

Расшифрованный текст:

ЧЕМ ДАЛЬШЕ, ТЕМ СИЛЬНЕЕ ОН ЧУВСТВОВАЛ НЕШУТОЧНОЕ РАЗДРАЖЕНИЕ, ПОРОЮ ПЕРЕХОДИВШЕЕ В ПРИЛИВЫ ЗЛОСТИ-ОТТОГО, ЧТО ОНИ ЧЕТВЕРО СУТОК, ОБРАТИВШИЕСЬ В ЗРЕНИЕ И СЛУХ, ТОРЧАЛИ В ЧАЩОБЕ, КАК ДИКИЕ ОБЕЗЬЯНЫ ИЗ БРАЗИЛИИ, ОТТОГО, ЧТО ПОДВЕРНУЛСЯ ТУПОЙ КАЙМАН, С ОДИНАКОВЫМ УСЕРДИЕМ НАПАДАВШИЙ И НА ЛЕСНУЮ СВИНЬЮ, И НА ОТЛИЧНОГО ПАРНЯ С ДРУГОГО КОНТИНЕНТА А В ЭТО ВРЕМЯ ТЕ, НА БАЗЕ, ЖИЛИ В СВОЕ УДОВОЛЬСТВИЕ, СПАЛИ НА ЧИСТЕНЗЫЗИХ ПРОСТЫНКАХ В КОНДИЦИОНИРОВАННОЙ ПРОХЛАДЕ, ПРИНИМАЛИ ДУШ, ЖРАЛИ НА ЗАВТРАК ФРУКТЫ, ДЖЕМ И БИФСТЕКСЫ В ТРИ ПАЛЬЦА ТОЛЩИНОЙ, И ОКНА ТАК УЮТНО СВЕТИЛИСЬ, И МУЗЫКА ИГРАЛА, И ФУТБОЛ ПО ТЕЛЕВИЗОРУ

НИЧЕГО В ЭТОЙ ЗЛОСТИ НЕ БЫЛО ПЛОХОГО, НАОБОРОТ - ТАКОЙ НАСТРОЙ КАК РАЗ И ПРИДАЕТ БОЕВОГО КУРАЖА...

А ПОТОМ ПРИШЕЛ КОНЕЦ И ПОСТОРОННИМ МЫСЛЯМ И БЕЗДЕЛЬЮ. МОРСКОЙ ЗМЕЙ НАКОНЕЦ-ТО ПОДАЛ ЗНАК, КОТОРОГО ОНИ ЖДАЛИ ЧЕТВЕРО СУТОК, И ЭТО БЫЛО СЛОВНО МЕДНЫЙ РЕВ БОЕВОЙ ТРУБЫ, ЭТО ОЗНАЧАЛО, ЧТО НАЧАЛИСЬ РАБОТЫ, И НИЧЕГО УЖЕ НЕ ИЗМЕНИТЬ, НЕ ОСТАНОВИТЬ, НЕ ПЕРЕИГРАТЬ...

7. ШИФРОВАНИЕ С СЕКРЕТНЫМ КЛЮЧОМ

7.1. Теория шифров с секретным ключом

Блочные и поточные системы шифрования

Используемые в настоящее время системы шифрования делятся на два класса: блочные и поточные системы. Основным критерий такого разделения – мощность алфавита, над знаками которого производится операция шифрования. Если открытый текст перед шифрованием разбивается на блоки, состоящие из нескольких знаков, то есть исходное сообщение обрабатывается блоками, то мы имеем дело с блочным шифром. Если каждый знак сообщения шифруется отдельно, то такой шифр – поточный.

Разделение шифров на поточные и блочные связано с алгоритмическими и техническими особенностями реализации шифрующих преобразований, использующими возможности существующей элементной базы (разрядность процессоров, быстродействие микросхем, объем памяти компьютера). При увеличении мощности алфавита необходимо исследовать, прежде всего, вопросы о выборе преобразований, реализуемых криптосхемой, и способе их практической реализации, влияющем на эффективность функционирования криптосхемы с точки зрения эксплуатационных характеристик.

Естественно, что точное значение мощности алфавита, начиная с которого шифр следует считать уже не поточным, а блочным, назвать нельзя. Более того, с развитием техники эта характеристика меняется в сторону увеличения. Например, в настоящее время используются 16- и 32-разрядные процессоры, а перспективная шифровальная техника проектируется уже на 64-разрядных процессорах. Поэтому при построении поточных шифров могут быть использованы алфавиты мощности 2^{32} и 2^{64} .

Принципы построения блочных шифров

Как правило, алфавитом, на котором действует блочный шифр, является множество двоичных векторов-блоков открытого текста одинаковой длины (64, 128 и т. д.).

К.Шеннон сформулировал общий принцип построения шифрующих преобразований – принцип "*перемешивания*". Суть его состоит в требовании, чтобы применение шифрующего преобразования к наборам аргументов, отличающихся в незначительном числе позиций, приводило к существенному изменению результата.

Блочные шифры реализуются путем многократного применения к блокам открытого текста некоторых базовых преобразований. Базовые преобразования должны удовлетворять ряду требований, обусловленных тем, что они, во-первых, должны быть просто реализуемым, в том числе, программным способом на ЭВМ, и, во-вторых, при небольшом числе итераций давать аналитически сложные преобразования.

Обычно используются базовые преобразования двух типов – сложные в криптографическом отношении локальные преобразования над отдельными частями шифруемых блоков и простые преобразования, переставляющие между собой части шифруемых блоков. В криптографической литературе первые преобразования получили название "*перемешивающих*", а вторые – "*рассеивающих*". Качественно можно сказать, что перемешивание усложняет восстановление взаимосвязи статистических и аналитических свойств открытого и шифрованного текстов, а рассеивание распространяет влияние одного знака открытого текста на большое число знаков шифртекста, что позволяет сгладить влияние статистических свойств открытого текста на свойства шифртекста.

Алгоритм шифрования выполняет некоторое число циклов (итераций). Каждый цикл состоит в применении преобразований первого и второго типов. Такой принцип построения дает возможность реализовать каждый цикл шифрования с использованием однотипных узлов, а также выполнять расшифрование путем обработки данных в обратном направлении.

Удобной моделью для реализации базовых преобразований служат *регистры сдвига*. При этом рассеивающие преобразования определяются функциями обратной связи, а перемешивающие – сдвигами информации в регистре.

Получили распространение алгоритмы, в которых осуществляются преобразования над векторами, представляющими собой левую и правую половины содержимого регистра сдвига. Для построения таких алгоритмов часто используется конструкция, называемая *сетью Фейстеля* (Feistel Network) (рисунок 4.1).

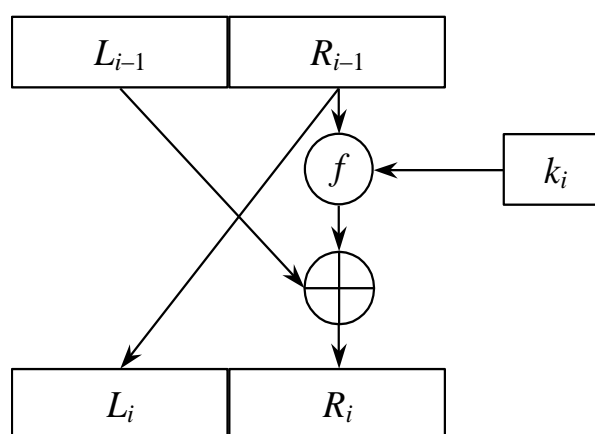


Рис. 7.1. Сеть Фейстеля

Преобразование, реализуемое сетью Фейстеля в i -м цикле шифрования, имеет вид

$$\begin{cases} L_i = R_{i-1}, \\ R_i = L_{i-1} \oplus f_i(R_{i-1}, k_i), \end{cases} \quad (5)$$

где L_{i-1} и R_{i-1} – левая и правая части входного блока L_i и R_i – левая и правая части блока, являющегося результатом зашифрования входного блока на ключе k_i с помощью функции f_i .

Алгоритм шифрования реализуется несколькими итерациями преобразования сети Фейстеля с использованием ключа k . При этом очередная (i -я) итерация использует в качестве входного блока результат предыдущей итерации и ключ k_i , вычисляемый определенным образом по ключу k . Функция f_i может зависеть или не зависеть от номера итерации.

Ценность преобразований подобного вида заключается в том, что даже если f_i не является обратимой функцией, преобразование сети Фейстеля обратимо. В самом деле, из (4.1) сразу следует, что

$$\begin{cases} L_{i-1} = R_i \oplus f_i(L_i, k_i), \\ R_{i-1} = L_i. \end{cases}$$

Примеры блочных шифров

Стандарт шифрования данных ГОСТ 28147-89

В России установлен единый алгоритм криптографического преобразования данных – ГОСТ 28147-89. Этот алгоритм предназначен для аппаратной и программной реализации, удовлетворяет необходимым криптографическим требованиям и не накладывает ограничений на степень секретности защищаемой информации. Алгоритм реализует шифрование 64-битовых блоков данных с помощью 256-битового ключа.

ГОСТ 28147-89 содержит описание алгоритмов нескольких уровней. На самом верхнем находятся практические алгоритмы, предназначенные для шифрования массивов данных и выработки для них имитовставки. Все они опираются на три алгоритма низшего уровня, называемые в тексте ГОСТа *циклами*. Будем называть эти фундаментальные алгоритмы *базовыми циклами*, чтобы отличать их от всех прочих циклов. Они имеют следующие названия и обозначения, последние приведены в скобках и смысл их будет объяснен позже:

- цикл зашифрования (32-З);
- цикл расшифрования (32-Р);
- цикл выработки имитовставки (16-З).

В свою очередь, каждый из базовых циклов представляет собой многократное повторение одной единственной процедуры, которую для определенности будем называть *основным шагом криптопреобразования*.

Таким образом, чтобы разобраться в ГОСТе, надо понять три следующие вещи:

1. что такое *основной шаг* криптопреобразования;
2. как из *основных шагов* складываются *базовые циклы*;

3. как из трех **базовых циклов** складываются все практические алгоритмы ГОСТа.

Прежде чем перейти к изучению этих вопросов, следует поговорить о ключевой информации, используемой алгоритмами ГОСТа. В соответствии с принципом Кирхгофа, которому удовлетворяют все современные известные широкой общественности шифры, именно ее секретность обеспечивает секретность зашифрованного сообщения. В ГОСТе ключевая информация состоит из двух структур данных. Помимо собственно **ключа**, необходимого для всех шифров, она содержит еще и **таблицу замен**:

1. **Ключ** является массивом из восьми 32-битных элементов кода и обозначается символом K : $K = \{K_i\}_{0 \leq i \leq 7}$. В ГОСТе элементы ключа используются как 32-разрядные целые числа без знака: $0 \leq K_i < 2^{32}$. Таким образом, размер ключа составляет $32 \times 8 = 256$ бит или 32 байта.

2. **Таблица замен** является матрицей 8×16 , содержащей 4-битовые элементы, которые можно представить в виде целых чисел от 0 до 15. Строки **таблицы замен** называются **узлами замен**, они должны содержать различные значения, то есть каждый **узел замен** должен содержать 16 различных чисел от 0 до 15 в произвольном порядке. Таблица замен обозначается символом H : $H = \{H_{i,j}\}_{\substack{0 \leq i \leq 7 \\ 0 \leq j \leq 15}}$, $0 \leq H_{i,j} \leq 15$. Таким образом, общий объем таблицы замен равен: 8 узлов \times 16 элементов/узел \times 4 бита/элемент = 512 бит или 64 байта.

Основной шаг криптопреобразования.

Основной шаг криптопреобразования по своей сути является оператором, определяющим преобразование 64-битового блока данных. Дополнительным параметром этого оператора является 32-битовый блок, в качестве которого используется какой-либо элемент ключа. Схема алгоритма основного шага приведена на рисунке 7.2.

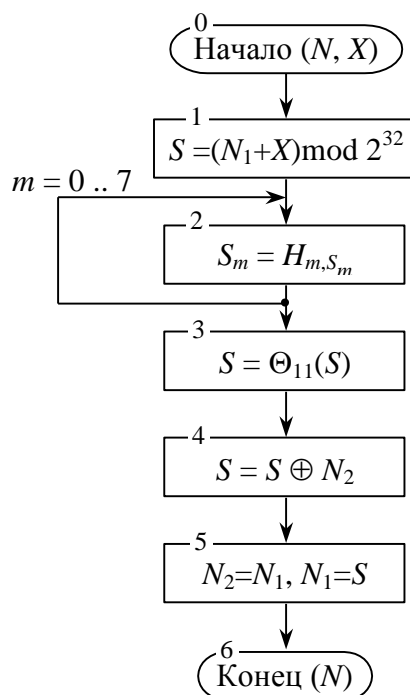


Рис. 7.2. Схема основного шага криптопреобразования алгоритма ГОСТ 28147-89

Ниже даны пояснения к алгоритму основного шага:

0. Определяет исходные данные для основного шага криптопреобразования:

- N – преобразуемый 64-битовый блок данных, в ходе выполнения шага его младшая (N_1) и старшая (N_2) части обрабатываются как отдельные 32-битовые целые числа без знака. Таким образом, можно записать $N=(N_1, N_2)$.
- X – 32-битовый элемент ключа;

1. Сложение с ключом. Младшая половина преобразуемого блока складывается по модулю 2^{32} с используемым на шаге элементом ключа, результат передается на следующий шаг;

2. Поблочная замена. 32-битовое значение, полученное на предыдущем шаге, интерпретируется как массив из восьми 4-битовых блоков кода:

$$S = (S_0, S_1, S_2, S_3, S_4, S_5, S_6, S_7).$$

Далее значение каждого из восьми блоков заменяется на новое, которое выбирается по таблице замен следующим образом: значение блока S_i заменяется на S_i -тый по порядку элемент (нумерация с нуля) i -того узла замен (т.е. i -той строки таблицы замен, нумерация также с нуля). Другими словами, в качестве замены для значения блока выбирается элемент из таблицы замен с номером строки, равным номеру заменяемого блока, и номером столбца, равным значению заменяемого

блока как 4-битового целого неотрицательного числа. Теперь становится понятным размер таблицы замен: число строк в ней равно числу 4-битных элементов в 32-битном блоке данных, то есть восьми, а число столбцов равно числу различных значений 4-битного блока данных, равному как известно 2^4 , шестнадцати.

3. Циклический сдвиг на 11 бит влево. Результат предыдущего шага сдвигается циклически на 11 бит в сторону старших разрядов и передается на следующий шаг. На схеме алгоритма символом Θ_{11} обозначена функция циклического сдвига своего аргумента на 11 бит в сторону старших разрядов.

4. Побитовое сложение: значение, полученное на шаге 3, побитно складывается по модулю 2 со старшей половиной преобразуемого блока.

5. Сдвиг по цепочке: младшая часть преобразуемого блока сдвигается на место старшей, а на ее место помещается результат выполнения предыдущего шага.

6. Полученное значение преобразуемого блока возвращается как результат выполнения алгоритма основного шага криптопреобразования.

Базовые циклы криптографических преобразований.

ГОСТ относится к классу блочных шифров, то есть единицей обработки информации в нем является блок данных. Следовательно, вполне логично ожидать, что в нем будут определены алгоритмы для криптографических преобразований, то есть для зашифрования, расшифрования и «учета» в контрольной комбинации одного блока данных. Именно эти алгоритмы и называются *базовыми циклами* ГОСТа, что подчеркивает их фундаментальное значение для построения этого шифра.

Базовые циклы построены из *основных шагов* криптографического преобразования, рассмотренного в предыдущем разделе. В процессе выполнения основного шага используется только один элемент ключа, в то время как ключ ГОСТ содержит восемь таких элементов. Следовательно, чтобы ключ был использован полностью, каждый из базовых циклов должен многократно выполнять основной шаг с различными его элементами. Вместе с тем кажется вполне естественным, что в каждом базовом цикле все элементы ключа должны быть использованы одинаковое число раз, по соображениям стойкости шифра это число должно быть больше одного.

Все сделанные выше предположения, опирающиеся просто на здравый смысл, оказались верными. Базовые циклы заключаются в многократном выполнении *основного шага* с использованием разных элементов ключа и отличаются друг от друга только числом

повторения шага и порядком использования ключевых элементов. Ниже приведен этот порядок для различных циклов.

1. Цикл зашифрования 32-З:

$K_0, K_1, K_2, K_3, K_4, K_5, K_6, K_7, K_0, K_1, K_2, K_3, K_4, K_5, K_6, K_7, K_0, K_1, K_2, K_3, K_4, K_5, K_6, K_7, K_7, K_6, K_5, K_4, K_3, K_2, K_1, K_0.$

2. Цикл расшифрования 32-Р:

$K_0, K_1, K_2, K_3, K_4, K_5, K_6, K_7, K_7, K_6, K_5, K_4, K_3, K_2, K_1, K_0, K_7, K_6, K_5, K_4, K_3, K_2, K_1, K_0, K_7, K_6, K_5, K_4, K_3, K_2, K_1, K_0.$

3. Цикл выработки имитовставки 16-З:

$K_0, K_1, K_2, K_3, K_4, K_5, K_6, K_7, K_0, K_1, K_2, K_3, K_4, K_5, K_6, K_7.$

Каждый из циклов имеет собственное буквенно-цифровое обозначение, соответствующее шаблону « $n-X$ », где первый элемент обозначения (n), задает число повторений основного шага в цикле, а второй элемент обозначения (X), буква, задает порядок зашифрования («З») или расшифрования («Р») в использовании ключевых элементов. Этот порядок нуждается в дополнительном пояснении:

Цикл расшифрования должен быть обратным циклу зашифрования, то есть последовательное применение этих двух циклов к произвольному блоку должно дать в итоге исходный блок, что отражается следующим соотношением:

$$C_{32-P}(C_{32-Z}(T))=T,$$

где T – произвольный 64-битный блок данных,

$C_X(T)$ – результат выполнения цикла X над блоком данных T .

Для выполнения этого условия для алгоритмов, подобных ГОСТу, необходимо и достаточно, чтобы порядок использования ключевых элементов соответствующими циклами был взаимно обратным. В справедливости записанного условия для рассматриваемого случая легко убедиться, сравнив приведенные выше последовательности для циклов 32-З и 32-Р. Из сказанного вытекает одно интересное следствие: свойство цикла быть обратным другому циклу является взаимным, то есть цикл 32-З является обратным по отношению к циклу 32-Р. Другими словами, зашифрование блока данных теоретически может быть выполнено с помощью цикла расшифрования, в этом случае расшифрование блока данных должно быть выполнено циклом зашифрования. Из двух взаимно обратных циклов любой может быть использован для зашифрования, тогда второй должен быть использован для расшифрования данных, однако стандарт ГОСТ28147-89 закрепляет роли за циклами и не предоставляет пользователю права выбора в этом вопросе.

Цикл выработки имитовставки вдвое короче циклов шифрования, порядок использования ключевых элементов в нем такой же, как в первых 16 шагах цикла

зашифрования, в чем нетрудно убедиться, рассмотрев приведенные выше последовательности, поэтому этот порядок в обозначении цикла кодируется той же самой буквой «З».

Схемы базовых циклов приведены на рисунках 4.3 а-в. Каждый из них принимает в качестве аргумента и возвращает в качестве результата 64-битный блок данных, обозначенный на схемах N .

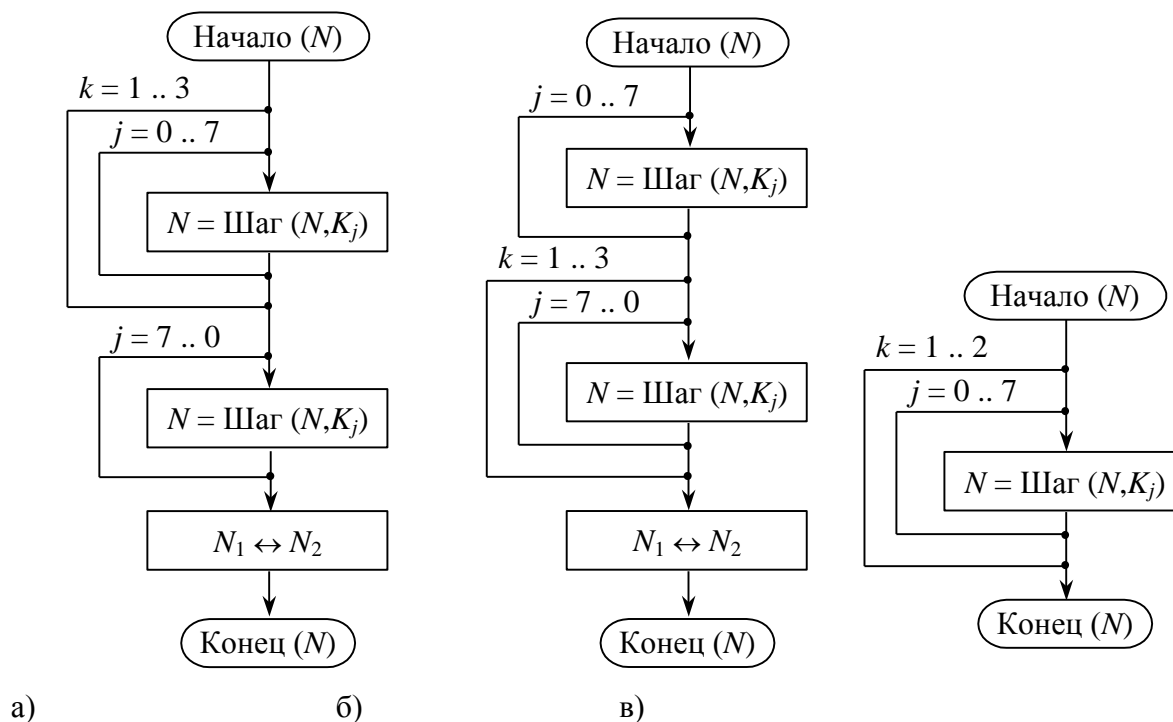


Рис. 7.3. Схемы цикла зашифрования 32-З (а), расшифрования 32-Р (б) и выработки имитовставки 16-З (в)

Символ Шаг(N, X) обозначает выполнение основного шага криптопреобразования для блока N с использованием ключевого элемента X . Между циклами шифрования и вычисления имитовставки есть еще одно отличие, не упомянутое выше: в конце базовых циклов шифрования старшая и младшая часть блока результата меняются местами, это необходимо для их взаимной обратимости.

Основные режимы шифрования.

Для решения разнообразных криптографических задач блочные шифры используют в нескольких режимах работы. ГОСТ 28147-89 предусматривает три следующих режима шифрования данных:

- простая замена,
- гаммирование,

- гаммирование с обратной связью,

и один дополнительный режим выработки имитовставки.

В любом из этих режимов данные обрабатываются блоками по 64 бита, на которые разбивается массив, подвергаемый криптографическому преобразованию, именно поэтому ГОСТ относится к блочным шифрам. Однако в двух режимах гаммирования есть возможность обработки неполного блока данных размером меньше 8 байт, что существенно при шифровании массивов данных с произвольным размером, который может быть не кратным 8 байтам.

Введем некоторые обозначения:

- $T_o, T_{ш}$ – массивы соответственно открытых и зашифрованных данных;
- $T_i^o, T_i^ш$ – i -тые по порядку 64-битные блоки соответственно открытых и зашифрованных данных: $T_o = (T_1^o, T_2^o, \dots, T_n^o), T_{ш} = (T_1^ш, T_2^ш, \dots, T_n^ш), 1 \leq i \leq n$.

Последний блок может быть неполным: $|T_i^o| = |T_i^ш| = 64$ при $1 \leq i < n, 1 \leq |T_n^o| = |T_n^ш| \leq 64$;

- n – число 64-битных блоков в массиве данных;
- Π_X – функция преобразования 64-битного блока данных по алгоритму базового цикла «X»;

Теперь опишем основные режимы шифрования.

Простая замена

Зашифрование в данном режиме заключается в применении цикла 32-З к блокам открытых данных, расшифрование – цикла 32-Р к блокам зашифрованных данных. Это наиболее простой из режимов, 64-битовые блоки данных обрабатываются в нем независимо друг от друга. Схемы алгоритмов зашифрования и расшифрования в режиме простой замены приведены на рисунках 2.4 а и 2.4 б соответственно.

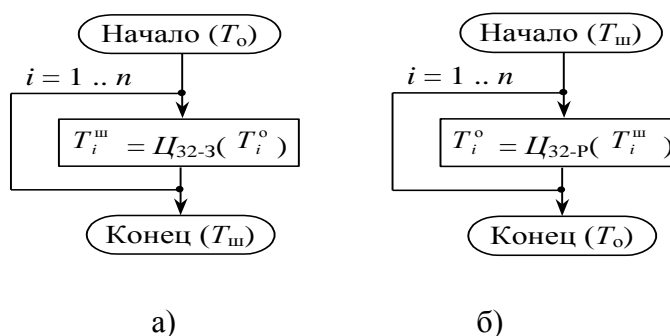


Рис. 7.4. Алгоритмы зашифрования (а) и расшифрования (б) данных в режиме простой замены

Размер массива открытых или зашифрованных данных, подвергающийся соответственно зашифрованию или расшифрованию, должен быть кратен 64 битам: $|T_o| = |T_{ш}| = 64 \times n$, после выполнения операции размер полученного массива данных не изменяется.

Режим шифрования простой заменой имеет следующие особенности:

1. Так как блоки данных шифруются независимо друг от друга и от их позиции в массиве, при зашифровании двух одинаковых блоков открытого текста получаются одинаковые блоки шифротекста и наоборот. Отмеченное свойство позволит криптоаналитику сделать заключение о тождественности блоков исходных данных, если в массиве зашифрованных данных ему встретились идентичные блоки, что является недопустимым для серьезного шифра.

2. Если длина шифруемого массива данных не кратна 8 байтам или 64 битам, возникает проблема, чем и как дополнять последний неполный блок данных массива до полных 64 бит. Эта задача не так проста, как кажется на первый взгляд, поскольку очевидные решения типа «дополнить неполный блок нулевыми битами» или, более обще, «дополнить неполный блок фиксированной комбинацией нулевых и единичных битов» могут при определенных условиях дать в руки криптоаналитика возможность методами перебора определить содержимое этого самого неполного блока, и этот факт означает снижение стойкости шифра. Кроме того, длина шифротекста при этом изменится, увеличившись до ближайшего целого, кратного 64 битам, что часто бывает нежелательным.

На первый взгляд, перечисленные выше особенности делают практически невозможным использование режима простой замены, ведь он может применяться только для шифрования массивов данных с размером кратным 64 битам, не содержащим повторяющихся 64-битных блоков. Кажется, что для любых реальных данных гарантировать выполнение указанных условий невозможно. Это почти так, но есть одно очень важное исключение: вспомните, что размер ключа составляет 32 байта, а размер таблицы замен – 64 байта. Кроме того, наличие повторяющихся 8-байтовых блоков в ключе или таблице замен будет говорить об их весьма плохом качестве, поэтому в реальных ключевых элементах такого повторения быть не может. Таким образом мы выяснили, что режим простой замены вполне подходит для шифрования ключевой информации, тем более, что прочие режимы для этой цели менее удобны, поскольку требуют наличия дополнительного синхронизирующего элемента данных – синхропосылки. ГОСТ предписывает использовать режим простой замены исключительно для шифрования ключевых данных.

Гаммирование

Как же можно избавиться от недостатков режима простой замены? Для этого необходимо сделать возможным шифрование блоков с размером менее 64 бит и обеспечить зависимость блока шифротекста от его номера, иными словами, *рандомизировать* процесс шифрования. В ГОСТе это достигается двумя различными способами в двух режимах шифрования, предусматривающих *гаммирование*. *Гаммирование* – это наложение (снятие) на открытые (зашифрованные) данные криптографической гаммы, то есть последовательности элементов данных, вырабатываемых с помощью некоторого криптографического алгоритма, для получения зашифрованных (открытых) данных. Для наложения гаммы при шифровании и ее снятия при расшифровании должны использоваться взаимно обратные бинарные операции, например, сложение и вычитание по модулю 2^{64} для 64-битных блоков данных. В ГОСТе для этой цели используется операция побитного сложения по модулю 2, поскольку она является обратной самой себе и к тому же наиболее просто реализуется. Гаммирование решает обе упомянутые проблемы; во первых, все элементы гаммы различны для реальных шифруемых массивов и, следовательно, результат шифрования даже двух одинаковых блоков в одном массиве данных будет различным. Во вторых, хотя элементы гаммы и вырабатываются одинаковыми порциями в 64 бита, использоваться может и часть такого блока с размером, равным размеру шифруемого блока.

Теперь перейдем непосредственно к описанию режима гаммирования. Гамма для этого режима получается следующим образом: с помощью некоторого алгоритмического рекуррентного генератора последовательности чисел (РГПЧ) вырабатываются 64-битные блоки данных, которые далее подвергаются преобразованию по циклу 32-3, то есть шифрованию в режиме простой замены, в результате получают блоки гаммы. Благодаря тому, что наложение и снятие гаммы осуществляется при помощи одной и той же операции побитового исключающего или, алгоритмы шифрования и расшифрования в режиме гаммирования идентичны, их общая схема приведена на рисунке 2.5.

РГПЧ, используемый для выработки гаммы, является рекуррентной функцией:

$$\Omega_{i+1} = f(\Omega_i),$$

где Ω_i – элементы рекуррентной последовательности,

f – функция преобразования.

Следовательно, неизбежно возникает вопрос о его инициализации, то есть об элементе Ω_0 . В действительности, этот элемент данных является параметром алгоритма для режимов

гаммирования, на схемах он обозначен как S , и называется в криптографии *синхропосылкой*, а в нашем ГОСТе – *начальным заполнением* одного из регистров шифрователя. По определенным соображениям разработчики ГОСТа решили использовать для инициализации РГПЧ не непосредственно синхропосылку, а результат ее преобразования по циклу 32-3:

$$\Omega_0 = U_{32-3}(S).$$

Последовательность элементов, вырабатываемых РГПЧ, целиком зависит от его начального заполнения, то есть элементы этой последовательности являются функцией своего номера и начального заполнения РГПЧ:

$$\Omega_i = f_i(\Omega_0),$$

$$\text{где } f_i(X) = f(f_{i-1}(X)), f_0(X) = X.$$

С учетом преобразования по алгоритму простой замены добавляется еще и зависимость от ключа:

$$\Gamma_i = U_{32-3}(\Omega_i) = U_{32-3}(f_i(\Omega_0)) = U_{32-3}(f_i(U_{32-3}(S))) = \varphi_i(S, K),$$

где Γ_i – i -тый элемент гаммы,

K – ключ.

Таким образом, последовательность элементов гаммы для использования в режиме гаммирования однозначно определяется ключевыми данными и синхропосылкой. Естественно, для обратимости процедуры шифрования в процессах за- и расшифрования должна использоваться одна и та же синхропосылка. Из требования уникальности гаммы, невыполнение которого приводит к катастрофическому снижению стойкости шифра, следует, что для шифрования двух различных массивов данных на одном ключе необходимо обеспечить использование различных синхропосылок. Это приводит к необходимости хранить или передавать синхропосылку по каналам связи вместе с зашифрованными данными, хотя в отдельных особых случаях она может быть предопределена или вычисляться особым образом, если исключается шифрование двух массивов на одном ключе.

Теперь подробно рассмотрим РГПЧ, используемый в ГОСТе для генерации элементов гаммы. Прежде всего надо отметить, что к нему не предъявляются требования обеспечения каких-либо статистических характеристик вырабатываемой последовательности чисел. РГПЧ спроектирован разработчиками ГОСТа исходя из необходимости выполнения следующих условий:

- период повторения последовательности чисел, вырабатываемой РГПЧ, не должен сильно (в процентном отношении) отличаться от максимально возможного при заданном размере блока значения 2^{64} ;
- соседние значения, вырабатываемые РГПЧ, должны отличаться друг от друга в каждом байте, иначе задача криптоаналитика будет упрощена;

- РГПЧ должен быть достаточно просто реализуем как аппаратно, так и программно на наиболее распространенных типах процессоров, большинство из которых, как известно, имеют разрядность 32 бита.

Исходя из перечисленных принципов создатели ГОСТа спроектировали весьма удачный РГПЧ, имеющий следующие характеристики:

- в 64-битовом блоке старшая (Ω_i^1) и младшая (Ω_i^0) части обрабатываются независимо друг от друга: $\Omega_i = (\Omega_i^0, \Omega_i^1)$, $|\Omega_i^0| = |\Omega_i^1| = 32$, $\Omega_{i+1}^0 = f_{РГПЧ1}(\Omega_i^0)$, $\Omega_{i+1}^1 = f_{РГПЧ2}(\Omega_i^1)$, фактически, существуют два независимых РГПЧ для старшей и младшей частей блока.

- рекуррентные соотношения для старшей и младшей частей следующие:

$$\Omega_{i+1}^0 = (\Omega_i^0 + C_1) \bmod 2^{32}, \text{ где } C_1 = 1010101_{16};$$

$$\Omega_{i+1}^1 = (\Omega_i^1 + C_2 - 1) \bmod (2^{32} - 1) + 1, \text{ где } C_2 = 1010104_{16};$$

Нижний индекс в записи числа означает его систему счисления, таким образом, константы, используемые на данном шаге, записаны в 16-ричной системе счисления.

- период повторения последовательности для младшей части составляет 2^{32} , для старшей части $2^{32}-1$, для всей последовательности период составляет $2^{32} \cdot (2^{32}-1)$, доказательство этого факта, весьма несложное, получите сами. Первая формула из двух реализуется за одну команду, вторая, несмотря на ее кажущуюся громоздкость, за две команды на всех современных 32-разрядных процессорах.

Схема алгоритма шифрования в режиме гаммирования приведена на рисунке 2.5.

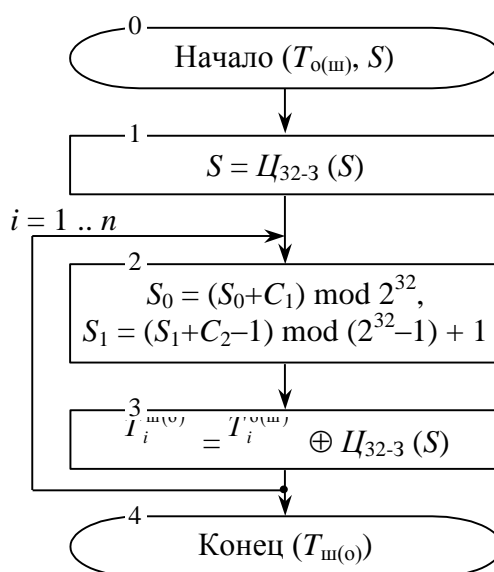


Рис. 7.5. Алгоритм зашифрования (расшифрования) данных в режиме гаммирования

Ниже изложены пояснения к схеме:

0. Определяет исходные данные для основного шага криптопреобразования:

- $T_{o(ш)}$ – массив открытых (зашифрованных) данных произвольного размера, подвергаемый процедуре зашифрования (расшифрования), по ходу процедуры массив подвергается преобразованию порциями по 64 бита;
- S – *синхросылка*, 64-битный элемент данных, необходимый для инициализации генератора гаммы;

1. Начальное преобразование синхросылки, выполняемое для ее «рандомизации», то есть для устранения статистических закономерностей, присутствующих в ней, результат используется как начальное заполнение РГПЧ;

2. Один шаг работы РГПЧ, реализующий его рекуррентный алгоритм. В ходе данного шага старшая (S_1) и младшая (S_0) части последовательности данных вырабатываются независимо друг от друга;

3. Гаммирование. Очередной 64-битный элемент, выработанный РГПЧ, подвергается процедуре зашифрования по циклу 32–3, результат используется как элемент гаммы для зашифрования (расшифрования) очередного блока открытых (зашифрованных) данных того же размера.

4. Результат работы алгоритма – зашифрованный (расшифрованный) массив данных.

Ниже перечислены особенности гаммирования как режима шифрования.

1. Одинаковые блоки в открытом массиве данных дадут при зашифровании различные блоки шифротекста, что позволит скрыть факт их идентичности.

2. Поскольку наложение гаммы выполняется побитно, шифрование неполного блока данных легко выполнимо как шифрование битов этого неполного блока, для чего используются соответствующие биты блока гаммы. Так, для зашифрования неполного блока в 1 бит можно использовать любой бит из блока гаммы.

3. Синхросылка, использованная при зашифровании, каким-то образом должна быть передана для использования при расшифровании. Это может быть достигнуто следующими путями:

- хранить или передавать синхросылку вместе с зашифрованным массивом данных, что приведет к увеличению размера массива данных при зашифровании на размер синхросылки, то есть на 8 байт;

- использовать predetermined значение синхропосылки или вырабатывать ее синхронно источником и приемником по определенному закону, в этом случае изменение размера передаваемого или хранимого массива данных отсутствует;

Оба способа дополняют друг друга, и в тех редких случаях, где не работает первый, наиболее употребительный из них, может быть использован второй, более экзотический. Второй способ имеет гораздо меньшее применение, поскольку сделать синхропосылку predetermined можно только в том случае, если на данном комплекте ключевой информации шифруется заведомо не более одного массива данных, что бывает в редких случаях. Генерировать синхропосылку синхронно у источника и получателя массива данных также не всегда представляется возможным, поскольку требует жесткой привязки к чему-либо в системе. Так, здравая на первый взгляд идея использовать в качестве синхропосылки в системе передачи зашифрованных сообщений номер передаваемого сообщения не подходит, поскольку сообщение может потеряться и не дойти до адресата, в этом случае произойдет десинхронизация систем шифрования источника и приемника. Поэтому в рассмотренном случае нет альтернативы передаче синхропосылки вместе с зашифрованным сообщением.

С другой стороны, можно привести и обратный пример. Допустим, шифрование данных используется для защиты информации на диске, и реализовано оно на низком уровне, для обеспечения независимого доступа данные шифруются по секторам. В этом случае невозможно хранить синхропосылку вместе с зашифрованными данными, поскольку размер сектора нельзя изменить, однако ее можно вычислять как некоторую функцию от номера считывающей головки диска, номера дорожки (цилиндра) и номера сектора на дорожке. В этом случае синхропосылка привязывается к положению сектора на диске, которое вряд ли может измениться без переформатирования диска, то есть без уничтожения данных на нем.

Режим гаммирования имеет еще одну интересную особенность. В этом режиме биты массива данных шифруются независимо друг от друга. Таким образом, каждый бит шифротекста зависит от соответствующего бита открытого текста и, естественно, порядкового номера бита в массиве:

$$t_i^m = t_i^o \oplus \gamma_i = f(t_i^o, i).$$

Из этого вытекает, что изменение бита шифротекста на противоположное значение приведет к аналогичному изменению бита открытого текста на противоположный:

$$\bar{t}_i^{\text{III}} = t_i^{\text{III}} \oplus 1 = (t_i^{\text{O}} \oplus \gamma_i) \oplus 1 = (t_i^{\text{O}} \oplus 1) \oplus \gamma_i = \bar{t}_i^{\text{O}} \oplus \gamma_i,$$

где \bar{t} обозначает инвертированное по отношению к t значение бита ($\bar{0} = 1, \bar{1} = 0$).

Данное свойство дает злоумышленнику возможность воздействуя на биты шифротекста вносить предсказуемые и даже целенаправленные изменения в соответствующий открытый текст, получаемый после его расшифрования, не обладая при этом секретным ключом. Это иллюстрирует хорошо известный в криптологии факт, что **«секретность и аутентичность суть различные свойства шифров»**. Иными словами, свойства шифров обеспечивать защиту от несанкционированного ознакомления с содержимым сообщения и от несанкционированного внесения изменений в сообщение являются независимыми и лишь в отдельных случаях могут пересекаться. Сказанное означает, что существуют криптографические алгоритмы, обеспечивающие определенную секретность зашифрованных данных и при этом никак не защищающие от внесения изменений и наоборот, обеспечивающие аутентичность данных и никак не ограничивающие возможность ознакомления с ними. По этой причине рассматриваемое свойство режима гаммирования не должно рассматриваться как его недостаток.

Гаммирование с обратной связью

Данный режим очень похож на режим гаммирования и отличается от него только способом выработки элементов гаммы – очередной элемент гаммы вырабатывается как результат преобразования по циклу 32-3 предыдущего блока зашифрованных данных, а для зашифрования первого блока массива данных элемент гаммы вырабатывается как результат преобразования по тому же циклу синхропосылки. Этим достигается сцепление блоков – каждый блок шифротекста в этом режиме зависит от соответствующего и всех предыдущих блоков открытого текста. Поэтому данный режим иногда называется **гаммированием с сцеплением блоков**. На стойкость шифра факт сцепления блоков не оказывает никакого влияния.

Схема алгоритмов за- и расшифрования в режиме гаммирования с обратной связью приведена на рисунке 2.6.

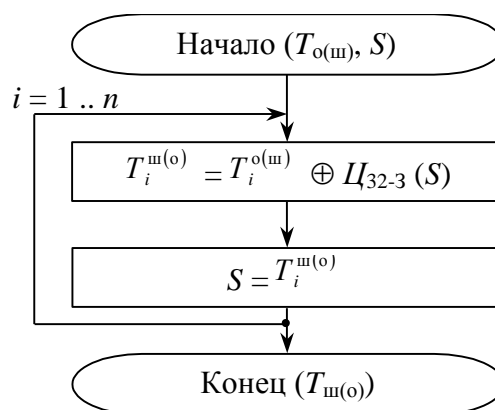


Рис. 7.6. Алгоритм зашифрования (расшифрования) данных в режиме гаммирования с обратной связью

Шифрование в режиме гаммирования с обратной связью обладает теми же особенностями, что и шифрование в режиме обычного гаммирования, за исключением влияния искажений шифротекста на соответствующий открытый текст. Для сравнения запишем функции расшифрования блока для обоих упомянутых режимов:

$$T_i^o = T_i^m \oplus \Gamma_i \text{ – гаммирование;}$$

$$T_i^o = T_i^m \oplus U_{32-3}(T_{i-1}^m) \text{ – гаммирование с обратной связью.}$$

Если в режиме обычного гаммирования изменения в определенных битах шифротекста влияют только на соответствующие биты открытого текста, то в режиме гаммирования с обратной связью картина несколько сложнее. Как видно из соответствующего уравнения, при расшифровании блока данных в режиме гаммирования с обратной связью, блок открытых данных зависит от соответствующего и предыдущего блоков зашифрованных данных. Поэтому, если внести искажения в зашифрованный блок, то после расшифрования искаженными окажутся два блока открытых данных – соответствующий и следующий за ним, причем искажения в первом случае будут носить тот же характер, что и в режиме гаммирования, а во втором случае – как в режиме простой замены. Другими словами, в соответствующем блоке открытых данных искаженными окажутся те же самые биты, что и в блоке зашифрованных данных, а в следующем блоке открытых данных все биты независимо друг от друга с вероятностью $1/2$ изменят свои значения.

Выработка имитовставки к массиву данных.

В предыдущих разделах мы обсудили влияние искажения зашифрованных данных на соответствующие открытые данные и установили, что при расшифровании в режиме простой замены соответствующий блок открытых данных оказывается искаженным

непредсказуемым образом, а при расшифровании блока в режиме гаммирования изменения предсказуемы. В режиме гаммирования с обратной связью искаженными оказываются два блока, один предсказуемым, а другой непредсказуемым образом. Значит ли это, что с точки зрения защиты от навязывания ложных данных режим гаммирования является плохим, а режимы простой замены и гаммирования с обратной связью хорошими? Ни в коем случае. При анализе данной ситуации необходимо учесть то, что непредсказуемые изменения в расшифрованном блоке данных могут быть обнаружены только в случае избыточности этих самых данных, причем чем больше степень избыточности, тем вероятнее обнаружение искажения. Очень большая избыточность имеет место, например, для текстов на естественных и искусственных языках, в этом случае факт искажения обнаруживается практически неизбежно. Однако в других случаях, например, при искажении сжатых звуковых образов, мы получим просто другой образ, который сможет воспринять наше ухо. Искажение в этом случае останется необнаруженным, если, конечно, нет априорной информации о характере звука. Вывод здесь такой: поскольку способность некоторых режимов шифрования обнаруживать искажения, внесенные в зашифрованные данные, существенным образом опирается на наличие и степень избыточности шифруемых данных, эта способность не является имманентным (внутренне присущим) свойством соответствующих режимов и не может рассматриваться как их достоинство.

Для решения задачи обнаружения искажений в зашифрованном массиве данных с заданной вероятностью в ГОСТе предусмотрен дополнительный режим криптографического преобразования – выработка имитовставки. Имитовставка – это контрольная комбинация, зависящая от открытых данных и секретной ключевой информации. Целью использования имитовставки является обнаружение всех случайных или преднамеренных изменений в массиве информации. Проблема, изложенная в предыдущем пункте, может быть успешно решена с помощью добавления к зашифрованным данным имитовставки. Для потенциального злоумышленника две следующие задачи практически неразрешимы, если он не владеет ключевой информацией:

- вычисление имитовставки для заданного открытого массива информации;
- подбор открытых данных под заданную имитовставку;

Схема алгоритма выработки имитовставки приведена на рисунке 2.7.

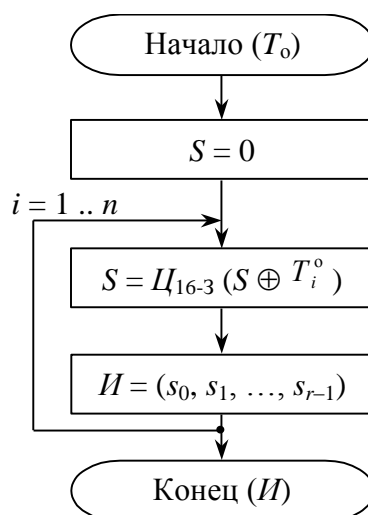


Рис. 7.7. Алгоритм выработки имитовставки для массива данных

В качестве имитовставки берется часть блока, полученного на выходе, обычно 32 его младших бита. При выборе размера имитовставки надо принимать во внимание, что вероятность успешного навязывания ложных данных равна величине $2^{-|И|}$ на одну попытку подбора. При использовании имитовставки размером 32 бита эта вероятность равна $2^{-32} \approx 0,23 \times 10^{-9}$.

Американский стандарт шифрования данных DES

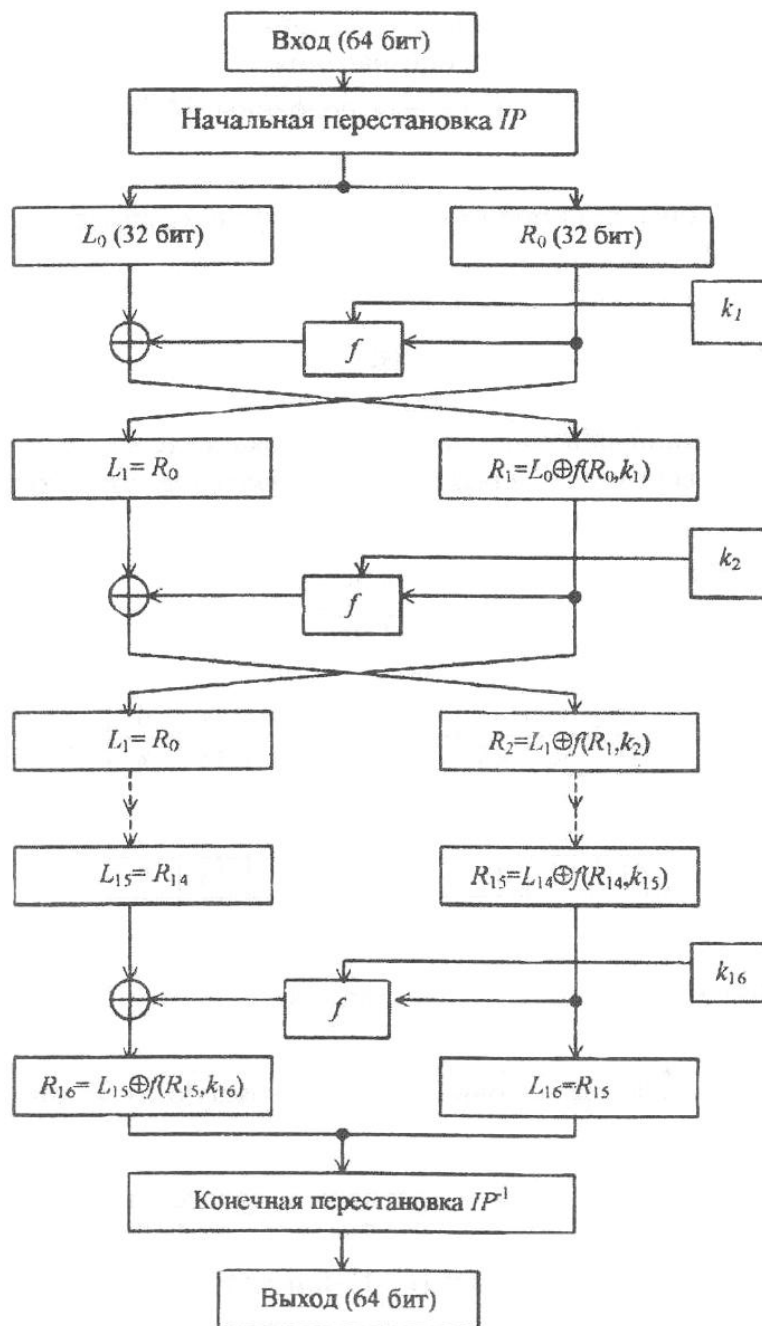
Стандарт шифрования данных DES (*Data Encryption Standard*) опубликован Национальным бюро стандартов США в 1977 г. В 1980 г. этот алгоритм был принят Национальным институтом стандартов и технологий США (НИСТ) в качестве стандарта шифрования данных для защиты от несанкционированного доступа к важной, но несекретной информации в государственных и коммерческих организациях США.

Основные достоинства алгоритма DES:

- простота ключевой системы (используется только один ключ длиной 56 бит);
- относительная простота алгоритма обеспечивает высокую скорость аппаратной и программной реализации;
- достаточно высокая криптографическая стойкость алгоритма шифрования.

Алгоритм DES использует комбинацию подстановок и перестановок. DES осуществляет шифрование 64-битовых блоков данных с помощью 56-битового ключа k , в котором значащими являются 56 бит (остальные 8 бит – проверочные биты для контроля на четность). Дешифрование в DES является операцией, обратной шифрованию, и выполняется

путем повторения операций шифрования в обратной последовательности. Структура алгоритма DES изображена на рисунке 2.8.



L_i, R_i – левая и правая половины 64-битового блока;

\oplus – операция побитового сложения блоков по модулю 2 (XOR);

k_i – 48-битовые ключи, получаемые из исходного ключа k ;

f – функция шифрования;

IP – начальная перестановка степени 64.

Рис. 7.8. Структура алгоритма DES

Процесс шифрования состоит в начальной перестановке битов 64-битового входного блока, шестнадцати циклах шифрования и, наконец, конечной перестановке битов.

Отметим, что все приводимые ниже таблицы являются стандартными и должны использоваться при реализации алгоритма DES в неизменном виде. Все перестановки и коды в таблицах подобраны разработчиками таким образом, чтобы максимально затруднить процесс вскрытия шифра путем подбора ключа.

При зашифровании очередного блока T его биты подвергаются *начальной перестановке* IP согласно приводимой таблице 7.1.

Таблица 7.1. Начальная перестановка IP

8	0	2	4	6	8	0	
0	2	4	6	8	0	2	
2	4	6	8	0	2	4	
4	6	8	0	2	4	6	
7	9	1	3	5	7		
9	1	3	5	7	9	1	
1	3	5	7	9	1	3	
3	5	7	9	1	3	5	

При этом бит 58 блока T становится битом 1, бит 50 – битом 2 и т. д. Полученный после перестановки блок $IP(T)$ разделяется на две половины: L_0 , состоящую из 32 старших бит, и R_0 , состоящую из 32 младших бит.

Затем выполняется итеративный процесс шифрования, состоящий из 16 циклов преобразования Фейстеля. Пусть $T_{i-1} = L_{i-1}R_{i-1}$ – результат $(i - 1)$ -й итерации.

Тогда результат i -й итерации $T_i = L_iR_i$ определяется формулами

$$\begin{cases} L_i = R_{i-1}, \\ R_i = L_{i-1} \oplus f_i(R_{i-1}, k_i), i = \overline{1,16}. \end{cases} \quad (6)$$

Функция f называется *функцией шифрования*. Ее аргументами являются 32-битовый вектор R_{i-1} и 48-битовый ключ k_i , который является результатом преобразования 56-битового ключа шифра k . Результатом последней итерации является блок $T_{16} = L_{16}R_{16}$. По окончании шифрования осуществляется восстановление позиций битов применением к T_{16} обратной перестановки IP^{-1} .

При расшифровании данных все действия выполняются в обратном порядке, при этом вместо соотношений (4.2) следует пользоваться соотношениями

$$\begin{cases} R_{i-1} = L_i, \\ L_{i-1} = R_i \oplus f_i(L_i, k_i), i = \overline{16,1}, \end{cases}$$

пользуясь которыми можно "спуститься" от L_{16} и R_{16} к R_0 и L_0 .

Схема вычисления значения функции шифрования $f(R_{i-1}, k_i)$ изображена на рисунке 2.9.

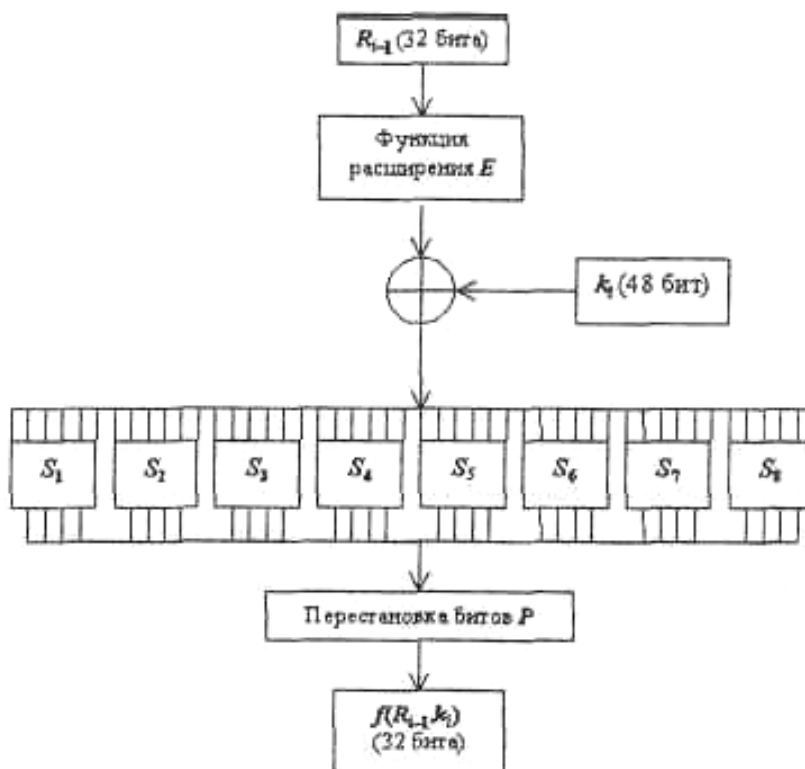


Рис. 7.9. Схема вычисления функции шифрования f

Для вычисления значения функции f используются:

- функция расширения E ;
- преобразование S , составленное из восьми преобразований S -блоков S_1, S_2, \dots, S_8 ;

- перестановка P .

Аргументами функции f являются вектор R_{i-1} (32 бита) и вектор k_i (48 бит). Функция E "расширяет" 32-битный вектор R_{i-1} до 48-битного вектора $E(R_{i-1})$ путем дублирования некоторых битов вектора R_{i-1} , при этом порядок следования битов в $E(R_{i-1})$ указан в таблице 2.2.

Таблица 7.2. Функция расширения E

32	1	2	3	4	5
4	5	6	7	8	9
8	9	10	11	12	13
12	13	14	15	16	17
16	17	18	19	20	21
20	21	22	23	24	25
24	25	26	27	28	29
28	29	30	31	32	1

Первые три бита в $E(R_{i-1})$ – это соответственно биты 32, 1 и 2 вектора R_{i-1} , а последние три бита – это соответственно биты 31, 32, 1 вектора R_{i-1} .

Полученный результат складывается побитно по модулю 2 с текущим значением ключа k_i и затем представляется в виде восьми последовательных 6-битовых блоков B_1, B_2, \dots, B_8 :

$$E(R_{i-1}) \oplus k_i = B_1 \dots B_8.$$

Далее каждый из блоков B_j трансформируется в 4-битовый блок B'_j с помощью подходящей таблицы S -блока S_j , список которых приведен в таблице 3.

Таблица 7.3. Функция преобразования S

		НОМЕР СТОЛБЦА																
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
Н О	0	14	4	13	1	2	15	11	8	3	10	6	12	5	9	0	7	S_1
	1	0	15	7	4	14	2	13	1	10	6	12	11	9	5	3	8	
	2	4	1	4	8	13	6	2	11	15	12	9	7	3	10	5	0	
	3	15	12	8	2	4	9	1	7	5	11	3	14	10	0	6	13	
М Е	0	15	1	8	14	6	11	3	4	9	7	2	13	12	0	5	10	S_2
	1	3	13	4	7	15	2	8	14	12	0	1	10	6	9	11	5	
	2	0	14	7	11	10	4	13	1	5	8	12	6	9	3	2	15	
	3	13	8	10	1	3	15	4	2	11	6	7	12	0	5	14	9	
Р	0	10	0	9	14	6	3	15	5	1	13	12	7	11	4	2	8	S_3
	1	13	7	0	9	3	4	6	10	2	8	5	14	12	11	15	1	
	2	13	6	4	9	8	15	3	0	11	1	2	12	5	10	14	7	
	3	1	10	13	0	6	9	8	7	4	15	14	3	11	5	2	12	
С Т	0	7	13	14	3	0	6	9	10	1	2	8	5	11	12	4	15	S_4
	1	13	8	11	5	6	15	0	3	4	7	2	12	1	10	14	9	
	2	10	6	9	0	12	11	7	13	15	1	3	14	5	2	8	4	
	3	3	15	0	6	10	1	13	8	9	4	5	11	12	7	2	14	
Р О	0	2	12	4	1	7	10	11	6	8	5	3	15	13	0	14	9	S_5
	1	14	11	2	12	4	7	13	1	5	0	15	10	3	9	8	6	
	2	4	2	1	11	10	13	7	8	15	9	12	5	6	3	0	14	
	3	11	8	12	7	1	14	2	13	6	15	0	9	10	4	5	3	
К И	0	12	1	10	15	9	2	6	8	0	13	3	4	14	7	5	11	S_6
	1	10	15	4	2	7	12	9	5	6	1	13	14	0	11	3	8	
	2	9	14	15	5	2	8	12	3	7	0	4	10	1	13	1	6	
	3	4	3	2	12	9	5	15	10	11	14	1	7	6	0	8	13	
	0	4	11	2	14	15	0	8	13	3	12	9	7	5	10	6	1	S_7
	1	13	0	11	7	4	9	1	10	14	3	5	12	2	15	8	6	
	2	1	4	11	13	12	3	7	14	10	15	6	8	0	5	9	2	
	3	6	11	13	8	1	4	10	7	9	5	0	15	14	2	3	12	
	0	13	2	8	4	6	15	11	1	10	9	3	14	5	0	12	7	S_8
	1	1	15	13	8	10	3	7	4	12	5	6	11	0	14	9	2	
	2	7	11	4	1	9	12	14	2	0	6	10	13	15	3	5	8	
	3	2	1	14	7	4	10	8	13	15	12	9	0	3	5	6	11	

Преобразование блока B_j и B'_j осуществляется следующим образом. Пусть, например, B_2 равен 111010. Первый и последний разряды B_2 являются двоичной записью числа a , $0 \leq a \leq 3$. Аналогично средние 4 разряда представляют число b , $0 \leq b \leq 15$. В нашем примере $a = 2$, $b = 13$.

Строки и столбцы таблицы S_2 пронумерованы числами a и b . Таким образом, пара (a, b) однозначно определяет число, находящееся на пересечении строки с номером a и столбца с номером b . В данном случае это число равно 3. Записывая его в двоичной форме, получаем B'_2 , равный 0011.

Значение $f(R_{i-1}, k_i)$ теперь получается применением перестановки битов P , заданной таблицей к результирующему 32-битовому блоку B'_1, B'_2, \dots, B'_8 .

Таблица 7.4. Функция P перестановки битов

16	7	20	21
29	12	28	17
1	15	23	26
5	18	31	10
2	8	24	14
32	27	3	9
19	13	30	6
22	11	4	25

На каждой итерации используется текущее значение ключа k_i (48 бит), получаемое из исходного ключа k следующим образом.

Сначала пользователи выбирают сам ключ k , содержащий 56 случайных значащих битов. Восемь битов, находящихся в позициях 8,16,...,64, добавляются в ключ таким образом, чтобы каждый байт содержал нечетное число единиц. Это используется для обнаружения ошибок при обмене и хранении ключей. Значащие 56 бит ключа подвергаются перестановке, приведенной в таблице 7.5.

Таблица 7.5. Функция первоначальной подготовки ключа

57			4	3	2	1	9
	9	1	3	5	7		
1			5	4	3	2	18
	8	0	2	4	6		
10			5	5	4	3	27
		9	1	3	5		
19			3	6	5	4	36
	1		0	2	4		
63			4	3	3	2	15
	5	7	9	1	3		
7			5	4	3	3	22
	2	4	6	8	0		

		6	5	4	3	
14		1	3	5	7	29
			2	2	1	
21	3	5	8	0	2	4

Эта перестановка определяется двумя блоками C_0 и D_0 по 28 бит в каждом (они занимают соответственно верхнюю и нижнюю половины таблицы). Так, первые три бита в C_0 есть соответственно 57, 49, 41 биты ключа. Затем индуктивно определяются блоки C_i и $D_i = \overline{1,16}$.

Если уже определены C_{i-1} и D_{i-1} , то C_i и D_i получаются из них одним или двумя левыми циклическими сдвигами согласно таблице 7.6.

Таблица 7. 6. Таблица сдвигов для вычисления ключа

i											0	1	2	3	4	5	6
Число сдвигов																	

Теперь определим ключи k_i , $1 \leq i \leq 16$. Ключ k_i состоит из 48 битов, выбираемых из битов блока $C_i D_i$ согласно таблице 7. Первыми тремя битами в k_i являются биты 14, 17, 11 из блока $C_i D_i$. Отметим, что 8 из 56 бит (с номерами 9, 18, 22, 25, 35, 38, 43, 54) из $C_i D_i$ отсутствуют в k_i .

Таблица 7.7. Функция завершающей обработки ключа

14	17	11	24	1	5
3	28	15	6	21	10
23	19	12	4	26	8
16	7	27	20	13	2
41	52	31	37	47	55
30	40	51	45	33	48
44	49	39	56	34	53
46	42	50	36	29	32

Официально DES являлся стандартом шифрования данных до 31 декабря 1998 г. В 1997 г. был объявлен конкурс на новый стандарт, который был назван AES (Advanced Encryption Standard). 2 октября 2000 г. Национальный институт стандартов и технологий (НИСТ) США объявил победителя конкурса AES. Однако для того, чтобы этот алгоритм завоевал мировое

признание, необходимы серьезные исследования его свойств специалистами различных стран.

Основные режимы шифрования

Алгоритм DES вполне подходит как для шифрования, так и для аутентификации данных. Он позволяет непосредственно преобразовывать 64-битовый входной открытый текст в 64-битовый выходной зашифрованный текст, однако данные редко ограничиваются 64 разрядами.

Чтобы использовать алгоритм DES для решения разнообразных криптографических задач, разработаны четыре режима:

- *режим электронной кодовой книги (ECB – Electronic Code Book);*
- *режим сцепления блоков шифра (CBC – Cipher Block Chaining);*
- *режим обратной связи по шифртексту (CFB – Cipher Feed Back);*
- *режим обратной связи по выходу (OFB – Output Feed Back).*

Лабораторная работа №1

Изучение стандарта криптографической защиты AES (Advanced Encryption Standart)

Цель работы

Ознакомиться с криптографическим стандартом шифрования AES и его особенностями, изучить различными режимами блочного шифрования.

Краткие теоретические сведения

Историческая справка

В 1997 г. Национальный институт стандартов и технологий США (NIST) объявил о начале программы по принятию нового стандарта криптографической защиты - стандарта XXI в. для закрытия важной информации правительственного уровня на замену существующему с 1974 г. алгоритму DES, самому распространенному криптоалгоритму в мире.

Требования к кандидатам были следующие:

- криптоалгоритм должен быть открыто опубликован;
- криптоалгоритм должен быть симметричным блочным шифром, допускающим размеры ключей в 128, 192 и 256 бит;

- криптоалгоритм должен быть предназначен как для аппаратной, так и для программной реализации;
- криптоалгоритм должен быть доступен для открытого использования в любых продуктах, а значит, не может быть запатентован, в противном случае патентные права должны быть аннулированы;
- криптоалгоритм подвергается изучению по следующим параметрам: стойкости, стоимости, гибкости, реализуемости в smart-картах.

В финал конкурса вышли следующие алгоритмы: MARS, TWOFISH и RC6 (США), RINDAEL (Бельгия), SERPENT (Великобритания, Израиль, Норвегия). По своей структуре TWOFISH является классическим шифром Фейстеля; MARS и RC6 можно отнести к модифицированным шифрам Фейстеля, в них используется новая малоизученная операция циклического "прокручивания" битов слова на число позиций, изменяющихся в зависимости от шифруемых данных и секретного ключа; RINDAEL и SERPENT являются классическими SP-сетями. MARS и TWOFISH имеют самую сложную конструкцию, RINDAEL и RC6 - самую простую.

В октябре 2000 г. конкурс завершился. Победителем был признан бельгийский шифр RINDAEL, как имеющий наилучшее сочетание стойкости, производительности, эффективности реализации и гибкости. Его низкие требования к объему памяти делают его идеально подходящим для встроенных систем. Авторами шифра являются Йон Дэмен (Joan Daemen) и Винсент Рюмен (Vincent Rijmen), начальные буквы фамилий которых и образуют название алгоритма - RINDAEL.

После этого NIST начал подготовку предварительной версии Федерального Стандарта Обработки Информации (Federal Information Processing Standard - FIPS) и в феврале 2001 г. опубликовал его на сайте <http://csrc.nist.gov/encryption/aes/>. В течение 90-дневного периода открытого обсуждения предварительная версия FIPS пересматривалась с учетом комментариев, после чего начался процесс исправлений и утверждения. Наконец 26 ноября 2001 г. была опубликована окончательная версия стандарта FIPS-197, описывающего новый американский стандарт шифрования AES. Согласно этому документу стандарт вступил в силу с 26 мая 2002 г.

Блочный криптоалгоритм RIJNDAEL и стандарт AES

Высочайшую надежность AES NIST подтверждает астрономическими числами. 128битный ключ обеспечивает 340 андециллионов ($340 \cdot 1036$) возможных комбинаций, а 256битный ключ увеличивает это число до $11 \cdot 1076$. Для сравнения, старый алгоритм DES, дает общее число комбинаций в $72 \cdot 1015$. На их перебор у специально построенной машины "DES Cracker" уходит несколько часов. Но даже если бы она делала это всего за одну секунду, то на перебор 128битного ключа машина потратила бы 149 триллионов лет. Между тем, возраст всей Вселенной ученые оценивают в менее, чем 20 миллиардов лет.

Математические предпосылки

Алгоритм оперирует байтами, которые рассматриваются как элементы конечного поля $GF(2^8)$.

Элементами поля $GF(2^8)$ являются многочлены степени не более 7, которые могут быть заданы строкой своих коэффициентов. Если представить байт в виде

$$b_7, b_6, b_5, b_4, b_3, b_2, b_1, b_0,$$

то элемент поля описывается многочленом с коэффициентами из $\{0, 1\}$:

$$b_7 x^7 + b_6 x^6 + b_5 x^5 + b_4 x^4 + b_3 x^3 + b_2 x^2 + b_1 x^1 + b_0$$

Например, байту $\{11001011\}$ (или $\{cb\}$ в шестнадцатеричной форме) соответствует многочлен $x^7 + x^6 + x^3 + x + 1$.

Для элементов конечного поля определены аддитивные и мультипликативные операции.

Сложение

Сложение суть операция поразрядного XOR и поэтому обозначается как \oplus . Пример выполнения операции сложения:

$$(x^6 + x^4 + x^2 + x + 1) \oplus (x^7 + x + 1) = x^1 + x^6 + x^4 + x^2 \text{ (в виде многочленов)}$$

$$\{01010111\} \oplus \{10000011\} = \{11010100\} \text{ (двоичное представление)}$$

$$\{57\} \oplus \{83\} = \{d4\} \text{ (шестнадцатеричное представление)}$$

В конечном поле для любого ненулевого элемента **a** существует обратный элемент **-a**, при этом $a + (-a) = 0$, где нулевой элемент -это $\{00\}$. В $GF(2^8)$ справедливо $a + a = 0$, т. е. каждый ненулевой элемент является своей собственной аддитивной инверсией.

Умножение

Умножение, обозначаемое далее как \bullet , более сложная операция. Умножение в $GF(2^8)$ - это операция умножения многочленов со взятием результата по модулю неприводимого многочлена $m(x)$ восьмой степени и с использованием операции XOR при приведении подобных членов. В RIJNDAEL выбран $m(x) = x^8 + x^4 + x^3 + x + 1$, или в шестнадцатеричной форме $1\{1b\}$ (такая запись обозначает, что присутствует «лишний» девятый бит). Пример операции умножения:

$$\{57\} \cdot \{83\} = \{c1\},$$

или

$$\begin{aligned} &(x^6 + x^4 + x^2 + x + 1)(x^7 + x + 1) = \\ &= x^{13} + x^{11} + x^9 + x^8 + x^7 + x^7 + x^5 + x^3 + x^2 + x + x^6 + x^4 + x^2 + x + 1 = \\ &= x^{13} + x^{11} + x^9 + x^8 + x^6 + x^5 + x^4 + x^3 + 1 \end{aligned}$$

Следовательно,

$$x^{13} + x^{11} + x^9 + x^8 + x^6 + x^5 + x^4 + x^3 + 1 \bmod (x^8 + x^4 + x^3 + x + 1) = x^7 + x^6 + 1$$

Ясно, что результат является двоичным полиномом не выше 8 степени. В отличие от сложения, простой операции умножения на уровне байтов не существует.

Умножение, определенное выше, является ассоциативным, и существует единичный элемент ('01'). Для любого двоичного полинома $b(x)$ не выше 8-й степени можно использовать расширенный алгоритм Евклида для вычисления полиномов $a(x)$ и $c(x)$ таких, что

$$b(x) a(x) + m(x) c(x) = 1$$

Следовательно,

$$a(x) * b(x) \bmod m(x) = 1$$

или

$$b^{-1}(x) = a(x) \bmod m(x)$$

Более того, можно показать, что

$$a(x) * (b(x) + c(x)) = a(x) * b(x) + a(x) * c(x)$$

Из всего этого следует, что множество из 256 возможных значений байта образует конечное поле GF (2⁸) с XOR в качестве сложения и умножением, определенным выше.

Умножение на x

Если умножить b(x) на полином x, мы будем иметь:

$$b_7 x^8 + b_6 x^7 + b_5 x^6 + b_4 x^5 + b_3 x^4 + b_2 x^3 + b_1 x^2 + b_0 x$$

x * b(x) получается понижением предыдущего результата по модулю m(x). Если b₇ = 0, то данное понижение является тождественной операцией. Если b₇ = 1, m(x) следует вычесть (т.е. XORed). Из этого следует, что умножение на x может быть реализовано на уровне байта как левый сдвиг и последующий побитовый XOR с '1B'. Данная операция обозначается как b = xtime (a).

Полиномы с коэффициентами из GF (2⁸)

Полиномы могут быть определены с коэффициентами из GF (2⁸). В этом случае четырехбайтный вектор соответствует полиному степени 4.

Полиномы могут быть сложены простым сложением соответствующих коэффициентов. Как сложение в GF (2⁸) является побитовым XOR, так и сложение двух векторов является простым побитовым XOR.

Умножение представляет собой более сложное действие. Предположим, что мы имеем два полинома в GF (2⁸).

$$a(x) = a_3 x^3 + a_2 x^2 + a_1 x + a_0$$

$$b(x) = b_3 x^3 + b_2 x^2 + b_1 x + b_0$$

$$c(x) = a(x) b(x)$$

определяется следующим образом

$$c(x) = c_6 x^6 + c_5 x^5 + c_4 x^4 + c_3 x^3 + c_2 x^2 + c_1 x + c_0$$

$$c_0 = a_0 * b_0$$

$$\begin{aligned}
c_1 &= a_1 * b_0 \oplus a_0 * b_1 \\
c_2 &= a_2 * b_0 \oplus a_1 * b_1 \oplus a_0 * b_2 \\
c_3 &= a_3 * b_0 \oplus a_2 * b_1 \oplus a_1 * b_2 \oplus a_0 * b_3 \\
c_4 &= a_3 * b_1 \oplus a_2 * b_2 \oplus a_1 * b_3 \\
c_5 &= a_3 * b_2 \oplus a_2 * b_3 \\
c_6 &= a_3 * b_3
\end{aligned}$$

Ясно, что в таком виде $c(x)$ не может быть представлен четырехбайтным вектором. Понижая $c(x)$ по модулю полинома 4-й степени, результат может быть полиномом степени ниже 4. В Rijndael это сделано с помощью полинома

$$M(x) = x^4 + 1$$

так как

$$x^j \bmod (x^4 + 1) = x^{j \bmod 4}$$

Модуль, получаемый из $a(x)$ и $b(x)$, обозначаемый $d(x) = a(x) \otimes b(x)$, получается следующим образом:

$$\begin{aligned}
d_0 &= a_0 * b_0 \oplus a_3 * b_1 \oplus a_2 * b_2 \oplus a_1 * b_3 \\
d_1 &= a_1 * b_0 \oplus a_0 * b_1 \oplus a_3 * b_2 \oplus a_2 * b_3 \\
d_2 &= a_2 * b_0 \oplus a_1 * b_1 \oplus a_0 * b_2 \oplus a_3 * b_3 \\
d_3 &= a_3 * b_0 \oplus a_2 * b_1 \oplus a_1 * b_2 \oplus a_0 * b_3
\end{aligned}$$

Операция, состоящая из умножения фиксированного полинома $a(x)$, может быть записана как умножение матрицы, где матрица является циклической. Мы имеем

$$\begin{bmatrix} d_0 \\ d_1 \\ d_2 \\ d_3 \end{bmatrix} = \begin{bmatrix} a_0 & a_3 & a_2 & a_1 \\ a_1 & a_0 & a_3 & a_2 \\ a_2 & a_1 & a_0 & a_3 \\ a_3 & a_2 & a_1 & a_0 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

Описание криптоалгоритма

Формат блоков данных и число раундов

RIJNDAEL - это итерационный блочный шифр, имеющий архитектуру "Квадрат". Шифр имеет переменную длину блоков и различные длины ключей. Длина ключа и длина блока могут быть равны независимо друг от друга 128, 192 или 256 битам. В стандарте AES определена длина блока данных, равная 128 битам.

Введем следующие обозначения:

- N_b - число 32-битных слов содержащихся во входном блоке, $N_b=4$;
- N_k - число 32-битных слов содержащихся в ключе шифрования, $N_k=4,6,8$;
- N_r - число раундов шифрования, как функция от N_b и N_k , $N_r=10,12,14$.

Входные (input), промежуточные (state) и выходные (output) результаты преобразований, выполняемых в рамках криптоалгоритма, называются состояниями (State). Состояние можно представить в виде прямоугольного массива байтов (рис. 1). При размере блока, равном 128 битам, этот 16-байтовый массив (рис. 2, а) имеет 4 строки и 4 столбца (каждая строка и каждый столбец в этом случае могут рассматриваться как 32-разрядные слова). Входные данные для шифра обозначаются как байты состояния в порядке $S_{00}, S_{10}, S_{20}, S_{30}, S_{01}, S_{11}, S_{21}, S_{31}$

После завершения действия шифра выходные данные получаются из байтов состояния в том же порядке. В общем случае число столбцов N_b равно длине блока, деленной на 32.

State			
a_0	a_4	a_8	a_{12}
a_1	a_5	a_9	a_{13}
a_2	a_6	a_{10}	a_{14}
a_3	a_7	a_{11}	a_{15}

Рис. 7.8. Пример представления 128-разрядного блока данных в виде массива State, где a_i - байт блока данных, а каждый столбец - одно 32-разрядное слово

s_{00}	s_{01}	s_{02}	s_{03}
s_{10}	s_{11}	s_{12}	s_{13}
s_{20}	s_{21}	s_{22}	s_{23}
s_{30}	s_{31}	s_{32}	s_{33}

a

k_{00}	k_{01}	k_{02}	k_{03}
k_{10}	k_{11}	k_{12}	k_{13}
k_{20}	k_{21}	k_{22}	k_{23}
k_{30}	k_{31}	k_{32}	k_{33}

б

Рис. 7.9. Форматы данных: а - пример представления блока ($N_b = 4$); б - ключа шифрования ($N_k = 4$), где s и k - соответственно байты массива State и ключа, находящиеся на пересечении i -й строки и j -го столбца

Рисунок 7.10 демонстрирует такое представление, носящее название архитектуры «Квадрат».

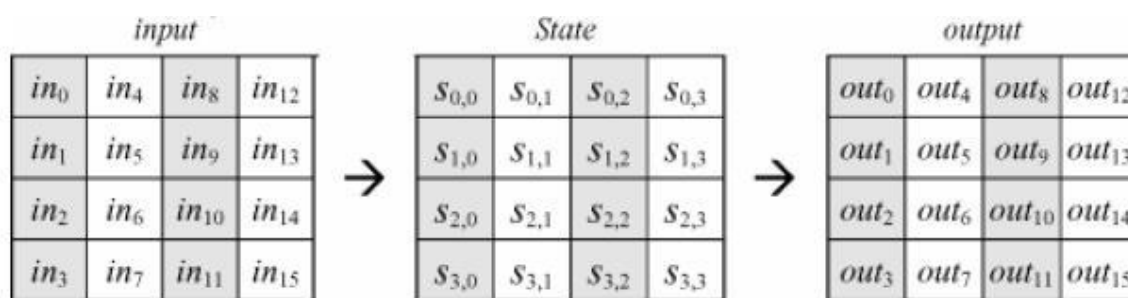


Рис. 7.10. Пример представления блока в виде матрицы $4N_b$

Ключ шифрования также представлен в виде прямоугольного массива с четырьмя строками (рис. 7.9, б). Число столбцов этого массива равно длине ключа, деленной на 32. В стандарте определены ключи всех трех размеров - 128 бит, 192 бита и 256 бит, то есть соответственно 4, 6 и 8 32-разрядных слова (или столбца – в табличной форме представления). В некоторых случаях ключ шифрования рассматривается как линейный массив 4-байтовых слов. Слова состоят из 4 байтов, которые находятся в одном столбце (при представлении в виде прямоугольного массива).

Число раундов N_r в алгоритме RIJNDAEL зависит от значений N_b и N_k , как показано в табл. 2.8. В стандарте AES определено соответствие между размером ключа, размером блока данных и числом раундов шифрования, как показано в табл.

Таблица 7.8. Число раундов N_r как функция от длины ключа N_k и длины блока N_b

N_r	$N_b = 4$	$N_b = 6$	$N_b = 8$
$N_k = 4$	10	12	14
$N_k = 6$	12	12	14
$N_k = 8$	14	14	14

Таблица 7.9. Соответствие между длиной ключа, размером блока данных и числом раундов в стандарте AES

Стандарт	Длина ключа (N_k слов)	Размер блока данных (N_b слов)	Число раундов (N_r)
AES-128	10	12	14
AES-192	12	12	14
AES-256	14	14	14

Раундовое преобразование

Раунд состоит из четырех различных преобразований:

- замены байтов `SubBytes()` - побайтовой подстановки в S-блоках с фиксированной таблицей замен размерностью 8×256 ;
- сдвига строк `ShiftRows()` - побайтового сдвига строк массива `State` на различное количество байт;
- перемешивания столбцов `MixColumns()` — умножения столбцов состояния, рассматриваемых как многочлены над $GF(2^8)$, на многочлен третьей степени $g(x)$ по модулю $x^4 + 1$;
- сложения с раундовым ключом `AddRoundKey()` - поразрядного XOR с текущим фрагментом развернутого ключа.

Замена байтов (*SubBytes*). Преобразование `SubBytes()` представляет собой нелинейную замену байтов, выполняемую независимо с каждым байтом состояния. Таблицы замены S-блока являются инвертируемыми и построены из композиции следующих двух преобразований входного байта:

1. получение обратного элемента относительно умножения в поле $GF(2^8)$, нулевой

элемент {00} переходит сам в себя;

2. применение преобразования над $GF(2^8)$, определенного следующим образом:

$$\begin{bmatrix} b'_0 \\ b'_1 \\ b'_2 \\ b'_3 \\ b'_4 \\ b'_5 \\ b'_6 \\ b'_7 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \\ b_6 \\ b_7 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}.$$

Другими словами суть преобразования может быть описана уравнениями:

$$b'_i = b_i \oplus b_{(i+4) \bmod 8} \oplus b_{(i+5) \bmod 8} \oplus b_{(i+6) \bmod 8} \oplus b_{(i+7) \bmod 8} \oplus c_i$$

где $c_0=c_1=c_5=c_6=0$, $c_2=c_3=c_4=c_7=1$, b'_i и b_i - соответственно преобразованное и исходное значение i -го бита, $i = 0, 1, 2, \dots, 7$.

Применение описанного S-блока ко всем байтам состояния обозначается как SubBytes(State). Рис. 4. иллюстрирует применение преобразования SubBytes() к состоянию. Логика работы S-блока при преобразовании байта {xy} отражена в табл. 3. Например, результат {ed} преобразования байта {53} находится на пересечении 5-й строки и 3-го столбца.

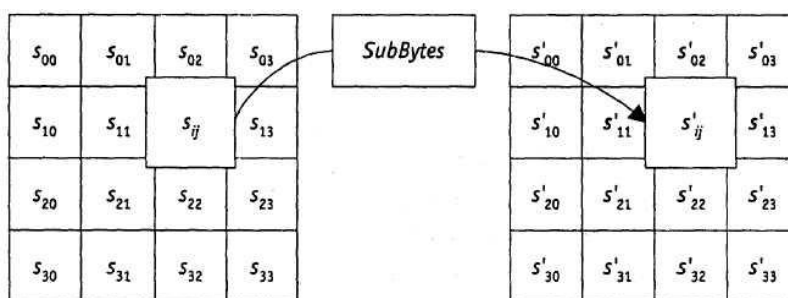


Рис. 7.10. SubBytes() действует на каждый байт состояния

Таблица 7.11. Таблица замен S-блока

x	y															
	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
0	63	7c	77	7b	f2	6b	6f	c5	30	01	67	2b	fe	d7	ab	76
1	ca	82	c9	7d	fa	59	47	f0	ad	d4	a2	af	9c	a4	72	c0
2	b7	fd	93	26	36	3f	f7	cc	34	a5	e5	f1	71	d8	31	15
3	04	c7	23	c3	18	96	05	9a	07	12	80	e2	eb	27	b2	75
4	09	83	2c	1a	1b	6e	5a	a0	52	3b	d6	b3	29	e3	2f	84
5	53	d1	00	ed	20	fc	b1	5b	6a	cb	be	39	4a	4c	58	cf
6	d0	ef	aa	fb	43	4d	33	85	45	f9	02	7f	50	3c	9f	a8
7	51	a3	40	8f	92	9d	38	f5	bc	b6	da	21	10	ff	f3	d2
8	cd	0c	13	ec	5f	97	44	17	c4	a7	7e	3d	64	5d	19	73
9	60	81	4f	dc	22	2a	90	88	46	ee	b8	14	de	5e	0b	db
a	e0	32	3a	0a	49	06	24	5c	c2	d3	ac	62	91	95	e4	79
b	e7	c8	37	6d	8d	d5	4e	a9	6c	56	f4	ea	65	7a	ae	08
c	ba	78	25	2e	1c	a6	b4	c6	e8	dd	74	1f	4b	bd	8b	8a
d	70	3e	b5	66	48	03	f6	0e	61	35	57	b9	86	c1	1d	9e
e	e1	f8	98	11	69	d9	8e	94	9b	1e	87	e9	ce	55	28	df
f	8c	a1	89	0d	bf	e6	42	68	41	99	2d	0f	b0	54	bb	16

Преобразование сдвига строк (ShiftRows). Последние 3 строки состояния циклически сдвигаются влево на различное число байтов. Строка 1 сдвигается на C1 байт, строка 2 - на C2 байт, и строка 3 - на C3 байт. Значения сдвигов C1, C2 и C3 в Rijndael зависят от длины блока Nb. Их величины приведены в табл. 2.12.

Таблица 7.12. Величина сдвига для разной длины блоков

Nb	C1	C2	C3
4	1	2	3
6	1	2	3
8	1	3	3

В стандарте AES, где определен единственный размер блока, равный 128 битам, C1 = 1, C2 = 2 и C3 = 3.

Операция сдвига последних трех строк состояния обозначена как ShiftRows(State). Рис. 2.11 показывает влияние преобразования на состояние.

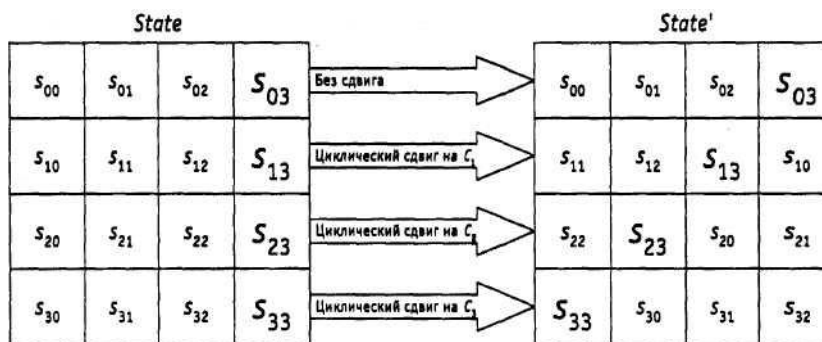


Рис. 7.11. ShiftRows() действует на строки состояния

Преобразование перемешивания столбцов (MixColumns). В этом преобразовании столбцы состояния рассматриваются как многочлены над $GF(2^8)$ и умножаются по модулю $x^4 + 1$ на многочлен $g(x)$, выглядящий следующим образом:

$$g(x) = \{03\}x^3 + \{01\}x^2 + \{01\}x + \{02\}.$$

Это может быть представлено в матричном виде следующим образом:

$$\begin{bmatrix} s'_{0c} \\ s'_{1c} \\ s'_{2c} \\ s'_{3c} \end{bmatrix} = \begin{bmatrix} 02 & 03 & 01 & 01 \\ 01 & 02 & 03 & 01 \\ 01 & 01 & 02 & 03 \\ 03 & 01 & 01 & 02 \end{bmatrix} \begin{bmatrix} s_{0c} \\ s_{1c} \\ s_{2c} \\ s_{3c} \end{bmatrix}, \quad 0 \leq c \leq 3,$$

где c - номер столбца массива State

В результате такого умножения байты столбца S_{0c} , S_{1c} , S_{2c} , S_{3c} заменяются соответственно на байты

$$S'_{0c} = (\{02\} \cdot S_{0c}) \oplus (\{03\} \cdot S_{1c}) \oplus S_{2c} \oplus S_{3c},$$

$$S'_{1c} = S_{0c} \oplus (\{02\} \cdot S_{1c}) \oplus (\{03\} \cdot S_{2c}) \oplus S_{3c},$$

$$S'_{2c} = S_{0c} \oplus S_{1c} \oplus (\{02\} \cdot S_{2c}) \oplus (\{03\} \cdot S_{3c}),$$

$$S'_{3c} = (\{03\} \cdot S_{0c}) \oplus S_{1c} \oplus S_{2c} \oplus (\{02\} \cdot S_{3c}).$$

Применение этой операции ко всем четырем столбцам состояния обозначено как MixColumns(State). Рис. 7.12 демонстрирует применение преобразования MixColumnsQ к столбцу состояния.

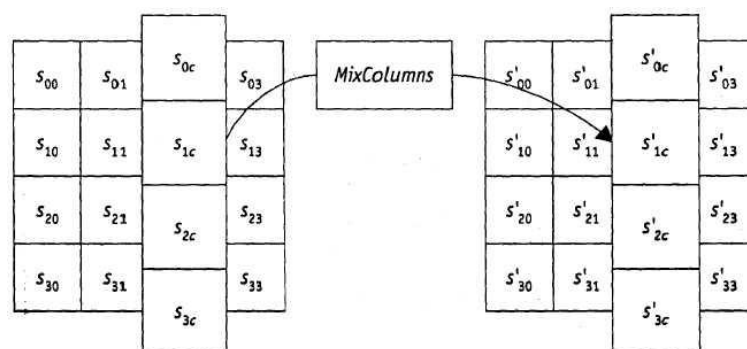


Рис. 7.12. MixColumnsQ действует на столбцы состояния

Добавление раундового ключа (AddRoundKey). В данной операции раундовый ключ добавляется к состоянию посредством простого поразрядного XOR. Раундовый ключ вырабатывается из ключа шифрования посредством алгоритма выработки ключей (key schedule). Длина раундового ключа (в 32-разрядных словах) равна длине блока Nb. Преобразование, содержащее добавление посредством XOR раундового ключа к состоянию (рис. 2.13), обозначено как $\text{AddRoundKey}(\text{State}, \text{RoundKey})$.

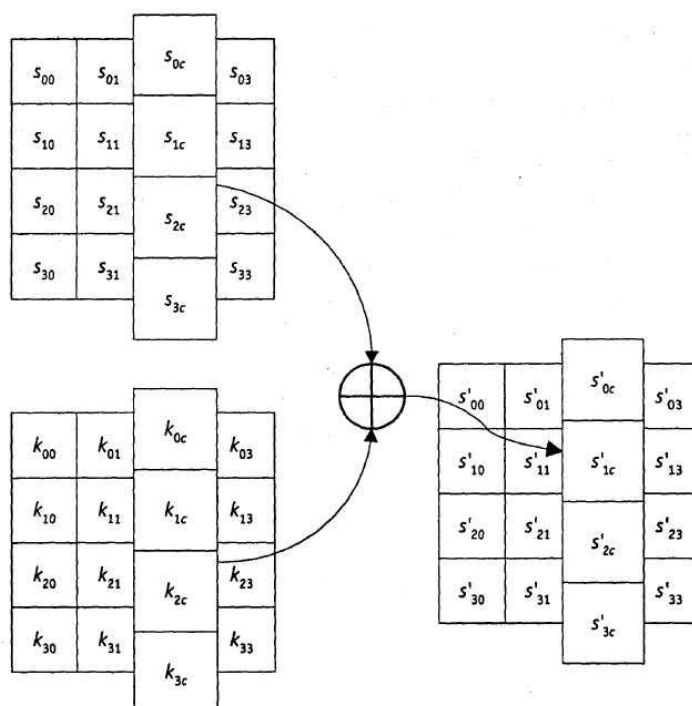


Рис. 7.13. При добавлении ключа раундовый ключ складывается посредством операции XOR с состоянием

Алгоритм выработки ключей (Key Schedule)

Раундовые ключи получаются из ключа шифрования посредством алгоритма выработки ключей. Он содержит два компонента: расширение ключа (Key Expansion) и выбор раундового ключа (Round Key Selection). основополагающие принципы алгоритма выглядят следующим образом:

- общее число битов раундовых ключей равно длине блока, умноженной на число раундов, плюс 1 (например, для длины блока 128 бит и 10 раундов требуется 1408 бит раундовых ключей);
- ключ шифрования расширяется в расширенный ключ (Expanded Key);
- раундовые ключи берутся из расширенного ключа следующим образом: первый раундовый ключ содержит первые Nb слов, второй - следующие Nb слов и т. д.

Расширение ключа (Key Expansion). Расширенный ключ в RIJNDAEL представляет собой линейный массив $w[i]$ из $Nb(N_r + 1)$ 4-байтовых слов, $i = 0, 1 \dots Nb(N_r + 1)$. В AES массив $w[i]$ состоит из $4(N_r + 1)$ 4-байтовых слов, $i = 0, 1 \dots 4(N_r + 1)$.

Примечание. Необходимо учитывать, что с целью полноты описания здесь приводится алгоритм для всех возможных длин ключей, на практике же полная его реализация нужна не всегда.

Первые N_k слов содержат ключ шифрования. Все остальные слова определяются рекурсивно из слов с меньшими индексами. Алгоритм выработки ключей зависит от величины N_k .

Как можно заметить (рис. 8, а), первые N_k слов заполняются ключом шифрования. Каждое последующее слово $w[i]$ получается посредством XOR предыдущего слова $w[i-1]$ и слова на N_k позиций ранее $w[i - N_k]$.

$$w[i] = w[i - 1] \oplus w[i - N_k].$$

Для слов, позиция которых кратна N_k , перед XOR применяется преобразование к $w[i-1]$, а затем еще прибавляется раундовая константа $Rcon$. Преобразование реализуется с помощью двух дополнительных функций: $RotWord()$, осуществляющей побайтовый сдвиг 32-разрядного слова по формуле $\{a_0 a_1 a_2 a_3\} \rightarrow \{a_1 a_2 a_3 a_0\}$, и $SubWord()$ осуществляющей

побайтовую замену с использованием 5-блока функции SubBytes(). Значение $Rcon[j]$ равно 2^j .¹ Значение $w[i]$ в этом случае равно

$$w[i] = \text{SubWord}(\text{RotWord}(w[i - 1])) \oplus Rcon[i/N_k] \oplus w[i - N_k].$$



Рис. 7.14. Процедуры:

а - расширения ключа (светло-серым цветом выделены слова расширенного ключа, которые формируются без использования функций SubWord() и RotWordQ; темно-серым цветом выделены слова расширенного ключа, при вычислении которых используются преобразования SubWordQ и RotWordQ);

б - выбора раундового ключа для $N_k - 4$

Выбор раундового ключа (Round Key Selection). Раундовый ключ i получается из слов массива раундового ключа от $W[N_b i]$ и до $W[N_b (i + 1)]$, как показано на рис. 8.

Примечание. Алгоритм выработки ключей можно осуществлять и без использования массива $w[i]$. Для реализаций, в которых существенно требование к занимаемой памяти, раундовые ключи могут вычисляться "на лету" посредством использования буфера из N_k слов. Расширенный ключ должен всегда получаться из ключа шифрования и никогда не указывается напрямую. Нет никаких ограничений на выбор ключа шифрования.

Функция зашифрования

Шифр RIJNDAEL состоит (рис. 2.15):

- из начального добавления раундового ключа;
- $N_r - 1$ раундов;
- заключительного раунда, в котором отсутствует операция MixColumns().

На вход алгоритма подаются блоки данных $State$, в ходе преобразований содержимое блока изменяется и на выходе образуется шифротекст, организованный опять же в виде блоков $State$, как показано на рис. 26, где $N_b = 4$, in_m и out_m - m -е байты соответственно входного и выходного блоков, $m = 0, 1 \dots 15$, s_{ij} - байт, находящийся на пересечении i -й строки и j -го столбца массива $State$, $i = j = 0, 1 \dots 3$.

Перед началом первого раунда происходит суммирование по модулю 2 с начальным ключом шифрования, затем - преобразование массива байтов $State$ в течение 10, 12 или 14 раундов в зависимости от длины ключа. Последний раунд несколько отличается от предыдущих тем, что не задействует функцию перемешивания байт в столбцах $MixColumns()$.

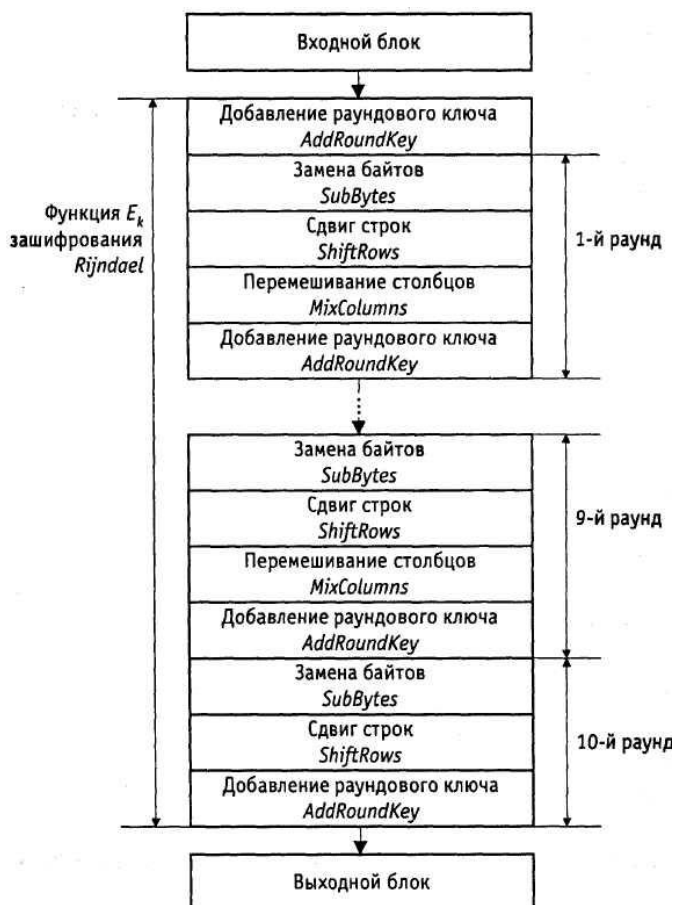


Рис. 7.15. Схема функции E_k зашифрования криптоалгоритма RIJNDAEL при $N_k = N_b = 4$



Рис. 7.16. Ход преобразования данных, организованных в виде блоков State

Рис. 7.17 демонстрирует и рассеивающие и перемешивающие свойства шифра. Видно, что два раунда обеспечивают полное рассеивание и перемешивание.

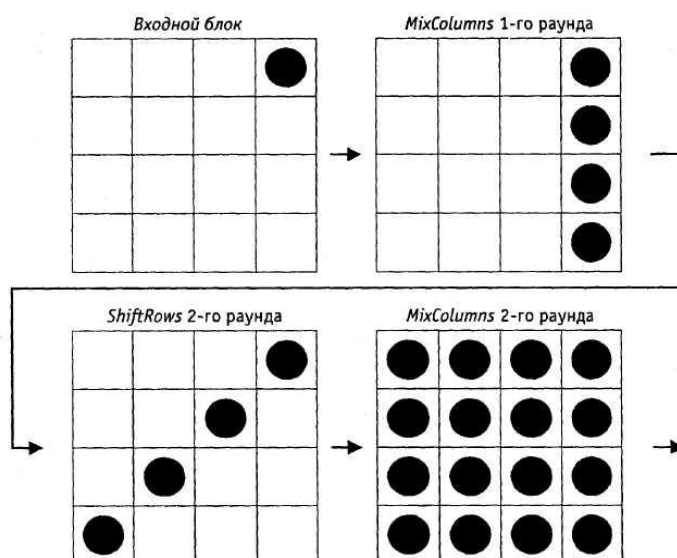


Рис. 7.17. Принцип действия криптоалгоритма RIJNDAEL; ● - измененный байт

Функция обратного расшифрования

Если вместо SubBytes(), ShiftRows(), MixColumns() и AddRoundKey() в обратной последовательности выполнить инверсные им преобразования, можно построить функцию обратного расшифрования. При этом порядок использования раундовых ключей является обратным по отношению к тому, который используется при зашифровании.

Далее приводится описание функций, обратных используемым при прямом зашифровании.

Функция AddRoundKey() обратна сама себе, учитывая свойства используемой в ней операции XOR.

Преобразование InvSubBytes. Логика работы инверсного S-блока при преобразовании байта {xy} отражена в табл. 7.13.

Таблица 7.13. Таблица замен инверсного S-блока

x	y															
	0	1	2	3	4	5	6	7	8	9	a	b	c	d	e	f
0	52	09	6a	d5	30	36	a5	38	bf	40	a3	9e	81	f3	d7	fb
1	7c	e3	39	82	9b	2f	ff	87	34	8e	43	44	c4	de	e9	cb
2	54	7b	94	32	a6	c2	23	3d	ee	4c	95	0b	42	fa	c3	4e
3	08	2e	a1	66	28	d9	24	b2	76	5b	a2	49	6d	8b	d1	25
4	72	f8	f6	64	86	68	98	16	d4	a4	5c	cc	5d	65	b6	92
5	6c	70	48	50	fd	ed	b9	da	5e	15	46	57	a7	8d	9d	84
6	90	d8	ab	00	8c	bc	d3	0a	f7	e4	58	05	b8	b3	45	06
7	d0	2c	1e	8f	ca	3f	0f	02	c1	af	bd	03	01	13	8a	6b
8	3a	91	11	41	4f	67	dc	ea	97	f2	cf	ce	f0	b4	e6	73
9	96	ac	74	22	e7	ad	35	85	e2	f9	37	e8	1c	75	df	6e
a	47	f1	1a	71	1d	29	c5	89	6f	b7	62	0e	aa	18	be	1b
b	fc	56	3e	4b	c6	d2	79	20	9a	db	c0	fe	78	cd	5a	f4
c	1f	dd	a8	33	88	07	c7	31	b1	12	10	59	27	80	ec	5f
d	60	51	7f	a9	19	b5	4a	0d	2d	e5	7a	9f	93	c9	9c	ef
e	a0	e0	3b	4d	ae	2a	f5	b0	c8	eb	bb	3c	83	53	99	61
f	17	2b	04	7e	ba	77	d6	26	e1	69	14	63	55	21	0c	7d

Преобразование InvShiftRows. Последние 3 строки состояния циклически сдвигаются вправо на различное число байтов. Строка 1 сдвигается на C1 байт, строка 2 - на C2 байт, и

строка 3 - на C3 байт. Значения сдвигов C1, C2, C3 зависят от длины блока Nb.
Преобразование *InvMixColumns*. В этом преобразовании столбцы состояния рассматриваются как многочлены над $GF(2^8)$ и умножаются по модулю $x^4 + 1$ на многочлен $g^{-1}(x)$, выглядящий следующим образом:

$$g^{-1}(x) = \{0b\}x^3 + \{0d\}x^2 + \{09\}x + \{0e\}.$$

Это может быть представлено в матричном виде следующим образом:

$$\begin{bmatrix} s'_{0c} \\ s'_{1c} \\ s'_{2c} \\ s'_{3c} \end{bmatrix} = \begin{bmatrix} 0e & 0b & 0d & 09 \\ 09 & 0e & 0b & 0d \\ 0d & 09 & 0e & 0b \\ 0b & 0d & 09 & 0e \end{bmatrix} \begin{bmatrix} s_{0c} \\ s_{1c} \\ s_{2c} \\ s_{3c} \end{bmatrix}, \quad 0 \leq c \leq 3.$$

В результате на выходе получаются следующие байты

$$\begin{aligned} s'_{0c} &= (\{0e\} \bullet s_{0c}) \oplus (\{0b\} \bullet s_{1c}) \oplus (\{0d\} \bullet s_{2c}) \oplus (\{09\} \bullet s_{3c}), \\ s'_{1c} &= (\{09\} \bullet s_{0c}) \oplus (\{0e\} \bullet s_{1c}) \oplus (\{0b\} \bullet s_{2c}) \oplus (\{0d\} \bullet s_{3c}), \\ s'_{2c} &= (\{0d\} \bullet s_{0c}) \oplus (\{09\} \bullet s_{1c}) \oplus (\{0e\} \bullet s_{2c}) \oplus (\{0b\} \bullet s_{3c}), \\ s'_{3c} &= (\{0b\} \bullet s_{0c}) \oplus (\{0d\} \bullet s_{1c}) \oplus (\{09\} \bullet s_{2c}) \oplus (\{0e\} \bullet s_{3c}). \end{aligned}$$

Функция прямого расшифрования

Алгоритм обратного расшифрования, описанный выше, имеет порядок приложения операций-функций, обратный порядку операций в алгоритме прямого шифрования, но использует те же параметры (развернутый ключ). Однако некоторые свойства алгоритма шифрования RIJNDAEL позволяют применить для расшифрования тот же порядок приложения функций (обратных используемым для зашифрования) за счет изменения некоторых параметров, а именно - развернутого ключа.

Два следующих свойства позволяют применить алгоритм прямого расшифрования.

Порядок приложения функций SubBytes() и ShiftRows() не играет роли. То же самое верно и для операций InvSubBytes() и InvShiftRows(). Это происходит потому, что функции SubBytes() и InvSubBytes() работают с байтами, а операции ShiftRows() и InvShiftRows() сдвигают целые байты, не затрагивая их значений.

Операция MixColumns() является линейной относительно входных данных, что означает

$$\text{InvMixColumns}(\text{State XOR RoundKey}) =$$

$$= \text{InvMixColumns}(\text{State}) \text{ XOR } \text{InvMixColumns}(\text{RoundKey})$$

Эти свойства функций алгоритма шифрования позволяют изменить порядок применения функций `InvSubBytes()` и `InvShiftRows()`. Функции `AddRoundKey()` и `InvMixColumns()` также могут быть применены в обратном порядке, но при условии, что столбцы (32-битные слова) развернутого ключа расшифрования предварительно пропущены через функцию `invMixColumns()`.

Таким образом, можно реализовать более эффективный способ расшифрования с тем же порядком приложения функций как и в алгоритме зашифрования.

При формировании развернутого ключа шифрования в процедуру развертывания ключа необходимо добавить следующий код

Примечание. В последнем операторе (в функции `InvMixColumn()`) происходит преобразование типа, так как развернутый ключ хранится в виде линейного массива 32-разрядных слов, в то время как входной параметр функции - двумерный массив байтов.

В табл. 6 приведена процедура зашифрования, а также два эквивалентных варианта процедуры расшифрования при использовании двухраундового варианта Rijndael. Первый вариант функции расшифрования суть обычная инверсия функции зашифрования. Второй вариант функции зашифрования получен из первого после изменения порядка следования операций в трех парах преобразований:

`invShiftRows` — `InvSubBytes` (дважды)

и `AddRoundKey` — `InvMixColumns`.

Таб. 7.14. Последовательность преобразований в двухраундовом варианте RIJNDAEL

Функция зашифрования двухраундового варианта RIJNDAEL	Функция обратного расшифрования двухраундового варианта RIJNDAEL	Эквивалентная функция прямого расшифрования двухраундового варианта RIJNDAEL
AddRoundKey	AddRoundKey	AddRoundKey
SubBytes	InvShiftRows	InvSubBytes
ShiftRows	InvSubBytes	InvShiftRows
MixColumns	AddRoundKey	InvMixColumns
AddRoundKey	InvMixColumns	AddRoundKey
SubBytes	InvShiftRows	InvSubBytes
ShiftRows	InvSubBytes	InvShiftRows
AddRoundKey	AddRoundKey	AddRoundKey

Очевидно, что результат преобразования при переходе от исходной к обратной последовательности выполнения операций в указанных парах не изменится.

Видно, что процедура зашифрования и второй вариант процедуры расшифрования совпадают с точностью до порядка использования раундовых ключей (при выполнении операций AddRoundKey), таблиц замен (при выполнении операций SubBytes и InvSubBytes) и матриц преобразования (при выполнении операций MixColumns и invMixColumns). Данный результат легко обобщить и на любое другое число раундов.

Атака “Квадрат”

Атака "Квадрат" была специально разработана для одноименного шифра SQUARE (авторы J. Daemen, L. Knudsen, V. Rijmen). Атака использует при своем проведении байт-ориентированную структуру шифра. Учитывая, что RIJNDAEL унаследовал многие свойства шифра SQUARE, эта атака применима и к нему. Далее приведено описание атаки "Квадрат" применительно к RIJNDAEL.

Атака "Квадрат" основана на возможности свободного подбора атакующим некоторого набора открытых текстов для последующего их зашифрования. Она независима от таблиц замен блоков, многочлена функции MixColumns() и способа разворачивания ключа. Эта атака для 6-раундового шифра RIJNDAEL, состоящего из 6 раундов, эффективнее, чем полный перебор по всему ключевому пространству. После описания базовой атаки на 4-раундовый RUNDAL, будет показано, как эту атаку можно продлить на 5 и даже 6 раундов. Но уже для 7 раундов “Квадрат” становится менее эффективным, чем полный перебор.

Предпосылки

Пусть L-набор - такой набор из 256 входных блоков (массивов State), каждый из которых имеет байты (назовем их активными), значения которых различны для всех 256 блоков. Остальные байты (будем называть их пассивными) остаются одинаковыми для всех 256 блоков из L-набора. То есть:

$$\forall x, y \in L \begin{cases} x_{ij} \neq y_{ij} , \text{ если байт с номером } ij \text{ активный;} \\ x_{ij} = y_{ij} , \text{ в противном случае.} \end{cases}$$

Будучи подвергнутыми обработке функциями SubBytes() и AddRoundKey() блоки L-набора дадут в результате другой L-набор с активными байтами в тех же позициях, что и у исходного. Функция ShiftRows() сместит эти байты соответственно заданным в ней смещениям в строках массивов State. После функции MixColumns() L-набор в общем случае необязательно останется L-набором (т. е. результат преобразования может перестать удовлетворять определению L-набора). Но поскольку каждый байт результата функции

MixColumns() является линейной комбинацией (с обратимыми коэффициентами) четырех входных байт того же столбца:

$$b_{ij} = 2a_{ij} \oplus 3a_{(i+1)j} \oplus a_{(i+2)j} \oplus a_{(i+3)j} \quad (7)$$

столбец с единственным активным байтом на входе даст в результате на выходе столбец со всеми четырьмя байтами - активными.

Базовая атака “Квадрат” на 4 раунда

Рассмотрим L-набор, во всех блоках которого активен только один байт. Иначе говоря, значение этого байта различно во всех 256 блоках, а остальные байты одинаковы (скажем, равны нулю). Проследим эволюцию этого байта на протяжении трех раундов. В первом раунде функция MixColumns() преобразует один активный байт в столбец из 4 активных байт. Во втором раунде эти 4 байта разойдутся по 4 различным столбцам в результате преобразования функцией ShiftRows(). Функция же MixColumns() следующего, третьего раунда преобразует эти байты в 4 столбца, содержащие активные байты. Этот набор все еще остается L-набором до того момента, когда он поступает на вход функции MixColumns() третьего раунда.

Основное свойство L-набора, используемое здесь, то, что поразрядная сумма по модулю 2 всех блоков такого набора всегда равна нулю. Действительно, поразрядная сумма неактивных (с одинаковыми значениями) байт равна нулю по определению операции поразрядного XOR, а активные байты, пробегаая все 256 значений, также при поразрядном суммировании дадут нуль. Рассмотрим теперь результат преобразования функцией MixColumns() в третьем раунде байтов входной массива данных a в байты выходного массива данных b . Покажем, что и в этом случае поразрядная сумма всех блоков выходного набора будет равна нулю, то есть:

$$\begin{aligned} \bigoplus_{b=\text{mixcolumns}(a), a \in L} b_{ij} &= \bigoplus_{a \in L} (2a_{ij} \oplus 3a_{(i+1)j} \oplus a_{(i+2)j} \oplus a_{(i+3)j}) = \\ &= 2 \bigoplus_{a \in L} a_{ij} \oplus 3 \bigoplus_{a \in L} a_{(i+1)j} \oplus \bigoplus_{a \in L} a_{(i+2)j} \oplus \bigoplus_{a \in L} a_{(i+3)j} = 2 \cdot 0 \oplus 3 \cdot 0 \oplus 1 \cdot 0 \oplus 1 \cdot 0 = 0 \end{aligned}$$

Таким образом, все данные на входе четвертого раунда сбалансированы (то есть их полная сумма равна нулю). Этот баланс в общем случае нарушается последующим преобразованием данных функцией SubBytes().

Мы предполагаем далее, что четвертый раунд является последним, то есть в нем нет функции MixColumns(). Тогда каждый байт выходных данных этого раунда зависит только от одного байта входных данных. Если обозначить через a байт входных данных четвертого раунда, через b байт выходных данных и через k - соответствующий байт раундового ключа, то можно записать:

$$a_{ij} = \text{SubBytes}(b_{ij}) \oplus k_{ij} \quad (8)$$

Отсюда, предполагая значение k_{ij} можно по известному a_{ij} вычислить b_{ij} , а затем проверить правильность догадки о значении k_{ij} ; если значения байта b_{ij} , полученные при данном k_{ij} не будут сбалансированы по всем блокам (то есть не дадут при поразрядном суммировании нулевой результат), значит догадка неверна. Перебрав максимум 2^8 вариантов байта раундового ключа, мы найдем его истинное значение.

По такому же принципу могут быть определены и другие байты раундового ключа. За счет того, что поиск может производиться отдельно (читай - параллельно) для каждого байта ключа, скорость подбора всего значения раундового ключа весьма велика. А по значению полного раундового ключа, при известном алгоритме его развертывания, не составляет труда восстановить сам начальный ключ шифрования.

Добавление пятого раунда в конец базовой атаки “Квадрат”

Если будет добавлен пятый раунд, то значения b_{ij} придется вычислять уже на основании выходных данных не четвертого, а пятого раунда. И дополнительно кроме байта раундового ключа четвертого раунда перебирать еще значения четырех байт столбца раундового ключа для пятого раунда. Только так мы сможем выйти на значения сбалансированных байт b_{ij} входных данных четвертого раунда.

Таким образом, теперь нам нужно перебрать 2^{40} значений - 2^{32} вариантов для 4 байт столбца раундового ключа пятого раунда и для каждого из них 2^8 вариантов для одного байта четвертого раунда. Эту процедуру нужно будет повторить для всех четырех столбцов пятого раунда. Поскольку при подборе "верного" значения байта раундового ключа четвертого раунда количество "неверных" ключей уменьшается в 2^8 раз, то работая одновременно с пятью L-наборами, можно с большой степенью вероятности правильно подобрать все 2^{40} бита. Поиск может производиться отдельно (т. е. параллельно) для каждого столбца каждого из пяти L-наборов, что опять же гораздо быстрее полного перебора всех возможных значений ключа.

Добавление шестого раунда в начало базовой атаки “Квадрат”

Основная идея заключается в том, чтобы подобрать такой набор блоков открытого текста, который на выходе после первого раунда давал бы L-набор с одним активным байтом. Это требует предположения о значении четырех байт ключа, используемых функцией `AddRoundKey()` перед первым раундом.

Для того чтобы на входе второго раунда был только один активный байт достаточно, чтобы в первом раунде один активный байт оставался на выходе функции `MixColumns()`. Это

означает, что на входе MixColumns() первого раунда должен быть такой столбец, байты а которого для набора из 256 блоков в результате линейного преобразования:

$$b_{ij} = 2a_i \oplus 3a_{i+1} \oplus a_{i+2} \oplus a_{i+3}, 0 \leq i \leq 3, \quad (9)$$

где i – номер строки, для одного определенного i давали 256 различных значений, в то время как для каждого из остальных трех значений i результат этого преобразования должен оставаться постоянным. Следуя обратно по порядку приложения функций преобразования в первом раунде, к ShiftRows() данное условие нужно применить к соответственно разнесенным по столбцам 4 байтам. С учетом применения функции SubBytes() и сложения с предполагаемым значением 4-байтового раундового ключа можно смело составлять уравнения и подбирать нужные значения байт открытого текста, подаваемых на зашифрование для последующего анализа результата:

$$b_{ij} = 2SubBytes(a_{ij} \oplus k_{ij}) \oplus 3SubBytes(a_{(i+1)(j+1)} \oplus k_{(i+1)(j+1)}) \oplus \\ \oplus SubBytes(a_{(i+2)(j+2)} \oplus k_{(i+2)(j+2)}) \oplus SubBytes(a_{(i+3)(j+3)} \oplus k_{(i+3)(j+3)}), 0 \leq i, j \leq 3$$

Таким образом, получаем следующий алгоритм взлома. Имеем всего 2^{32} различных значений a для определенных i и j . Остальные байты для всех блоков одинаковы (пассивные байты). Предположив значения четырех байт k ключа первого раунда, подбираем (исходя из вышеописанного условия) набор из 256 блоков. Эти 256 блоков станут L-набором после первого раунда. К этому L-набору применима базовая атака для 4 раундов. Подобранный с ее помощью один байт ключа последнего раунда фиксируется. Теперь подбирается новый набор из 256 блоков для того же значения 4 байт k ключа первого раунда. Опять осуществляется базовая атака, дающая один байт ключа последнего раунда. Если после нескольких попыток значение этого байта не меняется, значит мы на верном пути. В противном случае нужно менять предположение о значении 4 байт k ключа первого раунда. Такой алгоритм действий достаточно быстро приведет к полному восстановлению всех байт ключа последнего раунда.

Таким образом, атака "Квадрат" может быть применена к 6 раундам шифра RIJNDAEL, являясь при этом более эффективной, чем полный перебор по всему ключевому пространству. Любое известное продолжение атаки "Квадрат" на 7 и более раундов становится более трудоемким, чем даже обычный полный перебор значений ключа.

Поточные системы шифрования

Поточные шифры обычно делят на синхронные и самосинхронизирующиеся.

В синхронном поточном шифре шифрующая последовательность (гамма, ключевой поток, keystream) генерируется независимо от потока открытого текста и потока шифруемого

текста [6].

Генератор гаммы в синхронной поточной криптосистеме может быть описан следующими уравнениями

$$s_{i+1} = F(s_i),$$

$$k_i = f(s_i),$$

где s_i – значение внутреннего состояния (state), F – функция переходов, f – выходная функция, k_i – очередной элемент гаммы.

Начальное состояние (initial state) s_0 может быть определено из ключа k и вектора инициализации IV (initial vector).

Цель генератора гаммы – развернуть короткий случайный ключ k в длинную псевдослучайную последовательность (ПСП) k_1, k_2, \dots, k_n .

Работа синхронной криптосистемы схематически изображена на рисунке 2.17.

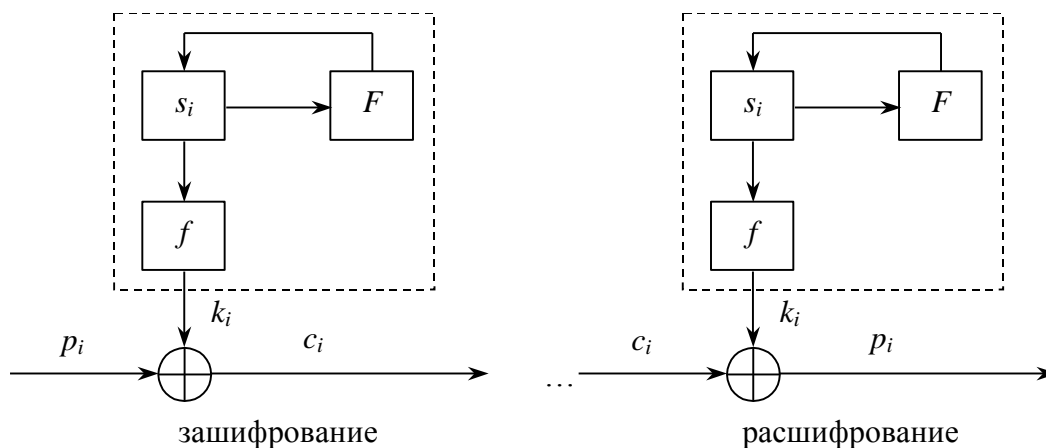


Рис. 7.17 Синхронный поточный шифр

В процессе зашифрования элементы гаммы и открытого текста подвергаются операции XOR (exclusive OR, исключающее ИЛИ), в результате чего создается шифртекст:

$$c_i = p_i \oplus k_i.$$

При расшифровании, для восстановления открытого текста, шифртекст и гамма также подвергаются операции XOR:

$$p_i = c_i \oplus k_i.$$

В самосинхронизирующемся поточном шифре гамма зависит не только от ключа, но еще и от определенного количества элементов ранее вычисленного шифртекста.

Генератор гаммы в самосинхронизирующейся поточной криптосистеме описывается следующим образом:

$$s_{i+1} = F(c_{i-1}, c_{i-2}, \dots, c_{i-N}),$$

$$k_i = f(k, s_i).$$

Работа самосинхронизирующейся криптосистемы схематически изображена на рисунке 7.18

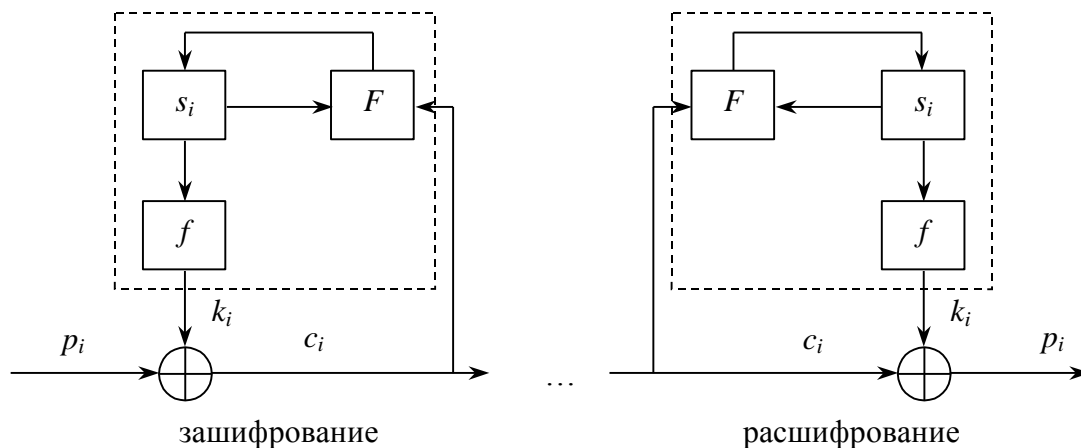


Рис. 7.18. Самосинхронизирующийся поточный шифр

Синхронные и самосинхронизирующиеся поточные шифры обладают как достоинствами так и недостатками. При использовании синхронного поточного шифра устройство зашифрования отправителя и устройство расшифрования получателя должны работать синхронно. Если в процессе передачи был потерян хотя бы один элемент шифртекста, то получатель обнаружит лишь бессмысленные данные, начиная с того места, где сбилась синхронизация. Обычно синхронизация достигается вставкой в передаваемое сообщение специальных “маркеров”. В результате этого элемент шифртекста, пропущенный в процессе передачи, приводит к неправильному расшифрованию лишь до тех пор, пока не будет принят один из маркеров. Другое решение – реинициализация состояний как устройства зашифрования, так и устройства расшифрования при некотором предварительно согласованном условии.

Самосинхронизирующийся поточный шифр обладает свойством, как свидетельствует его название, автоматически синхронизировать себя. Внутреннее состояние такого шифра зависит от N предыдущих элементов шифртекста. Если в процессе передачи был потерян или изменен элемент шифртекста, то после приема N правильных последовательных элементов шифртекста внутреннее состояние устройства расшифрования становится идентичным внутреннему состоянию устройства зашифрования, т.е. синхронизация восстановлена. Недостатком самосинхронизирующегося поточного шифра является распространение ошибки. Для каждого элемента шифртекста, искаженного при передаче, устройство расшифрования произведет N некорректных элементов гаммы. Следовательно, пока испорченный элемент влияет на внутреннее состояние, каждой ошибке шифртекста будет соответствовать N ошибок в открытом тексте. Синхронный поточный шифр имеет свойство не распространять ошибки. Расшифрование искаженного элемента шифртекста влияет

только на соответствующий элемент открытого текста. Хотя такое свойство может показаться желательным, у него есть и другая сторона. В этом случае ограничивается возможность обнаружения ошибки при расшифровании, но что еще более важно, противник имеет возможность произвести управляемые изменения в части шифртекста, совершенно точно зная, какие это вызовет изменения в соответствующем открытом тексте.

Поточные режимы блочных шифров

Для различных ситуаций, встречающихся на практике, разработано значительное количество режимов работы блочных шифров. Используя некоторый из них, блочные шифры можно реализовать как синхронные или самосинхронизирующиеся поточные шифры.

Для зарубежных стандартов блочного шифрования DES (Data Encryption Standard) и AES (Advanced Encryption Standard) существуют следующие основные режимы:

- Режим электронной кодовой книги, ECB (Electronic Code Book).
- Режим сцепления блоков шифртекста, CBC (Ciphertext Block Chaining).
- Режим обратной связи по шифртексту, CFB (Ciphertext Feedback).
- Режим обратной связи по выходу, OFB (Output Feedback).
- Режим счетчика, CTR (Counter mode).

В режимах OFB и CTR блочный шифр работает как синхронный поточный шифр, а в режиме CFB – как самосинхронизирующийся.

Отечественный стандарт блочного шифрования ГОСТ 28147-89 может работать в следующих режимах:

- Режим простой замены.
- Режим гаммирования.
- Режим гаммирования с обратной связью.
- Режим выработки имитовставки.

В режиме гаммирования алгоритм шифрования ГОСТ 28147-89 работает как синхронный поточный шифр, а в режиме гаммирования с обратной связью – как самосинхронизирующийся.

Строительные блоки поточных шифров

Рассмотрим основные блоки, используемые для построения поточных шифров.

Регистры сдвига с обратной связью

Большинство предложенных до настоящего времени алгоритмов поточного шифрования

так или иначе основаны на использовании регистров сдвига с обратной связью.

Регистр сдвига с обратной связью (feedback shift register, FSR) состоит из двух частей: регистра сдвига и функции обратной связи (рисунок 3). Регистр сдвига представляет собой последовательность битов. Длина регистра сдвига выражается числом битов. Если длина регистра равна n битам, регистр называют n -битовым регистром сдвига. При каждом извлечении бита все биты регистра сдвига сдвигаются вправо на 1 позицию. Новый старший бит рассчитывается как функция от всех остальных битов регистра. На выходе регистра сдвига оказывается 1 бит [4, 6].

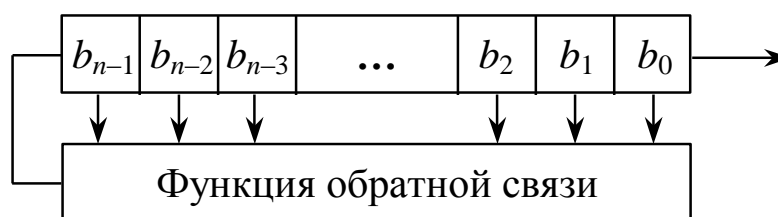


Рис. 7.19. Регистр сдвига с обратной связью

Регистры сдвига с линейной обратной связью

К простейшему типу FSR относится регистр сдвига с линейной обратной связью (linear feedback shift register, LFSR) [4]. Подавляющее большинство предложенных до настоящего времени генераторов поточного шифрования так или иначе основано на LFSR.

На это существует несколько причин:

- LFSR хорошо подходят для аппаратной реализации;
- LFSR могут производить последовательности большого периода;
- LFSR могут производить последовательности с хорошими статистическими свойствами;
- LFSR могут быть легко проанализированы с помощью алгебраических техник.
- LFSR длины n состоит из n элементов задержки (ячеек) $b_{n-1}, b_{n-2}, \dots, b_1, b_0$, каждый из которых может хранить один бит и имеет по одному входу и выходу.

Исходной информацией для построения LFSR является образующий многочлен. Степень этого многочлена определяет разрядность регистра сдвига, а ненулевые коэффициенты – характер обратных связей (номера отводов сигналов обратной связи). В общем случае двоичный образующий многочлен степени n имеет вид [6]:

$$c(x) = \sum_{i=0}^n c_i x^i = 1 + c_1 x + c_2 x^2 + \dots + c_{n-1} x^{n-1} + c_n x^n,$$

где $c_n = c_0 = 1$, $c_j \in \{0, 1\}$ для $j = 1, \dots, (n-1)$.

В общем случае двоичному образующему многочлену $c(x)$ соответствует две схемы: Фибоначчи (рисунок 4) и Галуа (рисунок 2.20).

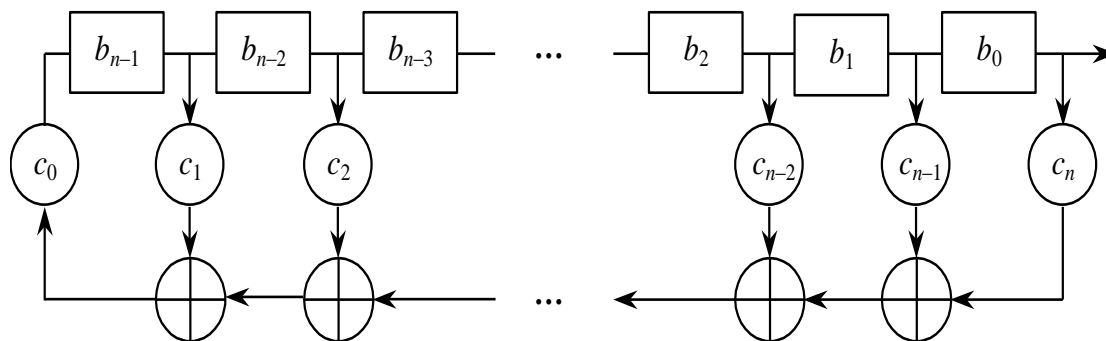


Рис. 2.20. LFSR, схема Фибоначчи

На каждом такте работы схемы Фибоначчи содержимое регистра сдвигается вправо на один бит, так что $b_i = b_{i+1}$ для $i = 0, \dots, (n - 2)$, а содержимое 0-ой ячейки b_0 поступает на выход LFSR. Новое содержимое $(n - 1)$ -ой ячейки b_{n-1} рассчитывается как сумма по модулю 2 предыдущих состояний определенных ячеек

$$b_{n-1} = \sum_{i=0}^{n-1} c_{n-i} b_i .$$

При $c_i = 1$ умножение на c_i равносильно наличию обратной связи, при $c_i = 0$ – отсутствию.

Схема Галуа представлена на рисунке 2.21.

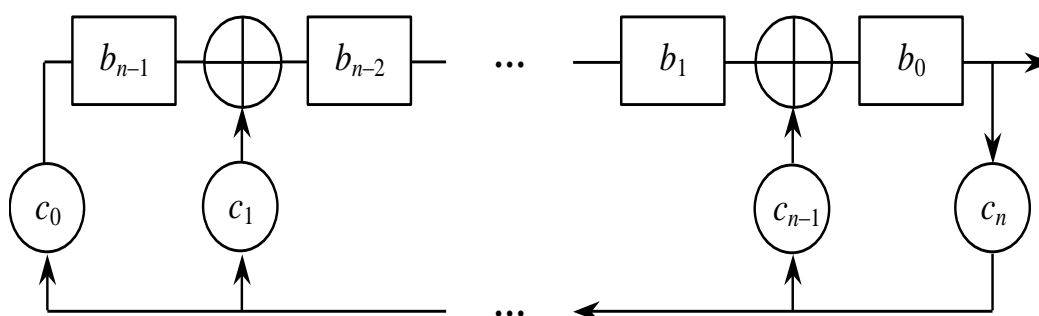


Рис. 2.21. LFSR, схема Галуа

На каждом такте работы схемы Галуа содержимое регистра сдвига сдвигается вправо на 1 бит так, что $b_i = b_{i+1} \oplus b_0 c_{n-i}$, а содержимое 0-ой ячейки b_0 поступает на выход LFSR и в $(n - 1)$ ячейку регистра, т.е. $b_{n-1} = b_0$.

n -битовый LFSR может находиться в одном из $(2^n - 1)$ внутренних состояний. Это значит, что теоретически такой регистр может генерировать псевдслучайную последовательность с периодом $(2^n - 1)$ битов. (Это число равно $(2^n - 1)$, а не 2^n , поскольку заполнение LFSR нулями влечет вывод регистром бесконечной последовательности нулей,

что совершенно бесполезно.) Только при определенных последовательностях отводов LFSR циклически пройдет через все $2^n - 1$ внутренних состояний. Такие регистры называются регистрами LFSR с максимальным периодом. Получившийся выход называют m -последовательностью.

Для обеспечения максимального периода конкретного LFSR, соответствующий многочлен, образованный из последовательности отводов регистра, должен быть примитивным по модулю 2. Степень многочлена является длиной регистра сдвига [4, 6].

Как бы ни был хорошо подобран полином обратной связи, LFSR остается линейным устройством. А такие устройства обычно легко поддаются криптоанализу независимо от того, насколько много параметров сохраняется в тайне. В современной криптографической литературе LFSR сами по себе не рекомендуются в качестве генераторов псевдослучайных шифрующих последовательностей [6].

Регистры LFSR сами по себе являются хорошими генераторами псевдослучайных последовательностей, но они обладают некоторыми нежелательными неслучайными свойствами. Последовательные биты линейны, что делает их бесполезными для шифрования. Внутреннее состояние LFSR длины n определяет следующие n выходных битов генератора. Даже если схема обратной связи хранится в секрете, она может быть определена по $2n$ выходным битам генератора с помощью высокоэффективного алгоритма Берлекампа-Мэсси [4]. В то же время подавляющее большинство реальных конструкций для поточного шифрования строится на основе LFSR [6].

Генераторы на основе LFSR

Поточные шифры на основе LFSR подразделяют на [4]:

- системы с генератором с равномерным движением регистров;
- системы с генератором с неравномерным движением регистров.

В генераторе с равномерным движением регистров каждый раз для получения нового бита следует однократно сдвинуть LFSR. Выходной бит представляет собой функцию некоторых битов LFSR. Генераторы с равномерным движением в свою очередь делятся на

[4, 5]:

- комбинирующий генератор;
- фильтрующий генератор.

Комбинирующий генератор (рисунок 6) состоит из нескольких параллельно работающих LFSR, выходы которых поступают на вход некоторой булевой функции f , комбинирующей эти выходы для генерации ключевого потока.

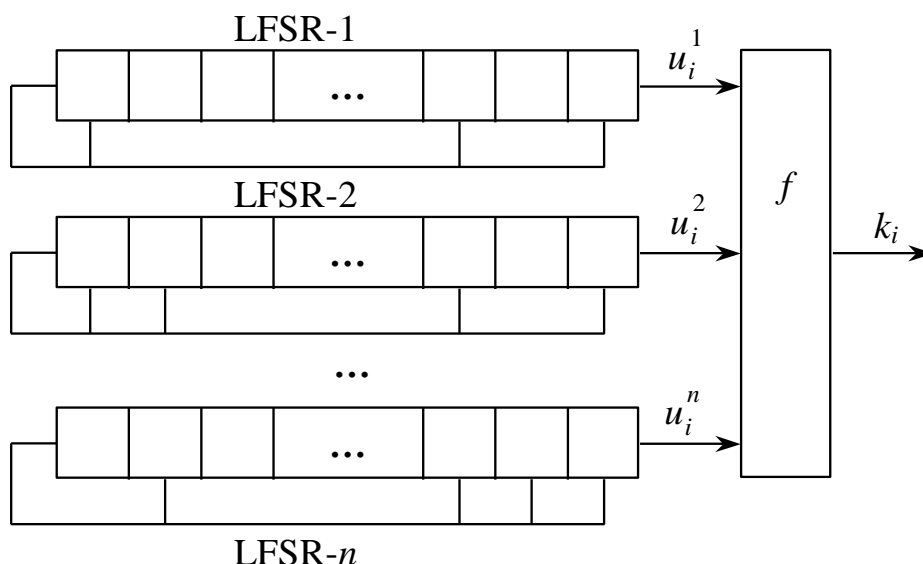


Рис. 2.22. Комбинирующий генератор

Ключом соответствующего поточного шифра является начальное заполнение каждого регистра, реже – начальное заполнение и функции обратных связей.

Результат работы комбинирующего генератора можно представить в виде

$$k_i = f(u_i^1, u_i^2, \dots, u_i^n),$$

где k_i – i -ый бит ключевого потока, производимого генератором;

n – количество LFSR;

u_i^j – i -ый бит, генерируемый j -ым LFSR.

Фильтрующий генератор (рисунок 2.23) состоит из одного LFSR. Для генерации ключевого потока используется нелинейная функция f , на вход которой подаются значения некоторых ячеек LFSR. Функция f в этом случае называется фильтрующей функцией.

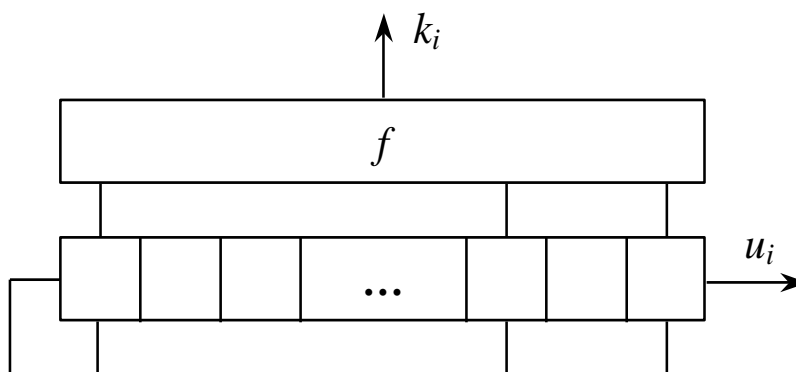


Рис. 7.23. Фильтрующий генератор

Ключом соответствующего поточного шифра является начальное заполнение регистра, реже – начальное заполнение и функция обратных связей.

Результат работы фильтрующего генератора можно представить в виде

$$k_i = f(u_i^1, u_i^2, \dots, u_i^n),$$

где k_i – i -ый бит ключевого потока, производимого генератором;

n – длина LFSR;

u_i^j – состояние j -ой ячейки LFSR.

Для того чтобы получить ключевую последовательность, обладающую хорошими статистическими свойствами, необходимо чтобы обратные связи в LFSR были сформированы в соответствии с примитивными многочленами, а комбинирующая и фильтрующая функции формировали равномерно распределенные последовательности.

Параметры LFSR и комбинирующей и фильтрующей функций обычно общеизвестны, секретными являются начальные состояния LFSR, зависящие от ключа. Поэтому целью большинства атак на комбинирующие и фильтрующие генераторы является восстановление начальных состояний всех LFSR.

Фильтрующий генератор можно рассматривать как частный случай комбинирующего генератора, у которого все LFSR одинаковы, а начальное состояние LFSR- i совпадает с состоянием LFSR-1 на i -ом такте работы. Но, поскольку криптоанализ этих схем несколько различается, принято рассматривать эти схемы как два различных типа генераторов ключевого потока.

Примерами комбинирующего генератора являются: генератор Геффе и генератор Дженнинга. Примером фильтрующего генератора является алгоритм Nanoteq.

Регистры сдвига с нелинейной обратной связью

Как уже говорилось выше, регистр сдвига с обратной связью состоит из двух частей: регистра сдвига и функции обратной связи. В качестве функции обратной связи выступает любая булева функция f от n переменных. В случае, когда функция обратной связи f линейна, регистр сдвига называется регистром сдвига с линейной обратной связью (LFSR), в противном случае – регистром сдвига с нелинейной обратной связью (non-linear feedback shift register, NLFSR).

Регистры сдвига с обратной связью по переносу

Регистр сдвига с обратной связью по переносу (feedback with carry shift register, FCSR) напоминает LFSR [4]. В обоих используется регистр сдвига и функция обратной связи, но в FCSR дополнительно предусмотрен еще и регистр переноса (рисунок 8).

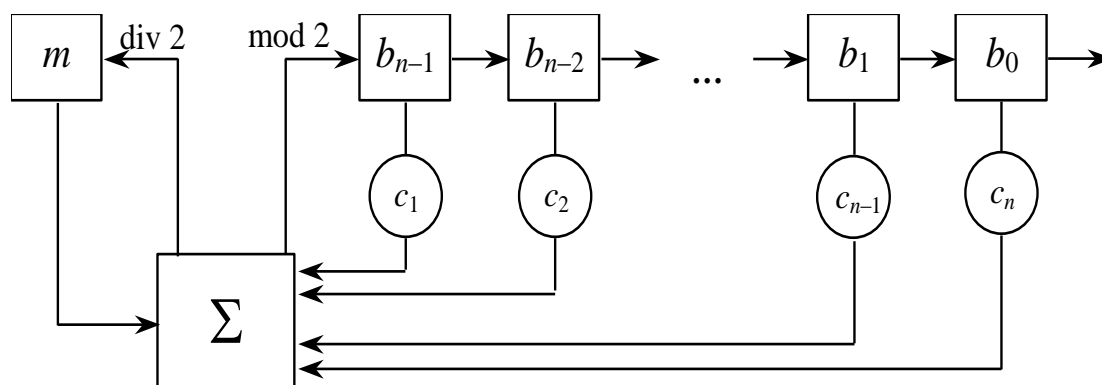


Рисунок 7.24. Регистр сдвига FCSR

На рисунке 2.24 знак Σ означает целочисленное сложение. Содержимое регистра сдвига состоит из n бит, обозначенных $b_{n-1}, b_{n-2}, \dots, b_1, b_0$.

Работа FCSR описывается следующим образом [6]:

1. Вычисляется сумма $\sigma_n = \sum_{k=1}^n c_k b_{n-k} + m$.
2. Содержимое регистра сдвигается на одну позицию вправо. Содержимое крайней правой ячейки FCSR b_0 поступает на выход.
3. Рассчитывается новое содержимое крайней левой ячейки $b_{n-1} = \sigma_n \pmod{2}$.
4. В регистр переноса записывается новое значение $m = \frac{\sigma_n - a_n}{2} = \left\lfloor \frac{\sigma_n}{2} \right\rfloor$.

Регистром переноса служит число, а не бит. Размер регистра переноса должен быть не менее $\log_2 t$, где t – количество отводов. Например, если отвода два, то регистр переноса однобитовый, а если отводов четыре, то регистр переноса должен состоять из 2 бит.

Максимальный период последовательности, генерируемой FCSR, равен $(c - 1)$, где c – число обратной связи (connection integer) FCSR. Число c определяется отводами обратной связи:

$$c = c_n 2^n + c_{n-1} 2^{n-1} + \dots + c_1 2^1 - 1.$$

Последовательность максимального периода, генерируемая FCSR, называется l -последовательностью. l -последовательность генерируется FCSR с числом обратной связи c , для которого 2 является корнем примитивным.

Поточные шифры

Обозначения

Ниже используются следующие обозначения:

- $+$: $(x + y)$ означает $x + y \pmod{2^{32}}$, где $0 \leq x < 2^{32}$ и $0 \leq y < 2^{32}$.
- \oplus : поразрядное исключающее ИЛИ (XOR).
- \otimes : поразрядное И (AND).
- $\&$: логическое И.
- $\|$: конкатенация.
- \gg : оператор сдвига вправо. $x \gg n$ означает, что x сдвигается вправо на n бит.
- \ll : оператор сдвига влево. $x \ll n$ означает, что x сдвигается влево на n бит.
- \ggg : оператор циклического сдвига вправо. $x \ggg n$ означает $((x \gg n) \oplus (x \ll (32 - n)))$, где $0 \leq n < 32$, $0 \leq x < 2^{32}$.
- \lll : оператор циклического сдвига влево. $x \lll n$ означает $((x \ll n) \oplus (x \gg (32 - n)))$, где $0 \leq n < 32$, $0 \leq x < 2^{32}$.
- \boxminus : $(x \boxminus y)$ означает $x - y \pmod{512}$.
- $A^{[g..h]}$ обозначает биты с g по h переменной A .

Поточный шифр HC-128

Поточный шифр HC-128 [6] – упрощенная версия поточного шифра HC-256 для 128-битового уровня безопасности. HC-128 – простой и свободно доступный шифр, ориентированный на программную реализацию. Поточный шифр HC-128 использует 128 битовый ключ и вектор инициализации IV.

HC-128 состоит из двух секретных таблиц, каждая из которых содержит 512 32-разрядных элемента. На каждом шаге один элемент таблицы обновляется с помощью нелинейной функции обратной связи. Все элементы этих двух таблиц обновляются каждые 1024 шага. На каждом шаге, нелинейной функцией фильтрации выхода генерируется один 32-разрядный выходной блок.

Инициализация

Процесс инициализации HC-128 состоит в расширении ключа и вектора инициализации в таблицы P и Q (подобно SHA-256) и выполнении цикла шифрования (1024 раза), без генерации выходной последовательности (выходные последовательности используются для обновления P и Q).

В процессе инициализации выполняются следующие шаги:

1. Пусть $K = K_0||K_1||K_2||K_3$ и $IV = IV_0||IV_1||IV_2||IV_3$, где K_i and IV_i обозначают 32-битовые числа. Ключ и IV расширяются в массив W_i ($0 \leq i \leq 1279$):

$$W_i = \begin{cases} K_i & 0 \leq i \leq 7 \\ IV_{i-8} & 8 \leq i \leq 15 \\ f_2(W_{i-2}) + W_{i-7} + f_1(W_{i-15}) + W_{i-16} + i & 16 \leq i \leq 1279 \end{cases}$$

где функции $f_1(x)$ и $f_2(x)$ при $x = x_3||x_2||x_1||x_0$, (x – 32-битное слово, x_0, x_1, x_2 , и x_3 – четыре байта. x_3 и x_0 обозначают соответственно самый старший байт и самый младший байт величины x) определяются как

$$f_1(x) = (x \gg \gg 7) \oplus (x \gg \gg 18) \oplus (x \gg 3),$$

$$f_2(x) = (x \gg \gg 17) \oplus (x \gg \gg 19) \oplus (x \gg 10).$$

2. Обновить таблицы P и Q массивом W .

$$P[i] = W_{i+256}, \text{ для } 0 \leq i \leq 511,$$

$$Q[i] = W_{i+768}, \text{ для } 0 \leq i \leq 511.$$

3. Выполнить цикл шифрования 1024 раза и использовать выходные последовательности для замены элементов таблицы следующим образом:

for $i = 0$ to 511 do

$$P[i] = (P[i] + g_1(P[i \boxminus 3], P[i \boxminus 10], P[i \boxminus 511])) \oplus h_1(P[i \boxminus 12]));$$

for $i = 0$ to 511 do

$$Q[i] = (Q[i] + g_2(Q[i \boxminus 3], Q[i \boxminus 10], Q[i \boxminus 511])) \oplus h_2(Q[i \boxminus 12]));$$

где функции $g_1(x)$, $g_2(x)$, $h_1(x)$ и $h_2(x)$ при $x = x_3||x_2||x_1||x_0$ определяются как

$$g_1(x, y, z) = ((x \gg \gg 10) \oplus (z \gg \gg 23)) + (y \gg \gg 8),$$

$$g_2(x, y, z) = ((x \ll \ll 10) \oplus (z \ll \ll 23)) + (y \ll \ll 8),$$

$$h_1(x) = Q[x_0] + Q[256 + x_2],$$

$$h_2(x) = P[x_0] + P[256 + x_2].$$

Для функции $h_1(x)$ таблица Q используется как S -блок. Для функции $h_2(x)$, таблица P используется как S -блок.

Процесс инициализации завершен, и шифр готов к генерации ключевой последовательности.

Генерация ключевого потока

На каждом шаге, обновляется один элемент таблицы, и генерируется один 32-битовый блок выходной последовательности. S -блок используется для генерации только 512 выходных последовательностей, затем он обновляется за следующие 512 шага. Процесс

генерации ключевой последовательности HC-128 описывается следующим псевдокодом (рисунок 1).

```

i = 0;
repeat {
  j = i mod 512;
  if (i mod 1024) < 512 {
    P[j] = P[j] + g1(P[j-3], P[j-10], P[j-511]);
    si = h1(P[j-12]) ⊕ P[j];
  }
  else {
    Q[j] = Q[j] + g2(Q[j-3], Q[j-10], Q[j-511]);
    si = h2(Q[j-12]) ⊕ Q[j];
  }
  i = i + 1;
}

```

Рисунок 1. Псевдокод генерации ключевой последовательности

s_i – 32-битовый выходной блок на i -м шаге процесса генерации ключевой последовательности.

Поточный шифр Rabbit

Алгоритм Rabbit в качестве входных данных использует 128-битовый секретный ключ и, если необходимо, 64-битовый вектор IV [6]. За одну итерацию генерирует блок 128 псевдослучайных бит. Шифрование/расшифрование производится путем сложения по модулю 2 (XOR) сгенерированной псевдослучайной последовательности с открытым/зашифрованным текстом. Размер внутреннего состояния – 513 бит, разделенных между восьмью 32-разрядными переменными состояниями, восьмью 32-разрядными счетчиками и одним битом переноса. Эти восемь переменных состояний обновляются посредством восьми парных нелинейных функций. Счетчики гарантируют нижнюю границу на длине периода для переменных состояний.

Внутреннее состояние поточного шифра состоит из 513 битов. 512 битов разделены между восьмью 32-разрядными переменными состояниями $x_{j,i}$ и восьмью 32-разрядными переменными счетчиками $c_{j,i}$, где $x_{j,i}$ – переменная состояния j -ой подсистемы в i -ой итерации, а $c_{j,i}$ обозначает соответствующую переменную счетчика. Также есть один бит счетчика по переносу $\phi_{i,i}$, который должен сохраняться/накапливаться между итерациями. Этот бит счетчика по переносу инициализируется путем обнуления. Восемь переменных состояний и восемь счетчиков формируются из ключа при инициализации.

Инициализация

Алгоритм инициализируется разворачиванием 128-битового ключа в восемь переменных состояний и восемь счетчиков так, что есть взаимно однозначное соответствие между ключом и переменными начального состояния $x_{j,i}$ и начальными значениями счетчиков $c_{j,i}$.

Ключ $K^{[127..0]}$ разделен на восемь 16-битовых подключей:

$$k_0 = K^{[15..0]}, k_1 = K^{[31..16]}, \dots, k_7 = K^{[127..112]}.$$

Переменные состояния и счетчика инициализируются из подключей следующим образом:

$$x_{j,0} = \begin{cases} k_{(j+1 \bmod 8)} \parallel k_j, & \text{для четных } j \\ k_{(j+5 \bmod 8)} \parallel k_{(j+4 \bmod 8)}, & \text{для нечетных } j \end{cases}$$

и

$$x_{j,0} = \begin{cases} k_{(j+4 \bmod 8)} \parallel k_{(j+5 \bmod 8)}, & \text{для четных } j \\ k_j \parallel k_{(j+1 \bmod 8)}, & \text{для нечетных } j \end{cases}.$$

Чтобы уменьшить корреляцию между битами ключа и битами переменных внутреннего состояния система повторяется четыре раза, в соответствии с функцией следующего состояния, определенной ранее. Наконец, переменные счетчика повторно инициализируются согласно выражению:

$$c_{j,4} = c_{j,4} \oplus x_{(j+4 \bmod 8),4}$$

для всех j , чтобы предотвратить восстановление ключа обратным преобразованием системы счетчика.

Обозначим внутреннее состояние после применения схемы установки ключа как основное состояние. Пусть копия этого основного состояния будет изменена в соответствии со схемой установки IV . Схема установки IV изменяет состояние счетчика как функция IV . Это реализуется путем применения операции XOR к 64-битовым IV и всеми 256 битами состояния счетчика. 64 бита IV обозначаются $IV^{[63..0]}$. Счетчики изменяются следующим образом:

$$c_{0,4} = c_{0,4} \oplus IV^{[31..0]},$$

$$c_{1,4} = c_{1,4} \oplus (IV^{[63..48]} \parallel IV^{[31..16]}),$$

$$c_{2,4} = c_{2,4} \oplus IV^{[63..32]},$$

$$c_{3,4} = c_{3,4} \oplus (IV^{[47..32]} \parallel IV^{[15..0]}),$$

$$c_{4,4} = c_{4,4} \oplus IV^{[31..0]},$$

$$c_{5,4} = c_{5,4} \oplus (IV^{[63..48]} \parallel IV^{[31..16]}),$$

$$c_{6,4} = c_{6,4} \oplus IV^{[63..32]},$$

$$c_{7,4} = c_{7,4} \oplus (IV^{[47..32]} \parallel IV^{[15..0]}).$$

Чтобы сделать все биты состояния нелинейно зависящими от всех битов IV система повторяется четыре раза. Модификация счетчика с помощью IV гарантирует, что все 2^{64} различные вектора IV приведут к уникальным ключевым последовательностям.

Генерация ключевого потока

Ядром алгоритма Rabbit является повторение функции выработки следующего состояния, определенной уравнениями:

$$x_{0,i+1} = g_{0,i} + (g_{7,i} \lll 16) + (g_{6,i} \lll 16),$$

$$x_{1,i+1} = g_{1,i} + (g_{0,i} \lll 8) + g_{7,i},$$

$$x_{2,i+1} = g_{2,i} + (g_{1,i} \lll 16) + (g_{0,i} \lll 16),$$

$$x_{3,i+1} = g_{3,i} + (g_{2,i} \lll 8) + g_{1,i},$$

$$x_{4,i+1} = g_{4,i} + (g_{3,i} \lll 16) + (g_{2,i} \lll 16),$$

$$x_{5,i+1} = g_{5,i} + (g_{4,i} \lll 8) + g_{3,i},$$

$$x_{6,i+1} = g_{6,i} + (g_{5,i} \lll 16) + (g_{4,i} \lll 16),$$

$$x_{7,i+1} = g_{7,i} + (g_{6,i} \lll 8) + g_{5,i},$$

$$g_{j,i} = ((x_{j,i} + c_{j,i+1})^2 \oplus ((x_{j,i} + c_{j,i+1})^2 \ggg 32)) \bmod 2^{32},$$

где все операции сложения приводятся по модулю 2^{32} . Эта двойная система приведена на рисунке 2.25. Перед каждой итерацией счетчики увеличиваются в соответствии с описанным ниже правилом.

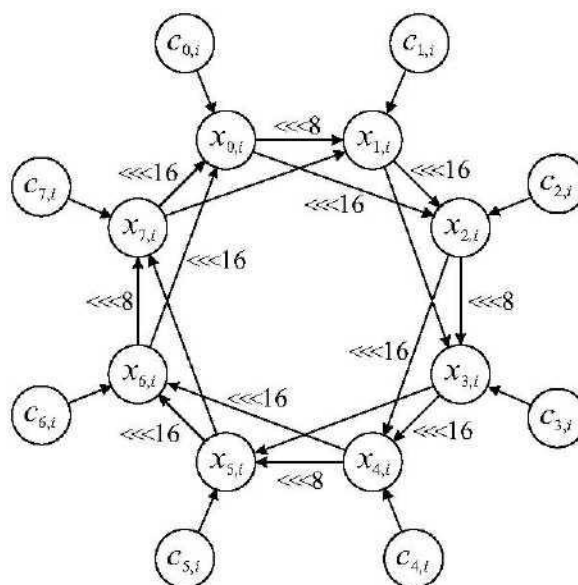


Рис. 2.25. Графическое представление системы

Работа счетчиков определяется следующим образом:

$$c_{0,i+1} = c_{0,i} + a_0 + \phi_{7,i} \bmod 2^{32},$$

$$c_{1,i+1} = c_{1,i} + a_0 + \phi_{0,i+1} \bmod 2^{32},$$

$$c_{2,i+1} = c_{2,i} + a_0 + \phi_{1,i+1} \bmod 2^{32},$$

$$c_{3,i+1} = c_{3,i} + a_0 + \phi_{2,i+1} \bmod 2^{32},$$

$$c_{4,i+1} = c_{4,i} + a_0 + \phi_{3,i+1} \bmod 2^{32},$$

$$c_{5,i+1} = c_{5,i} + a_0 + \phi_{4,i+1} \bmod 2^{32},$$

$$c_{6,i+1} = c_{6,i} + a_0 + \phi_{5,i+1} \bmod 2^{32},$$

$$c_{7,i+1} = c_{7,i} + a_0 + \phi_{6,i+1} \bmod 2^{32},$$

где бит счетчика по переносу $\phi_{j,i+1}$ задан выражением

$$\phi_{j,i+1} = \begin{cases} 1 & \text{если } c_{0,i} + a_0 + \phi_{7,i} \geq 2^{32} \wedge j = 0 \\ 1 & \text{если } c_{j,i} + a_j + \phi_{j-1,i+1} \geq 2^{32} \wedge j > 0 \\ 0 & \text{иначе} \end{cases}$$

Кроме того, константы a_j определены как:

$$a_0 = 0x4D34D34D, a_1 = 0xD34D34D3,$$

$$a_2 = 0x34D34D34, a_3 = 0x4D34D34D,$$

$$a_4 = 0xD34D34D3, a_5 = 0x34D34D34,$$

$$a_6 = 0x4D34D34D, a_7 = 0xD34D34D3.$$

После каждой итерации результат извлекается следующим образом:

$$s_i^{[15..0]} = x_{0,i}^{[15..0]} \oplus x_{5,i}^{[31..16]}, s_i^{[31..16]} = x_{0,i}^{[31..16]} \oplus x_{3,i}^{[15..0]},$$

$$s_i^{[47..32]} = x_{2,i}^{[15..0]} \oplus x_{7,i}^{[31..16]}, s_i^{[63..48]} = x_{2,i}^{[31..16]} \oplus x_{5,i}^{[15..0]},$$

$$s_i^{[79..64]} = x_{4,i}^{[15..0]} \oplus x_{1,i}^{[31..16]}, s_i^{[95..80]} = x_{4,i}^{[31..16]} \oplus x_{7,i}^{[15..0]},$$

$$s_i^{[111..96]} = x_{6,i}^{[15..0]} \oplus x_{3,i}^{[31..16]}, s_i^{[127..112]} = x_{6,i}^{[31..16]} \oplus x_{1,i}^{[15..0]}.$$

Поточный шифр Salsa20

Ядром шифра Salsa20 является хеш-функция с 64-байтовым входом и 64-байтовым выходом [6]. Хеш-функция в режиме счетчика используется как поточный шифр: Salsa20 шифрует 64-байтовый блок открытого текста хешированием ключа, в данном случае, и номера блока, складывая результат по модулю 2 (XOR) с открытым текстом.

Хеш-функция Salsa20

Хеш-функция $Salsa20(x)$ определяется следующим выражением:

$$Salsa20(x) = x + \text{doubleround}^{10}(x),$$

где каждая 4-байтовая последовательность x рассматривается как слово в форме littleendian.

Если $b = (b_0, b_1, b_2, b_3)$ – 32-битовое слово, где b_3 и b_0 обозначают соответственно самый старший байт и самый младший байт величины b , тогда

$$\text{littleendian}(b) = 2^{24}b_3 + 2^{16}b_2 + 2^8b_1 + b_0.$$

Функция $\text{doubleround}(x)$ вычисляется путем последовательного применения к последовательности x из 16 слов (слово – 32-битовый элемент) функций $\text{columnround}(x)$ и $\text{rowround}(x)$:

$$\text{doubleround}(x) = \text{rowround}(\text{columnround}(x)).$$

Функций $\text{columnround}(x)$ и $\text{rowround}(x)$ в свою очередь строятся на основе функции $\text{quarterround}(y)$. Функция $\text{quarterround}(y)$ оперирует последовательностями из 4 слов.

Если $y = (y_0, y_1, y_2, y_3)$, тогда $\text{quarterround}(y) = (z_0, z_1, z_2, z_3)$, где

$$z_1 = y_1 \oplus ((y_0 + y_3) \lll 7),$$

$$z_2 = y_2 \oplus ((z_1 + y_0) \lll 9),$$

$$z_3 = y_3 \oplus ((z_2 + z_1) \lll 13),$$

$$z_0 = y_0 \oplus ((z_3 + z_2) \lll 18).$$

Функцию quarterround можно представить как изменение y следующим образом: сначала y_1 изменяется на z_1 , затем y_2 изменяется на z_2 , затем y_3 изменяется на z_3 , затем y_0 изменяется на z_0 . Каждое изменение является обратимым, таким образом, вся функция является обратимой.

Функция $\text{rowround}(y)$ оперирует последовательностями из 16 слов.

Если $y = (y_0, y_1, y_2, y_3, \dots, y_{15})$, тогда $\text{rowround}(y) = (z_0, z_1, z_2, z_3, \dots, z_{15})$, где

$$(z_0, z_1, z_2, z_3) = \text{quarterround}(y_0, y_1, y_2, y_3),$$

$$(z_5, z_6, z_7, z_4) = \text{quarterround}(y_5, y_6, y_7, y_4),$$

$$(z_{10}, z_{11}, z_8, z_9) = \text{quarterround}(y_{10}, y_{11}, y_8, y_9),$$

$$(z_{15}, z_{12}, z_{13}, z_{14}) = \text{quarterround}(y_{15}, y_{12}, y_{13}, y_{14}).$$

Можно представить вход $(y_0, y_1, \dots, y_{15})$ в виде квадратной матрицы:

$$\begin{pmatrix} y_0 & y_1 & y_2 & y_3 \\ y_4 & y_5 & y_6 & y_7 \\ y_8 & y_9 & y_{10} & y_{11} \\ y_{12} & y_{13} & y_{14} & y_{15} \end{pmatrix}$$

Функция rowround изменяет строки матрицы параллельно, пропуская перестановку каждой строки через функцию quarterround . В первой строке функция rowround изменяет y_1 , затем y_2 , затем y_3 , затем y_0 ; во второй строке функция rowround изменяет y_6 , затем y_7 , затем y_4 , затем y_5 ; в третьей строке функция rowround изменяет y_{11} , затем y_8 , затем y_9 , затем y_{10} ; в четвертой строке функция rowround изменяет y_{12} , затем y_{13} , затем y_{14} , затем y_{15} .

Функция $\text{columnround}(x)$ также как и функция $\text{rowround}(y)$ оперирует последовательностями из 16 слов.

Если $x = (x_0, x_1, x_2, x_3, \dots, x_{15})$ тогда $\text{columnround}(x) = (y_0, y_1, y_2, y_3, \dots, y_{15})$, где

$$(y_0, y_4, y_8, y_{12}) = \text{quarterround}(x_0, x_4, x_8, x_{12}),$$

$$(y_5, y_9, y_{13}, y_1) = \text{quarterround}(x_5, x_9, x_{13}, x_1),$$

$$(y_{10}, y_{14}, y_2, y_6) = \text{quarterround}(x_{10}, x_{14}, x_2, x_{16}),$$

$$(y_{15}, y_3, y_7, y_{11}) = \text{quarterround}(x_{15}, x_3, x_7, x_{11}).$$

Эквивалентная формула: $(y_0, y_4, y_8, y_{12}, y_1, y_5, y_9, y_{13}, y_2, y_6, y_{10}, y_{14}, y_3, y_7, y_{11}, y_{15}) = \text{rowround}(x_0, x_4, x_8, x_{12}, x_1, x_5, x_9, x_{13}, x_2, x_6, x_{10}, x_{14}, x_3, x_7, x_{11}, x_{15})$.

Можно представить вход $(x_0, x_1, \dots, x_{15})$ в виде квадратной матрицы:

$$\begin{pmatrix} x_0 & x_1 & x_2 & x_3 \\ x_4 & x_5 & x_6 & x_7 \\ x_8 & x_9 & x_{10} & x_{11} \\ x_{12} & x_{13} & x_{14} & x_{15} \end{pmatrix}$$

Функция columnround с этого ракурса представляется просто заменой (транспонированием) функции rowround . Функция columnround , изменяет столбцы матрицы параллельно, пропуская перестановку каждого столбца через функцию quarterround . В первом столбце, функция columnround изменяет y_4 , затем y_8 , затем y_{12} , затем y_0 ; во втором столбце, функция columnround изменяет y_9 , затем y_{13} , затем y_1 , затем y_5 ; в третьем столбце, функция columnround изменяет y_{14} , затем y_2 , затем y_6 , затем y_{10} ; в четвертом столбце, функция columnround изменяет y_3 , затем y_7 , затем y_{11} , затем y_{15} .

Инициализация

Если ключ k – 32-байтовая или 16-байтовая последовательность, а iv – 16-байтовая последовательность, тогда $\text{Salsa20}_k(iv)$ является 64-байтовой последовательностью.

Определим $\sigma_0 = (101, 120, 112, 97)$, $\sigma_1 = (110, 100, 32, 51)$, $\sigma_2 = (50, 45, 98, 121)$, и $\sigma_3 = (116, 101, 32, 107)$. Если k_0, k_1, iv являются 16-байтовыми последовательностями, тогда

$$\text{Salsa20}_{k_0, k_1}(iv) = \text{Salsa20}(\sigma_0, k_0, \sigma_1, iv, \sigma_2, k_1, \sigma_3).$$

Определим $\tau_0 = (101, 120, 112, 97)$, $\tau_1 = (110, 100, 32, 49)$, $\tau_2 = (54, 45, 98, 121)$, и $\tau_3 = (116, 101, 32, 107)$. Если k, iv являются 16-байтовыми последовательностями, тогда

$$\text{Salsa20}_k(iv) = \text{Salsa20}(\tau_0, k, \tau_1, iv, \tau_2, k, \tau_3).$$

Константы $\sigma_0\sigma_1\sigma_2\sigma_3$ и $\tau_0\tau_1\tau_2\tau_3$, равны “expand 32-byte k ” и “expand 16-byte k ” (в ASCII) соответственно.

Функция шифрования Salsa20

Пусть k – это 32- или 16-байтовый секретный ключ, iv – 8-байтовый вектор инициализации, m – открытый текст. Тогда шифрование последовательности m шифром Salsa20 с помощью вектор инициализации iv на ключе k , обозначается $Salsa20_k(iv) \oplus m$ – зашифрованный текст. Либо m может быть зашифрованным текстом, в случае, когда $Salsa20_k(iv) \oplus m$ является исходным открытым текстом.

Поточный шифр SOSEMANUK

Sosemanuk – синхронный поточный шифр [6]. Шифр Sosemanuk использует принципы базовой конструкции поточного шифра SNOW 2.0 преобразования блочного шифра SERPENT. Поэтому разработчики решили, что название этого шифра должно относиться и к SERPENT и к SNOW. Однако, известно, что снежных змей не существуют, т.к. змеи либо впадают в спячку, либо перемещаются в более теплые страны на зимы. С другой стороны Sosemanuk – популярный спорт, в который играют восточные канадские племена. Он состоит в броске деревянной палки по сугробу как можно дальше. Его название на языке Cree означает snowsnake, так как палка на снегу похожа на змею. Kwakweco-cime win – вариант той же самой игры, но для названия шифра не подходит.

Sosemanuk – синхронный поточный шифр, ориентированный на программную реализацию. Размер ключа варьируется между 128 и 256 битами. Утверждается, что при любой длине ключа достигается 128-битная безопасность. Sosemanuk стремится улучшить SNOW 2.0 в двух отношениях. Во-первых, в Sosemanuk избегаются некоторые свойства структуры, которые могут проявиться как потенциальные слабости, даже если шифр SNOW 2.0 с 128-битным ключом сопротивляется всем известным нападениям. Во-вторых, эффективность улучшена относительно нескольких архитектур, уменьшением размера внутреннего состояния, таким образом допускается более прямое отображение данных относительно регистров процессора. Sosemanuk также требует меньшего количества статических данных; соответственно более низкого использования кэша данных, что приводит к лучшей работе на некоторых архитектурах. Другое достоинство Sosemanuk – то, что его процедура установки ключа основана на сокращенной версии блочного шифра SERPENT, улучшая классические процедуры инициализации и с точки зрения эффективности и с точки зрения безопасности.

SERPENT и его производные

SERPENT – блочный шифр, предложенный в качестве AES кандидата. SERPENT работает с 128-битовыми блоками, которые разбиваются на четыре 32-разрядных слова, а

затем объединяются в так называемом “секционном” (“bitslice”) режиме. Таким образом SERPENT может быть определен как шифр, работающий с четверкой 32-битовых слов. Мы нумеруем входные и выходные четверки SERPENT’а от 0 до 3, и пишем им в порядке: (Y_3, Y_2, Y_1, Y_0) . Y_0 – самое младшее слово, содержащее младшие разряды 32-ух 4-битовых входных данных для S-блоков SERPENT’а. После того как выходной поток SERPENT’а записывается в 16 байтов, в значения Y_i записываются в соответствии со следующим условным обозначением littleendian (вначале самый младший байт), и сначала выводится Y_0 , затем Y_1 , и так далее.

Из SERPENT разработчики определили два примитива по имени Serpent1 и Serpent24.

Циклы SERPENT’а состоят из:

- добавления подключа поразрядным исключаяющим или (XOR);
- применения s -блока, которое заключается в ряде поразрядных комбинаций между четырьмя обрабатываемыми 32-битовыми словами, в секционном режиме (bitslice mode);
- линейного биективного преобразования, которое равнозначно нескольким XOR’ам, сдвигам и циклическим сдвигам в секционном режиме (bitslice mode).

Serpent1 – один раунд шифра SERPENT, без добавления ключа и линейного преобразования. SERPENT использует восемь различных S -блоков, пронумерованных от S_0 до S_7 , рассчитанных на 4-битовые слова. Мы определяем Serpent1 как применение S_2 , в секционном режиме (bitslice mode). Это третий S -блоковый уровень шифра SERPENT. Serpent1 использует в качестве входных данных четыре 32-битовых слова, и вырабатывает четыре 32-битовых слова в качестве выхода.

Serpent24 – это SERPENT, сокращенный до 24 раундов, вместо полной 32-х раундовой версии SERPENT. Serpent24 соответствует первым 24 раундам шифра SERPENT, причем последний (24-й) раунд полный – с линейным преобразованием и XOR’ом с 25-ым подключом. Другими словами, 24-ый раунд Serpent24 эквивалентен тридцать второму раунду шифра SERPENT, за исключением того, что содержит линейное преобразование и использует 24-ый и 25-ый подключи (32-ый и 33-ий подключи в SERPENT). Соответствующее уравнение последнего раунда приведено в [32_Serpent]

$$R_{23}(X) = L \left(S_{23} \left(X \oplus \hat{K}_{23} \right) \right) \oplus \hat{K}_{24}$$

Serpent24 использует только 25 128-битовых подключей, которые являются первыми 25 подключами, производящимися в соответствии со схемой разворачивания ключа шифра

SERPENT. В Sosemanuk Serpent24 используется на этапе инициализации, только в режиме шифрования. При расшифровании не используется.

Инициализация

Процесс инициализации Sosemanuk разбит на два шага:

- схема разворачивания ключа, которая подвергает обработке секретный ключ, но не зависит от вектора инициализации IV ;
- добавление вектора инициализации IV , которое использует выход схемы разворачивания ключа и вектор IV .

Таким образом, инициализируется внутреннее состояние поточного шифра.

Установка ключа соответствует схеме разворачивания ключа в Serpent24, которая производит 25 128-битовых подключей, как 100 32-битовых слова. Эти 25 128-битовых подключей идентичны первым 25 128-битовым подключам, производящимся в соответствии с простой схемой разворачивания ключа в SERPENT.

SERPENT допускает любую длину ключа длину от 1 до 256 битов; следовательно, Sosemanuk может работать с точно такими же ключами. Однако, так как Sosemanuk стремится к 128-битовой безопасности, длина его ключа должна быть по крайней мере 128 битов. Поэтому, 128 битов – стандартная длина ключа. Поддерживается любая длина ключа от 128 битов до 256 битов. Но, уровень безопасности все же соответствует 128-битовой безопасности. Другими словами, использование более длинного секретного ключа не гарантирует обеспечение уровня безопасности, обычно ожидаемого от такого ключа.

Вектор IV является 128-битовым значением. Используется в качестве входа для блочного шифра Serpent24, инициализированный в соответствии со схемой разворачивания ключа. Serpent24 состоит из 24 раундов, используются выходы 12-го, 18-го и 24-го раундов. Мы обозначим эти выходы следующим образом:

- $(Y_3^{12}, Y_2^{12}, Y_1^{12}, Y_0^{12})$ – выход 12-ого раунда;
- $(Y_3^{18}, Y_2^{18}, Y_1^{18}, Y_0^{18})$ – выход 18-ого раунда;
- $(Y_3^{24}, Y_2^{24}, Y_1^{24}, Y_0^{24})$ – выход 24-ого раунда.

Выход каждого раунда состоит из четырех 32-битовых слов, берущихся только после линейного преобразования, за исключением 24-ого раунда, для которого выход берется только после добавления 25-ого подключа.

Эти значения используются для инициализации внутреннего состояния Sosemanuk'a со следующими значениями:

$$(s_7, s_8, s_9, s_{10}) = (Y_3^{12}, Y_2^{12}, Y_1^{12}, Y_0^{12})$$

$$(s_5, s_6) = (Y_1^{18}, Y_3^{18})$$

$$(s_1, s_2, s_3, s_4) = (Y_3^{24}, Y_2^{24}, Y_1^{24}, Y_0^{24})$$

$$R1_0 = Y_0^{18}$$

$$R2_0 = Y_2^{18}$$

Генерация ключевого потока

Для формирования выходных значений z_t используется конечный автомат (FSM, Finite State Machine) и функция Serpent1.

Конечный автомат (FSM) – компонент с 64 битами памяти, аналогичный двум 32-битовым регистрам $R1$ и $R2$. В каждом шаге FSM берет в качестве входных данных некоторые слова из состояния LFSR; обновляет биты памяти и производит 32-битовый выход. FSM оперирует состоянием LFSR в моменты времени $t \geq 1$ следующим образом:

$$FSM_t: (R1_{t-1}, R2_{t-1}, s_{t+1}, s_{t+8}, s_{t+9}) \mapsto (R1_t, R2_t, f_t)$$

где

$$R1_t = (R2_{t-1} + \text{mux}(\text{lsb}(R1_{t-1}), s_{t+1}, s_{t+1} \oplus s_{t+8})) \bmod 2^{32}$$

$$R2_t = \text{Trans}(R1_{t-1})$$

$$f_t = (s_{t+9} + R1_t \bmod 2^{32}) \oplus R2_t$$

где $\text{lsb}(x)$ – младший бит x ,

$\text{mux}(c, x, y)$ равен x если $c = 0$, или y если $c = 1$.

Внутренняя переходная функция Trans над полем $F_{2^{32}}$ определяется как

$$\text{Trans}(z) = (M \times z \bmod 2^{32})_{\ll\ll 7}$$

где M – постоянное значение, равное $0x54655307$ (шестнадцатеричное выражение первых десяти десятичных чисел π).

Выходы конечного автомата FSM группируются по четыре, и Serpent1 применяется к каждой группе; затем результат объединяется XOR'ом с соответствующими отбрасываемыми значениями LFSR, для производства выходных значений z_t :

$$(z_{t+3}, z_{t+2}, z_{t+1}, z_t) = \text{Serpent1}(f_{t+3}, f_{t+2}, f_{t+1}, f_t) \oplus (s_{t+3}, s_{t+2}, s_{t+1}, s_t)$$

Четыре последовательных раунда Sosemanuk изображены на рисунке 2.26

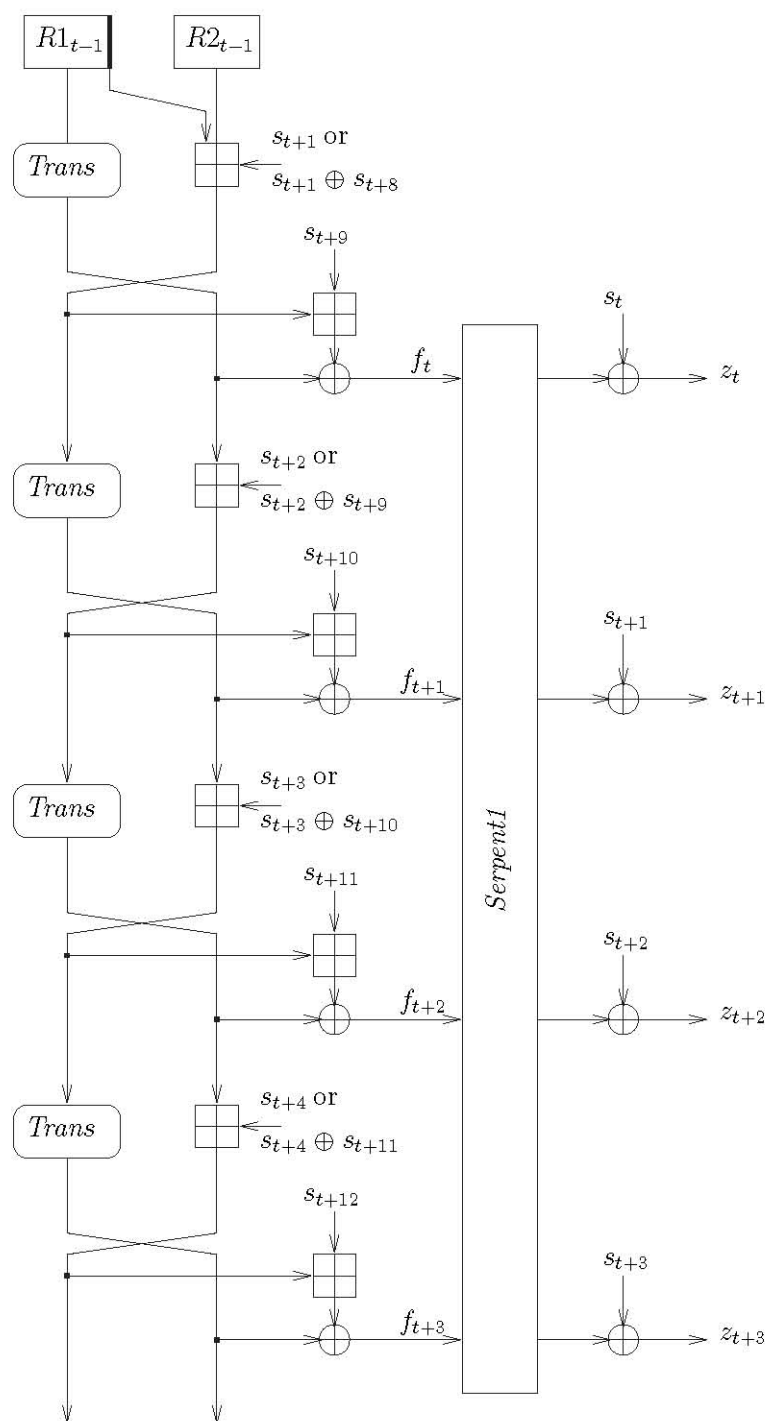


Рис. 7.26. Выходное преобразование четырех последовательных раундов Sosemanuk

Шифр Sosemanuk объединяет FSM и LFSR для производства выходных значений z_t . Время $t = 0$ определяет внутреннее состояние после инициализации; первое выходное значение — z_1 . На рисунке 3 представлено описание шифра Sosemanuk.

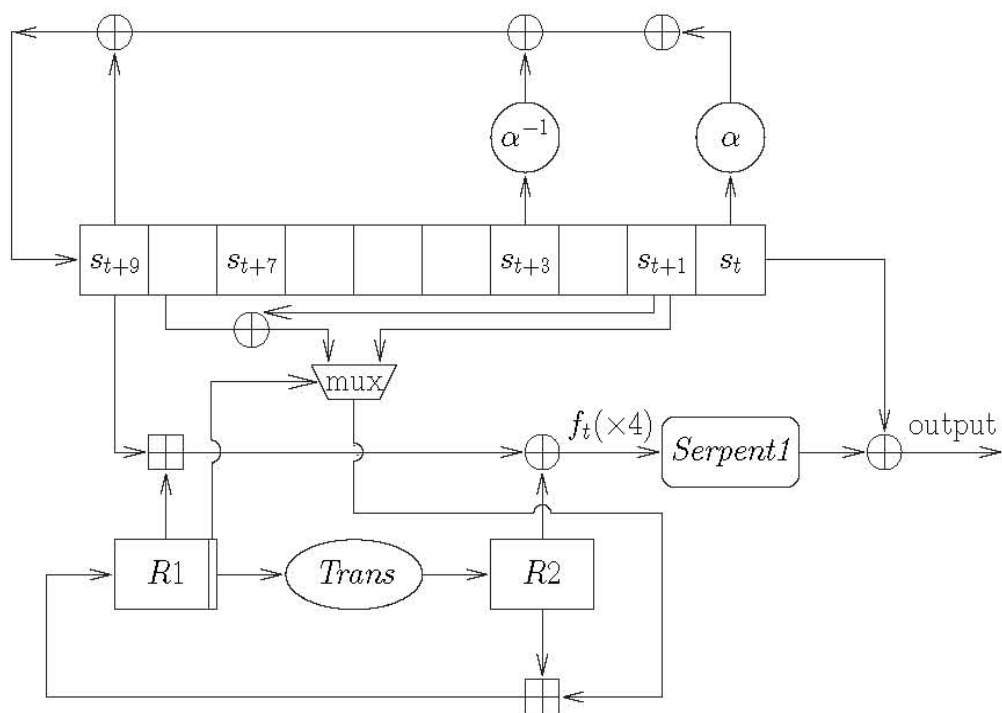


Рис. 7.27. Описание шифра Sosemanuk

В моменты времени $t \geq 1$ выполняются следующие операции:

- обновляется FSM: из $R1_{t-1}$, $R2_{t-1}$, s_{t+1} , s_{t+8} и s_{t+9} вычисляются $R1_t$, $R2_t$ и промежуточные значения f_t .
- обновляется LFSR: из s_t , s_{t+3} и s_{t+9} вычисляется s_{t+10} . Значение s_t передается внутреннему буферу, LFSR сдвигается.

Каждые четыре шага, из накопленных значений $f_t, f_{t+1}, f_{t+2}, f_{t+3}$ и $s_t, s_{t+1}, s_{t+2}, s_{t+3}$ производятся четыре выходных значения z_t, z_{t+1}, z_{t+2} , и z_{t+3} . Таким образом, Sosemanuk производит 32-битовые значения.

Соответственно первые четыре итерации Sosemanuk'a следующие:

- Начальное состояние LFSR содержит значения s_1, \dots, s_{10} ; значение s_0 не определено. Начальное состояние FSM содержит $R1_0$ и $R2_0$.
- В течение первого шага из $R1_0, R2_0, s_2, s_9$ и s_{10} вычисляются $R1_1, R2_1$ и f_1 .
- Первый шаг производит промежуточные значения s_1 и f_1 , сохраняемые в буфере.
- В течение первого шага, вычисляется слово обратной связи s_{11} из s_{10}, s_4 и s_1 , обновляется внутреннее состояние LFSR, что приводит к новому состоянию, формируемое из s_2, \dots, s_{11} .

- Первые четыре выходные значения z_1, z_2, z_3 и z_4 , вычисляются, используя однократное применение Serpent1'a к (f_4, f_3, f_2, f_1) , выход которого объединяется с (s_4, s_3, s_2, s_1) при помощи XOR.

Поточный шифр F-FCSR-H

F-FCSR-H – аддитивный поточный шифр [5].

Для генерации ключевого потока из ключа длиной 80 бит и вектора инициализации длиной от 32 до 80 бит используется фильтрующий автомат FCSR.

Генерация ключевого потока

В основе шифра F-FCSR-H лежит регистр сдвига с обратной связью по переносу (FCSR) – автомат, который вычисляет двоичное разложение 2-адического числа p/q , где p и q – некоторые целые числа, с q нечетное. Допустим, что $q < 0 < p < |q|$. Размер FCSR n такой, что $n + 1$ является длиной q в битах.

В данном шифре p зависит от секретного ключа (и IV), а q – открытый параметр. Выбор q порождает много свойств ключевого потока. Самое важное – то, что это полностью определяет длину периода ключевого потока. Условия для оптимального выбора:

- q – (отрицательное) простое число размером $(n + 1)$ бит.
- $(|q| - 1)$ – порядок 2 по модулю q .
- $T = (q - 1)/2$ также является простым числом.
- $d = (1 + |q|)/2$. $W(d)$ – вес Хемминга двоичного разложения, $W(d) > n/2$.

Автомат FCSR содержит два регистра (наборы ячеек): основной регистр M и регистр переноса C . Основной регистр M содержит n ячеек. Обозначим m_i ($0 \leq i \leq n - 1$) двоичные знаки $(n-1)$, содержащиеся в этих ячейках, и назовем числами $m = \sum_{i=0}^{n-1} m_i 2^i$ содержимое (или состояние) регистра M .

Пусть d – положительное целое число $d = (1 - q)/2$, а $d = \sum_{i=0}^{n-1} d_i 2^i$ его двоичное разложение. Регистр переносов содержит l ячеек, где $(l + 1)$ – количество чисел d_i отличных от нуля. Более строго, регистр переносов содержит одну ячейку для каждого ненулевого d_i , для $0 \leq i \leq (n - 2)$. Обозначим c_i – двоичный знак, содержащийся в этой ячейке. Мы также устанавливаем $c_i = 0$ когда $d_i = 0$ или когда $i = (n - 1)$. Назовем числом $c = \sum_{i=0}^{n-2} c_i 2^i$ содержимое (или состояние) регистра C . Вес Хемминга двоичного разложения c не больше, чем l . Функция перехода может быть описана выражениями

$$m(t+1) = (m(t) \gg 1) \oplus c(t) \oplus m_0(t)d,$$

$$c(t+1) = (m(t) \gg 1) \otimes c(t) \oplus c(t) \otimes m_0(t)d \oplus m_0(t)d \otimes (m(t) \gg 1).$$

Отметим, что $m_0(t)$ – самый младший бит в $m(t)$. Числа $m(t)$, $c(t)$ и d – целые числа размером в n бит (или меньше).

Для извлечения псевдослучайных бит ключевого потока из основного регистра автомата FCSR используется фильтр. Этот фильтр описывает, какие ячейки выбраны для производства псевдослучайных битов. Чтобы получить мультиразрядный выход, используются восемь или шестнадцать одноразрядных фильтров, чтобы извлечь 8- или 16-битовые выходные слова после каждого перехода автомата.

Фильтр F – это битовая строка (f_0, \dots, f_{n-1}) длиной n (что эквивалентно числу $\sum_{i=0}^{n-1} f_i 2^i$).

Выходной бит z получается вычислением веса *parity* поразрядного И состояния M основного регистра и фильтра F :

$$z = \bigoplus_{i=0}^{n-1} f_i m_i.$$

Это эквивалентно следующему:

$$S = M \otimes F,$$

$$z = \text{parity}(S).$$

Аналогичным способом, предлагается метод извлечения s -битового слова из состояния FCSR. Значение s будет равно 8 для F-FCSR-H, и 16 для F-FCSR-16.

Фильтр F также является битовой строкой (f_0, \dots, f_{n-1}) длиной n (которая является кратной числу s). Он разбивается на s подфильтров F_0, \dots, F_{s-1} каждый определяется как

$$F_j = \sum_{i=0}^{n/s-1} f_{si+j} 2^i.$$

Каждый подфильтр F_j выбирает несколько ячеек m_i в основном регистре среди тех, что удовлетворяют выражению $i \equiv j \pmod{s}$. *Parity* полученного двоичного слова дает j -й псевдослучайный бит:

$$z_j = \bigoplus_{i=0}^{n/s-1} f_{si+j} m_{si+j}.$$

Так как есть s подфильтров, то мы получаем s битов при каждом переходе автомата.

Эта процедура может быть описана эквивалентно следующим образом. Фильтр F и состояние M комбинируются функцией И. Результат разбивается на n/s слова. Псевдослучайное слово получается операцией XOR этих n/s слов:

$$S = M \otimes F$$

Определим S_i с помощью выражения $S = \sum_{i=0}^{n/s-1} S_i \cdot 2^{si}$, для $0 \leq S_i \leq 2^8 - 1$.

Выходное слово z будет равно:

$$z = \bigoplus_{i=0}^{n/s-1} S_i.$$

Отметим, что целое слово извлекается быстрее, чем отдельный бит.

Шифр F-FCSR-N использует ключи длиной 80 бит и IV размером от 32 до 80 бит. Если IV не используется, то по умолчанию можно использовать значение 0.

Длина FCSR (размер основного регистра) – $n = 160$. Регистр переносов содержит $l = 82$ ячейки. Обратное простое число

$$q = -1993524591318275015328041611344215036460140087963$$

таким образом сложение полей и ячеек переносов присутствуют в позициях, соответствующих тем (кроме ведущего) в следующей строке 160 битов (которая имеет вес Хемминга 83),

$$d = (1 + |q|)/2 = (\text{AE985DFF 26619FC5 8623DC8A AF46D590 3DD4254E})_{16}.$$

Фильтрация

Чтобы извлечь один псевдослучайный байт, мы используем статический фильтр

$$F = d = (\text{AE985DFF 26619FC5 8623DC8A AF46D590 3DD4254E})_{16}$$

Фильтр F разбит на 8 подфильтров (подфильтр j получен выбором бита j в каждом байте F),

$$F_0 = (0011\ 0111\ 0100\ 1010\ 1010)_2,$$

$$F_1 = (1001\ 1010\ 1101\ 1100\ 0001)_2,$$

$$F_2 = (1011\ 1011\ 1010\ 1110\ 1111)_2,$$

$$F_3 = (1111\ 0010\ 0011\ 1000\ 1001)_2,$$

$$F_4 = (0111\ 0010\ 0010\ 0011\ 1100)_2,$$

$$F_5 = (1001\ 1100\ 0100\ 1000\ 1010)_2,$$

$$F_6 = (0011\ 0101\ 0010\ 0110\ 0101)_2,$$

$$F_7 = (1101\ 0011\ 1011\ 1011\ 0100)_2.$$

Повторный вызов бита b_i ($0 \leq i \leq 7$) каждого извлеченного байта выражается

$$b_i = \bigoplus_{j=0}^{19} f_i^{(j)} m_{8j+i}, \text{ где } F_i = \sum_{j=0}^{19} f_i^{(j)} 2^j$$

где m_k – биты, содержащиеся в основном регистре.

Инициализация

Инициализация шифра F-FCSR-N производится в следующем порядке:

$$v \leq 80)$$

1. Основной регистр M инициализируется ключом и IV :

$$M := K + 2^{80} \cdot IV = (0^{80-v} \| IV \| K)$$

2. Регистр переносов инициализируется в 0:

$$C := 0 = (0^{82})$$

3. FCSR тактируется 160 раз. (На этом шаге выход отбрасывается),

После фазы установки, псевдослучайный поток производится тактированием FCSR и извлечением по одному псевдослучайных байт, используя фильтр F , как описано выше.

Поточный шифр Grain-128

Grain – двоичный аддитивный поточный шифр, ориентированный на аппаратную реализацию. Версия, которая обозначается Grain v.1 [6], предназначена для приложений, которые имеют очень ограниченные аппаратные ресурсы. Grain v.1 поддерживает 80-битовый размер ключа.

Версия нифра Grain-128 поддерживает размер ключа – 128 бит и размер вектора IV – 96 битов. Шифр Grain-128 также является очень компактным и легко осуществимым в аппаратных средствах. Кроме того, возможно достаточно просто увеличивать скорость за счет большего количества аппаратных средств. Это является отличительной особенностью семейства поточных шифров Grain, и во многих других шифрах явно не обосновано. Grain-128 использует линейный регистр сдвига с обратной связью, чтобы гарантировать хорошие статистические свойства и гарантировать нижнюю границу периода ключевой последовательности. Чтобы ввести нелинейность вместе с нелинейным фильтром используется нелинейный регистр сдвига с обратной связью (NFSR). Нелинейный фильтр берет в качестве входных данных значения от обоих сдвиговых регистров.

Генерация ключевого потока

Краткий обзор различных блоков, используемых в шифре, приведен на рисунке 2.28.

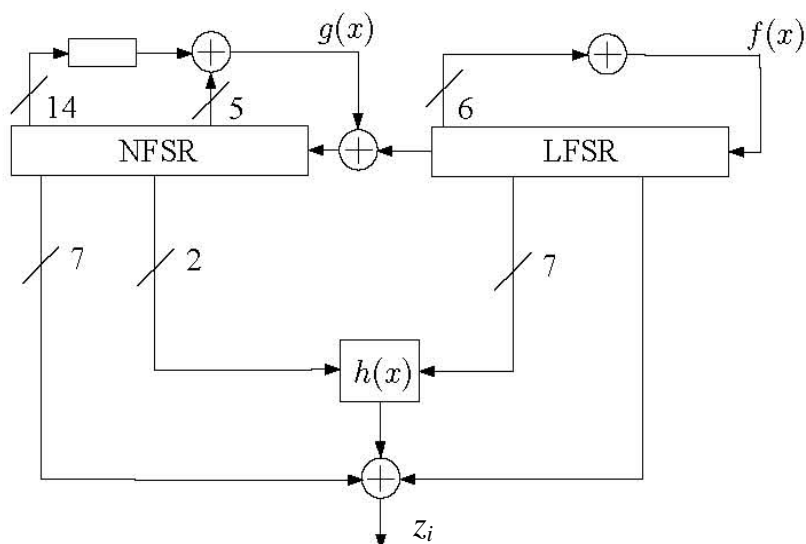


Рис. 7.28. Краткий обзор шифра.

Шифр состоит из трех основных строительных блоков, а именно: LFSR, NFSR и выходной функции. Содержимое LFSR обозначается $s_i, s_{i+1}, \dots, s_{i+127}$. Аналогичным образом содержание NFSR обозначается $b_i, b_{i+1}, \dots, b_{i+127}$. Полином обратной связи LFSR, обозначенный $f(x)$, является примитивным полиномом степени 128. Он определен как

$$f(x) = 1 + x^{32} + x^{47} + x^{58} + x^{90} + x^{121} + x^{128}.$$

Чтобы избежать любой возможной неопределенности, разработчики в [6] привели соответствующую функцию обновления LFSR:

$$s_{i+128} = s_i + s_{i+7} + s_{i+38} + s_{i+70} + s_{i+81} + s_{i+96}.$$

Нелинейный полином обратной связи NFSR, $g(x)$, является суммой одной линейной и одной нелинейной функций. Он определен как

$$g(x) = 1 + x^{32} + x^{37} + x^{72} + x^{102} + x^{128} + x^{44}x^{60} + x^{61}x^{125} + \\ + x^{63}x^{67} + x^{69}x^{101} + x^{80}x^{88} + x^{110}x^{111} + x^{115}x^{117}.$$

Аналогично, чтобы избежать любой возможной неопределенности, разработчики также привели в [6] соответствующую функцию обновления NFSR. В функции обновления, приведенной ниже, бит s_i , опущенный в полиноме обратной связи, включен и замаскирован со входом NFSR.

$$b_{i+128} = s_i + b_i + b_{i+26} + b_{i+56} + b_{i+91} + b_{i+96} + b_{i+3}b_{i+67} + b_{i+11}b_{i+13} + \\ + b_{i+17}b_{i+18} + b_{i+27}b_{i+59} + b_{i+40}b_{i+48} + b_{i+61}b_{i+65} + b_{i+68}b_{i+84}.$$

256 элементов памяти в этих двух сдвиговых регистрах представляют состояние шифра. Из этого состояния берется 9 переменных в качестве входа булевой функции $h(x)$. Два входа $h(x)$ берутся из NFSR, а семь – из LFSR. Эта функция имеет степень 3 и является очень простой. Она определяется как

$$h(x) = x_0x_1 + x_2x_3 + x_4x_5 + x_6x_7 + x_0x_4x_8,$$

где переменные $x_0, x_1, x_2, x_3, x_4, x_5, x_6, x_7$ и x_8 соответствуют позициям сигнала $b_{i+12}, s_{i+8}, s_{i+13}, s_{i+20}, b_{i+95}, s_{i+42}, s_{i+60}, s_{i+79}$ и s_{i+95} соответственно. Выходная функция определяется как

$$z_i = \sum_{j \in A} b_{i+j} + h(x) + s_{i+93},$$

где $A = \{2, 15, 36, 45, 64, 73, 89\}$.

Инициализация

Перед генерацией ключевой последовательности шифр должен быть инициализирован с помощью ключа и вектора IV . Пусть биты ключа k будут обозначаться $k_i, 0 \leq i \leq 127$, а биты вектора IV будут обозначаться $IV_i, 0 \leq i \leq 95$. Тогда инициализация ключа и вектора IV выполняется следующим образом. 128 элементов NFSR заполняются битами ключа, $b_i = k_i, 0 \leq i \leq 127$, затем первые 96 элементов LFSR заполняются битами IV битами, $s_i = IV_i, 0 \leq i \leq 95$. Последние 32 бита LFSR заполняются единицами, $s_i = 1, 96 \leq i \leq 127$. После загрузки битов ключа и IV , шифр тактируется 256 раз, без производства ключевой последовательности. Вместо выходной функции производится подача назад и сложение по модулю 2 (XOR'тся) со входом и LFSR и NFSR, см. рисунок 7.29

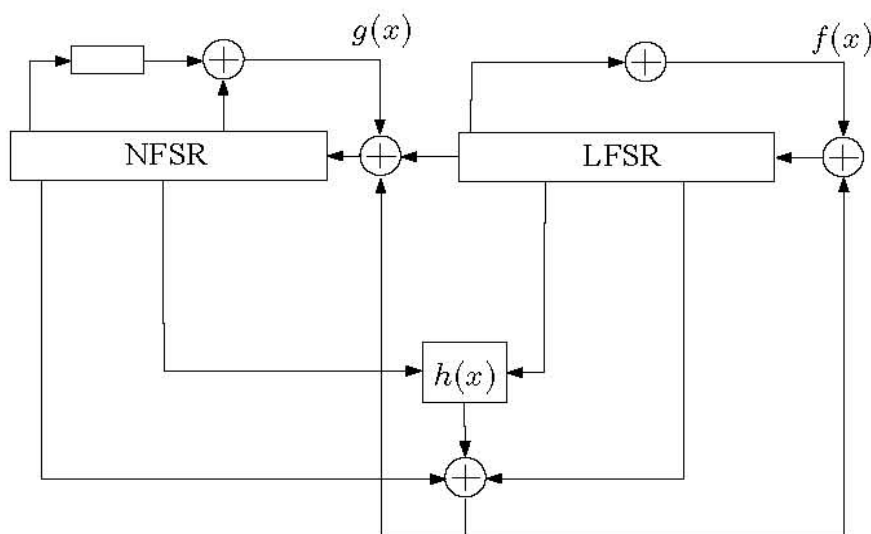


Рис. 7.29. Инициализация ключа

Поточный шифр MICKEY-128

MICKEY-128 – поточный шифр, ориентированный на аппаратную реализацию. Для генерации ключевого потока MICKEY-128 использует 128 битовый ключ [6]. Этот шифр использует нерегулярно тактируемые регистры сдвига с новыми методиками, предназначенными обеспечить баланс между необходимостью в гарантированных

периоде и псевдослучайностью и необходимостью избежать определенных криптоаналитических атак.

Инициализация

MISKEY128 2.0 использует два входных параметра:

- 128-битовый секретный ключ K , биты которого обозначаются $k_0 \dots k_{127}$;
- вектор инициализации IV (initialisation variable), длиной от 0 до 128 битов, биты которой обозначаются $iv_0 \dots iv_{IVLENGTH-1}$, где $IVLENGTH$ – размер вектора инициализации в битах.

Генератор строится из двух регистров R и S . Каждый регистр длиной 160 разрядов. Биты в регистрах обозначаются $r_0 \dots r_{159}$ и $s_0 \dots s_{159}$ соответственно.

Разработчики заявляют, что R – это “линейный регистр”, а S – это “нелинейный регистр”.

Инициализация генератора начинается с обнуления всех разрядов регистров R и S . Затем производится загрузка вектора инициализации IV путем тактирования всего генератора $IVLENGTH$ раз с помощью операции $clock_KG(R, S, Mixing = True, Input_bit = iv_i)$, описанной в п.1.3.8.2. Затем производится загрузка ключа: весь генератор тактируется 128 раз, так же с помощью операции $clock_KG(R, S, Mixing = True, Input_bit = k_i)$. Заканчивается процесс инициализации 160 кратным применением операции $clock_KG(R, S, Mixing = True, Input_bit = 0)$ тактирования всего генератора.

Генерация ключевого потока

После выполнения операции инициализации можно приступить к генерации битов ключевой последовательности $z_0 \dots z_{L-1}$:

For $0 \leq i \leq L-1$:

- $z_i = r_0 \oplus s_0$,
- $clock_KG(R, S, Mixing = False, Input_bit = 0)$,

где $clock_KG(R, S, Mixing, Input_bit)$ и – операция тактирования всего генератора, которая определяется в соответствии со следующим псевдокодом:

- $Control_bit_R = s_{54} \oplus r_{106}$
- $Control_bit_S = s_{106} \oplus r_{53}$
- If $Mixing = True$,
 - $clock_R(R, Input_bit_R = Input_bit \oplus s_{80}, Control_bit_R = Control_bit)$
 - $clock_S(S, Input_bit_S = Input_bit, Control_bit_S = Control_bit)$
- If instead $Mixing = False$,
 - $clock_R(R, Input_bit_R = Input_bit, Control_bit_R = Control_bit)$
 - $clock_S(S, Input_bit_S = Input_bit, Control_bit_S = Control_bit)$

clock_R и clock_S – это операции тактирования регистров R и S соответственно.

Операция тактирования регистра R clock_R(R , Input_bit_R, Control_bit_R) определяется следующим псевдокодом:

- Пусть $r_0 \dots r_{159}$ будет состоянием регистра R до тактирования, а $r'_0 \dots r'_{159}$ будет состоянием регистра R после тактирования.
- Feedback_bit = $r_{159} \oplus$ Input_bit_R
- For $1 \leq i \leq 159$ { $r'_i = r_{i-1}$; $r'_0 = 0$;
- For $0 \leq i \leq 159$ {if $i \in Rtaps$, $r'_i = r'_i \oplus$ Feedback_bit}
- If Control_bit_R = 1
 - For $0 \leq i \leq 159$ { $r'_i = r'_i \oplus r_i$ }

$Rtaps$ – это набор позиций отводов обратной связи для R :

$Rtaps = \{0, 4, 5, 8, 10, 11, 14, 16, 20, 25, 30, 32, 35, 36, 38, 42, 43, 46, 50, 51, 53, 54, 55, 56, 57, 60, 61, 62, 63, 65, 66, 69, 73, 74, 76, 79, 80, 81, 82, 85, 86, 90, 91, 92, 95, 97, 100, 101, 105, 106, 107, 108, 109, 111, 112, 113, 115, 116, 117, 127, 128, 129, 130, 131, 133, 135, 136, 137, 140, 142, 145, 148, 150, 152, 153, 154, 156, 157\}$.

Операция тактирования регистра S clock_S(S , Input_bit_S, Control_bit_S) определяется следующим псевдокодом:

- Пусть $s_0 \dots s_{159}$ будет состоянием регистра S до тактирования, а $s'_0 \dots s'_{159}$ будет состоянием регистра после тактирования. Также, для упрощения описания шифра, мы будем использовать $\hat{s}_0 \dots \hat{s}_{159}$ как промежуточные переменные.
- Feedback_bit = $s_{159} \oplus$ Input_bit_R
- For $1 \leq i \leq 158$ { $\hat{s}_i = s_{i-1} \oplus ((s_i \oplus comp0_i) \cdot (s_{i+1} \oplus comp1_i))$; $\hat{s}_i = 0$; $\hat{s}_{159} = s_{159}$;
- If Control_bit_S = 0:
 - For $0 \leq i \leq 159$ { $s'_i = \hat{s}_i \oplus (fb0_i \cdot$ Feedback_bit)}
- If instead Control_bit_S = 1
 - For $0 \leq i \leq 159$ { $s'_i = \hat{s}_i \oplus (fb1_i \cdot$ Feedback_bit)}

Значения $comp0$, $comp1$, $fb0$ и $fb1$, определенные разработчиками, приведены в [6].

Поточный шифр Trivium

Trivium – синхронный поточный шифр, использующий 80-битовый ключ и 80-битовый вектор инициализации IV [6].

Инициализация

Алгоритм инициализируется загрузкой 80-битового ключа и 80-битового IV в 288-битовое начальное состояние и устанавливает все оставшиеся биты в 0, за исключением s_{286} , s_{287} , и s_{288} . Затем состояние изменяется за 4 полных цикла (без генерации бит ключевого потока). В итоге это можно представить в виде псевдокода, приведенного ниже:

```

( $s_1, s_2, \dots, s_{93}$ )  $\leftarrow$  ( $K_1, \dots, K_{80}, 0, \dots, 0$ )
( $s_{94}, s_{95}, \dots, s_{177}$ )  $\leftarrow$  ( $IV_1, \dots, IV_{80}, 0, \dots, 0$ )
( $s_{178}, s_{279}, \dots, s_{288}$ )  $\leftarrow$  ( $0, \dots, 0, 1, 1, 1$ )
for  $i = 0$  to  $4 \cdot 288$  do
{
   $t_1 \leftarrow s_{66} \oplus s_{91} \otimes s_{92} \oplus s_{93} \oplus s_{171}$ 
   $t_2 \leftarrow s_{162} \oplus s_{175} \otimes s_{176} \oplus s_{177} \oplus s_{264}$ 
   $t_3 \leftarrow s_{243} \oplus s_{286} \otimes s_{287} \oplus s_{288} \oplus s_{69}$ 
  ( $s_1, s_2, \dots, s_{93}$ )  $\leftarrow$  ( $t_3, s_1, \dots, s_{92}$ )
  ( $s_{94}, s_{95}, \dots, s_{177}$ )  $\leftarrow$  ( $t_1, s_{94}, \dots, s_{176}$ )
  ( $s_{178}, s_{279}, \dots, s_{288}$ )  $\leftarrow$  ( $t_2, s_{178}, \dots, s_{287}$ )
}

```

Генерация ключевого потока

Предложенная конструкция содержит 288-битовое внутреннее состояние, обозначенное (s_1, \dots, s_{288}) . Генерация ключевого потока состоит из итерационного процесса, который извлекает значения 15 определенных битов состояния и использует их для обновления 3 битов состояния и вычисления 1 бита ключевого потока z_i . Затем биты состояния циклически сдвигаются и процесс повторяет, пока не будут сгенерированы требуемые N бит ключевого потока. Полное описание дается следующим простым псевдокодом:

```

for  $i = 1$  to  $N$  do
{
   $t_1 \leftarrow s_{66} \oplus s_{93}$ 
   $t_2 \leftarrow s_{162} \oplus s_{177}$ 
   $t_3 \leftarrow s_{243} \oplus s_{288}$ 
   $z_i \leftarrow t_1 \oplus t_2 \oplus t_3$ 
   $t_1 \leftarrow t_1 \oplus s_{91} \otimes s_{92} \oplus s_{171}$ 
   $t_2 \leftarrow t_2 \oplus s_{175} \otimes s_{176} \oplus s_{264}$ 
   $t_3 \leftarrow t_3 \oplus s_{286} \otimes s_{287} \oplus s_{69}$ 
  ( $s_1, s_2, \dots, s_{93}$ )  $\leftarrow$  ( $t_3, s_1, \dots, s_{92}$ )
  ( $s_{94}, s_{95}, \dots, s_{177}$ )  $\leftarrow$  ( $t_1, s_{94}, \dots, s_{176}$ )
  ( $s_{178}, s_{279}, \dots, s_{288}$ )  $\leftarrow$  ( $t_2, s_{178}, \dots, s_{287}$ )
}

```

Графическое представление процесса генерации ключевого потока приведено на рис.

2.30.

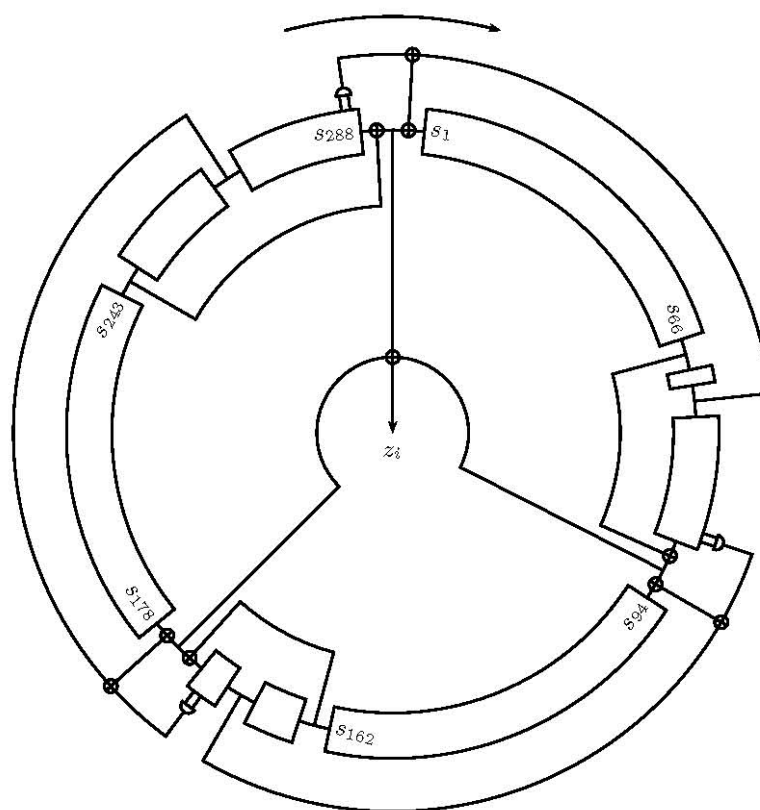


Рис. 7.30. Поточный шифр Trivium

Блочные шифры в поточных режимах

Российский блочный шифр ГОСТ 28147-89 в поточном режиме

В России в качестве стандарта шифрования принят алгоритм криптографического преобразования ГОСТ 28147-89. Этот алгоритм предназначен для аппаратной и программной реализации, удовлетворяет необходимым криптографическим требованиям и не накладывает ограничений на степень секретности защищаемой информации.

Для шифрования используются 256 битовый ключ, который разбивается на 8 32-битовых подключа: K_0, K_1, \dots, K_7 , и таблица блока подстановки H . Заполнение таблиц блока подстановки H является долговременным ключевым элементом.

Алгоритм криптографического преобразования ГОСТ 28147-89 опирается на цикл шифрования (рисунок 7.31). Цикл шифрования построен по принципу сети Фейстеля.

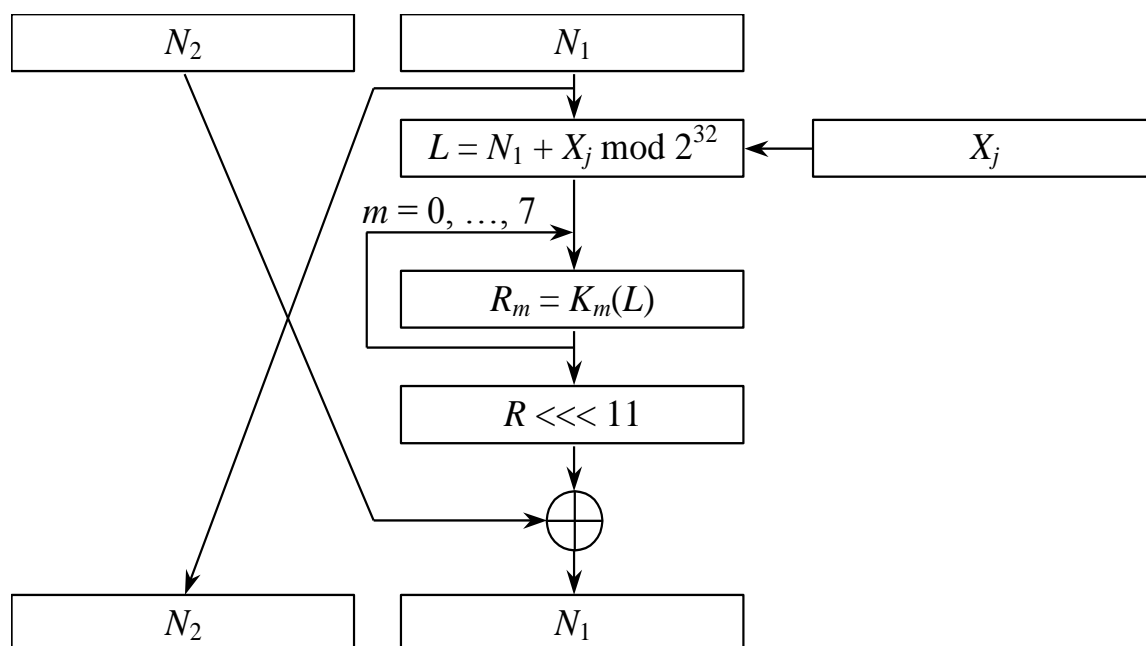


Рис. 7.31. Цикл шифрования

Один цикл шифрования преобразует 64-битовый блок данных, используя один 32-битовый элемент ключа K_j и блок подстановки H .

Преобразуемый 64-битовый блок данных (N) разбивается на две части: старшую (N_2) и младшую (N_1). Значение N_1 суммируется по модулю 2^{32} с одним из 32-битовых элементов ключа K_j , $j = 0, \dots, 7$. Результат суммирования преобразуется в блоке подстановки H . Блок подстановки H состоит из 8 узлов замены H_1, \dots, H_8 размером 64 бита каждый. Поступающий на блок подстановки 32-разрядный блок разбивается на восемь последовательно идущих 4-разрядных блоков, каждый из которых преобразуется в 4-разрядный блок соответствующим узлом замены, представляющим собой таблицу из шестнадцати строк, содержащих по четыре бита заполнения в строке. Входной блок определяет адрес строки в таблице, заполнение данной строки является выходным блоком. Затем 4-разрядные выходные блоки последовательно объединяются в 32-разрядный блок. Полученное значение R циклически сдвигается на 11 шагов в сторону старших разрядов. Результат сдвига суммируется поразрядно по модулю 2 со значением N_2 . Полученный результат N_2 записывается в N_1 , при этом старое значение N_1 переписывается в N_2 .

Многократное повторение цикла шифрования, при использовании различных элементов ключа, позволяет построить циклы зашифрования (32-3) и расшифрования (32-Р). В цикле зашифрования 32-3 отдельный цикл шифрования повторяется 32 раза,

используя 32-битовые элементы ключа в следующем порядке: $K_0, \dots, K_7, K_0, \dots, K_7, K_0, \dots, K_7, K_7, \dots, K_0$. Цикл расшифрования 32-Р отличается от цикла порядком использования 32-битовых элементов ключа. В цикле 32-Р этот порядок следующий: $K_0, \dots, K_7, K_7, \dots, K_0, K_7, \dots, K_0, K_7, \dots, K_0$.

Ниже приведено описание двух поточных режимов работы криптографического алгоритма ГОСТ 28147-89, а именно: гаммирование и гаммирование с обратной связью.

Гаммирование

Криптосхема, реализующая алгоритм зашифрования в режиме гаммирования, имеет вид, указанный на рисунке 7.32а.

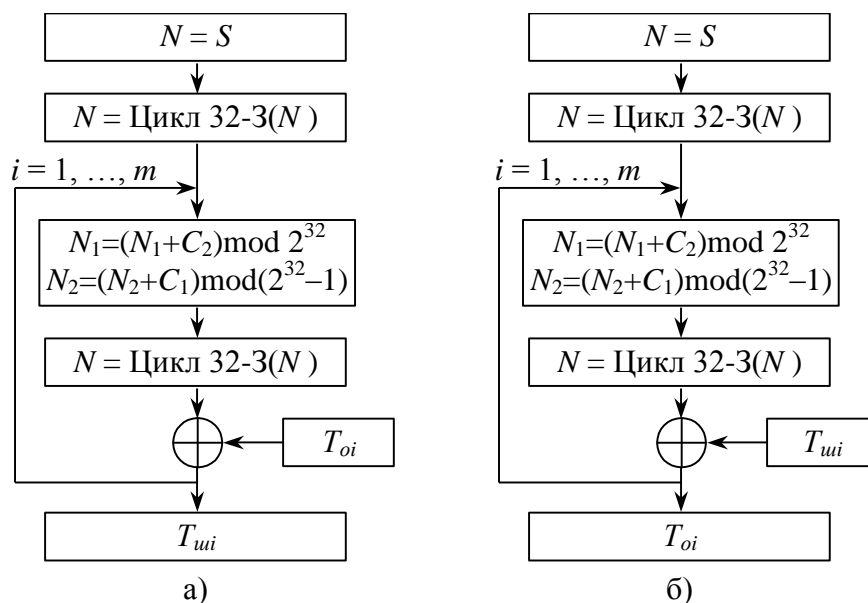


Рис. 7.32. Схема алгоритма зашифрования (а) и расшифрования (б) в режиме гаммирования

Открытые данные, разбитые на 64-разрядные блоки T_{oi} , зашифровываются в режиме гаммирования путем поразрядного суммирования по модулю 2 с гаммой шифра Γ_u , которая вырабатывается блоками по 64 бита Γ_{ui} , $i = 1, \dots, m$. m определяется объемом шифруемых данных. Число двоичных разрядов в блоке T_{om} может быть меньше 64, при этом неиспользованная для зашифрования часть гаммы шифра из блока Γ_{um} отбрасывается.

Для инициализации процесса генерации гаммы используется 64-разрядная двоичная последовательность (синхропосылка) S . К синхропосылке применяется цикл зашифрования 32-3. Результат шифрования N разбивается на две части: старшую (N_2) и младшую (N_1). Значение N_2 суммируется по модулю $(2^{32} - 1)$ с 32-разрядной константой $C_1 = 01010104_{16}$. Значение N_1 суммируется по модулю 2^{32} с 32-разрядной константой $C_2 = 01010101_{16}$. К полученному значению N применяется цикл зашифрования 32-3. Полученное в результате зашифрования значение N образует первый 64-разрядный блок гаммы шифра Γ_{u1} , который

суммируется поразрядно по модулю 2 с первым 64-разрядным блоком открытых данных T_{o1} . В результате суммирования получается 64-разрядный блок зашифрованных данных T_{ui1} . Для получения следующего 64-разрядного блока гаммы шифра Γ_{ui2} старшая часть N_2 значения N суммируется по модулю $(2^{32} - 1)$ с константой C_1 , а младшая часть N_1 суммируется по модулю 2^{32} с константой C_2 . К полученному значению применяется цикл зашифрования 32-3. Полученное в результате зашифрования значение N образует второй 64-разрядный блок гаммы шифра Γ_{ui2} , который суммируется поразрядно по модулю 2 со вторым блоком открытых данных T_{o2} . Аналогично вырабатываются блоки гаммы шифра $\Gamma_{ui3}, \Gamma_{ui4}, \dots, \Gamma_{uim}$ и зашифровываются блоки открытых данных $T_{o3}, T_{o4}, \dots, T_{om}$. Если длина последнего m -го блока открытых данных T_{om} меньше 64 бит, то из последнего m -го блока гаммы шифра Γ_{uim} для зашифрования используется только соответствующее число разрядов гаммы шифра, остальные разряды отбрасываются.

Аналогичным образом производится расшифрование в режиме гаммирования (рисунок 7.32б).

Гаммирование с обратной связью

Криптосхема, реализующая алгоритм зашифрования в режиме гаммирования с обратной связью, имеет вид, указанный на рисунке 3а.

Открытые данные, разбитые на 64-разрядные блоки T_{oi} , зашифровываются в режиме гаммирования с обратной связью путем поразрядного суммирования по модулю 2 с гаммой шифра Γ_{ui} , которая вырабатывается блоками по 64 бита Γ_{ui} , $i = 1, \dots, m$. m определяется объемом шифруемых данных. Число двоичных разрядов в блоке T_{om} может быть меньше 64.

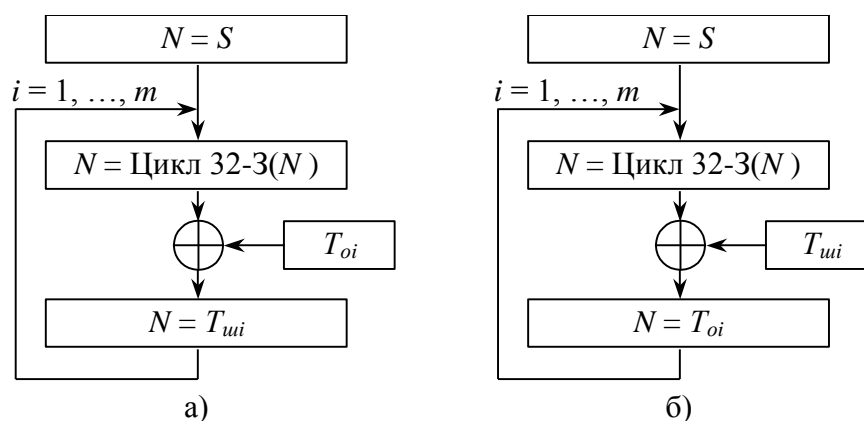


Рис. 7.33. Схема алгоритма зашифрования (а) и расшифрования (б) в режиме гаммирования с обратной связью

Исходное значение – синхропосылка S зашифровывается с помощью цикла зашифрования 32-3. Полученное в результате зашифрования значение N образует первый 64-разрядный блок гаммы шифра Γ_{ui1} , который суммируется поразрядно по модулю 2 с первым

64-разрядным блоком открытых данных T_{o1} . В результате получается 64-разрядный блок зашифрованных данных T_{u1} .

Блок зашифрованных данных T_{u1} одновременно является также исходным состоянием N для выработки второго блока гаммы шифра Γ_{u2} и по обратной связи передается на вход. Новое значение N зашифровывается с помощью цикла зашифрования 32-3. Полученное в результате зашифрования значение N образует второй 64-разрядный блок гаммы шифра Γ_{u2} , который суммируется поразрядно по модулю 2 со вторым блоком открытых данных T_{o2} .

Выработка последующих блоков данных шифра Γ_{ui} и зашифрование соответствующих блоков открытых данных T_{oi} ($i = 3, \dots, m$) производится аналогично. Если длина последнего m -го блока открытых данных T_{om} меньше 64 бит, то из последнего m -го блока гаммы шифра Γ_{um} используется только соответствующее число разрядов гаммы шифра, остальные разряды отбрасываются.

Аналогичным образом производится расшифрование в режиме гаммирования с обратной связью (рисунок 7.33б).

Блочный шифр AES в поточном режиме

В основе стандарта шифрования AES лежит алгоритм Rijndael. Rijndael – это итерационный блочный шифр, имеющий архитектуру “Квадрат”. Шифр имеет переменную длину блоков и различные длины ключей. Длина ключа и длина блока могут быть равны независимо друг от друга 128, 192 или 256 битам. В стандарте AES-128 определена длина блока данных, равная 128 битам.

Операции алгоритма AES выполнены на основе двумерного массива байтов, называемого состоянием (*state*). Состояние состоит из четырех строк, каждая из которых содержит N_b байт, где N_b – длина блока в битах, деленная на 32. В массиве состояния, обозначенном символом s , каждый байт имеет два индекса: номер строки r , $0 \leq r < 4$, и номер столбца c , $0 \leq c < N_b$. Это позволяет обращаться к отдельному байту состояния как s_{rc} . Для AES-128 $N_b = 4$, т.е. $0 \leq c < 4$.

Перед применением операции зашифрования или расшифрования, входной блок данных, представленный в виде массива байтов in_0, \dots, in_{15} , копируется в массив состояния (рисунок 2.34):

$$s_{rc} = in_{r+4c}, \quad 0 \leq r < 4, \quad 0 \leq c < N_b,$$

а в конце операций зашифрования или расшифрования, данные из состояния копируются в выходной массив (рисунок 2.34):

$$out_{r+4c} = s_{rc}, \quad 0 \leq r < 4, \quad 0 \leq c < N_b.$$



Рис. 7.34. Входной и выходной блоки данных, промежуточное состояние

Ключ шифрования также представляется в виде прямоугольного массива с четырьмя строками (рисунок 2). Число столбцов N_k этого массива равно длине ключа в битах, деленной на 32. В стандарте определены ключи всех трех размеров – 128 бит, 192 бита и 256 бит, то есть соответственно 4, 6 и 8 32-разрядных слова (или столбца – в табличной форме представления). В некоторых случаях ключ шифрования рассматривается как линейный массив 4-байтовых слов. Слова состоят из 4 байтов, которые находятся в одном столбце (при представлении в виде прямоугольного массива).

k_{00}	k_{01}	k_{02}	k_{03}
k_{10}	k_{11}	k_{12}	k_{13}
k_{20}	k_{21}	k_{22}	k_{23}
k_{30}	k_{31}	k_{32}	k_{33}

Рис. 7.35. Формат представления ключа шифрования

Для алгоритма AES длина входного и выходного блоков, а также блока состояния – 128 битов, т.е. $N_b = 4$. Длина ключа шифрования равна 128, 192 или 256 битов, т.е. $N_k = 4, 6$ или 8 . Количество раундов, которые будут выполнены в течение выполнения алгоритма, зависит от размера ключа. Число раундов обозначается N_r , где $N_r = 10$ при $N_k = 4$, $N_r = 12$ при $N_k = 6$ и $N_r = 14$ при $N_k = 8$.

И для шифрования, и для расшифрования, алгоритм AES использует раундовую функцию, которая составлена из четырех различных преобразований над байтами: 1) замена байта, используя таблицу замены (S-блок), 2) сдвиг строк массива состояния на различные смещения, 3) смешивание данных в пределах каждого столбца массива состояния, и 4) добавление раундового ключа к состоянию. Эти преобразования (и их инверсии) описаны ниже.

Функция зашифрования

Перед зашифрованием входные данные копируются в массив состояния. После начального добавления раундового ключа, массив состояния преобразуется с помощью раундовой функции 10, 12 или 14 раз (в зависимости от длины ключа). Заключительный раунд немного отличается от первых $N_r - 1$ раундов. Затем полученное состояние копируется в выходной массив.

Раундовая функция состоит из 4 различных преобразований: *SubBytes*, *ShiftRows*, *MixColumns* и *AddRoundKey*. Все N_r раундов идентичны за исключением последнего, который не включает преобразование *MixColumns* (рисунок 7.36).

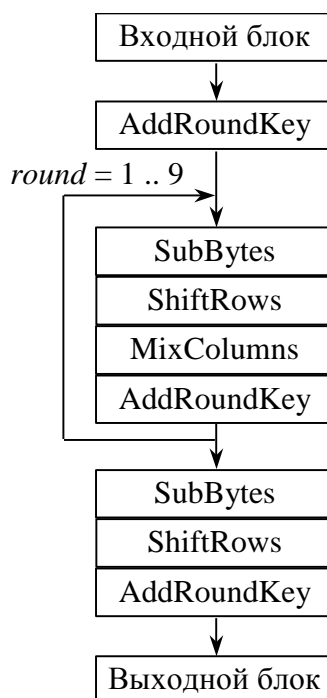


Рис. 7.36. Схема функции зашифрования ($N_b = 4$, $N_k = 4$)

Преобразование *SubBytes* – нелинейная замена байта, которая работает независимо на каждом байте состояния, используя таблицу замены (S-блок). Эта таблица замен создается с использованием двух преобразований:

1. получение обратного элемента относительно умножения в поле $GF(2^8)$, нулевой элемент переходит сам в себя;
2. применение преобразования над $GF(2)$, определенного следующим образом:

$$b'_i = b_i \oplus b_{(i+4)\text{mod}8} \oplus b_{(i+5)\text{mod}8} \oplus b_{(i+6)\text{mod}8} \oplus b_{(i+7)\text{mod}8} \oplus c_i,$$

где $c_0 = c_1 = c_5 = c_6 = 1$, $c_2 = c_3 = c_4 = c_7 = 0$, b_i и b'_i – соответственно исходное и преобразованное значение i -го бита, $i = 0, \dots, 7$.

В преобразовании *ShiftRows* байты в последних трех строках состояния циклически сдвигаются влево на различное число байт. Первая строка ($r = 0$) не сдвигается. Вторая

строка ($r = 1$) сдвигается на 1 байт, третья строка ($r = 2$) – на 2 байта, четвертая строка ($r = 3$) – на 3 байта.

Преобразование *MixColumns* работает с состоянием столбец за столбцом, обрабатывая каждый столбец как 4-элементный полином. Столбцы рассматриваются как полиномы над $GF(2^8)$ и умножаются по модулю $x^4 + 1$ на фиксированный полином $a(x)$, приведенный ниже

$$a(x) = \{03\}x^3 + \{01\}x^2 + \{01\}x + \{02\}.$$

Это может быть представлено в матричном виде:

$$\begin{bmatrix} s'_{0c} \\ s'_{1c} \\ s'_{2c} \\ s'_{3c} \end{bmatrix} = \begin{bmatrix} 02 & 03 & 01 & 01 \\ 01 & 02 & 03 & 01 \\ 01 & 01 & 02 & 03 \\ 03 & 01 & 01 & 02 \end{bmatrix} \begin{bmatrix} s_{0c} \\ s_{1c} \\ s_{2c} \\ s_{3c} \end{bmatrix}, \quad 0 \leq c \leq 3.$$

В результате этого умножения байты столбца заменяются следующими:

$$s'_{0c} = (\{02\} \cdot s_{0c}) \oplus (\{03\} \cdot s_{1c}) \oplus s_{2c} \oplus s_{3c},$$

$$s'_{1c} = s_{0c} \oplus (\{02\} \cdot s_{1c}) \oplus (\{03\} \cdot s_{2c}) \oplus s_{3c},$$

$$s'_{2c} = s_{0c} \oplus s_{1c} \oplus (\{02\} \cdot s_{2c}) \oplus (\{03\} \cdot s_{3c}),$$

$$s'_{3c} = (\{03\} \cdot s_{0c}) \oplus s_{1c} \oplus s_{2c} \oplus (\{02\} \cdot s_{3c}).$$

В преобразовании *AddRoundKey* раундовый ключ добавляется к состоянию простой поразрядной операцией XOR. Каждый раундовый ключ состоит из N_b 4-байтовых слов развернутого ключа. Сложение раундового ключа и состояния производится следующим образом:

$$[s'_{0c}, s'_{1c}, s'_{2c}, s'_{3c}] = [s_{0c}, s_{1c}, s_{2c}, s_{3c}] \oplus [w_{round \cdot N_b + c}], \quad 0 \leq c < N_b,$$

где $[w_i]$ – слова развернутого ключа, описанного ниже, $round$ – номер раунда, $0 \leq round \leq N_r$. При зашифровании, перед первым применением раундовой функции, происходит начальное добавление раундового ключа ($round = 0$). Применение преобразования *AddRoundKey* в N_r раундах зашифрования происходит при $1 \leq round \leq N_r$.

Расширение ключа

В алгоритме AES для получения раундовых ключей используется алгоритм расширения ключа. Расширенный ключ представляет собой линейный массив $w[i]$ из $N_b(N_r + 1)$ 4-байтовых слов, $i = 0, \dots, N_b(N_r + 1)$. Алгоритм выработки ключей зависит от величины N_k . Первые N_k слов расширенного ключа заполняются ключом шифрования. Каждое последующее слово $w[i]$ получается посредством XOR предыдущего слова $w[i-1]$ и слова на N_k позиций ранее $w[i - N_k]$

$$w[i] = w[i - 1] \oplus w[i - N_k].$$

Для слов, позиция которых кратна N_k , перед применением операции XOR слово $w[i-1]$ подвергается воздействию двух дополнительных функций: *RotWord*, осуществляющей побайтовый сдвиг 32-разрядного слова по формуле $\{a_0 a_1 a_2 a_3\} \rightarrow \{a_1 a_2 a_3 a_0\}$, и *SubWord*, осуществляющей побайтовую замену с использованием S-блока функции *SubBytes*. Затем к полученному значению прибавляется раундовая константа $Rcon[j] = 2^{j-1}$. В итоге значение $w[i]$ равно

$$w[i] = SubWord(RotWord(w[i-1])) \oplus Rcon[i/N_k] \oplus w[i - N_k].$$

i -ый раундовый ключ получается из слов массива раундового ключа от $w[N_b i]$ и до $w[N_b(i+1)]$.

Функция расшифрования

Если преобразования, используемые в функции зашифрования, инвертировать и затем применить в обратном порядке, то можно произвести обратное расшифрование. При расшифровании используются следующие преобразования: *InvShiftRows*, *InvSubBytes*, *InvMixColumns* и *AddRoundKey* (рисунок 4а).

InvShiftRows – инверсия преобразования *ShiftRows*. Байты в последних трех строках состояния циклически сдвигаются вправо на различное число байт. Первая строка ($r = 0$) не сдвигается. Вторая строка ($r = 1$) сдвигается на 1 байт, третья строка ($r = 2$) – на 2 байта, четвертая строка ($r = 3$) – на 3 байта.

InvSubBytes – инверсия преобразования замены байта, в котором к каждому байту состояния применяется обратный S-блок.

InvMixColumns – инверсия преобразования *MixColumns*. *InvMixColumns* оперирует состоянием столбец за столбцом, обрабатывая каждый столбец как 4-элементный полином, как описано в ранее. Столбцы рассматривают как полиномы над $GF(2^8)$ и умножаются по модулю $x^4 + 1$ на фиксированный полином $a^{-1}(x)$

$$a^{-1}(x) = \{0b\}x^3 + \{0d\}x^2 + \{09\}x + \{0e\}.$$

Это может быть представлено в матричном виде:

$$\begin{bmatrix} s'_{0c} \\ s'_{1c} \\ s'_{2c} \\ s'_{3c} \end{bmatrix} = \begin{bmatrix} 0e & 0b & 0d & 09 \\ 09 & 0e & 0b & 0d \\ 0d & 09 & 0e & 0b \\ 0b & 0d & 09 & 0e \end{bmatrix} \begin{bmatrix} s_{0c} \\ s_{1c} \\ s_{2c} \\ s_{3c} \end{bmatrix}, \quad 0 \leq c \leq 3.$$

В результате на выходе получаются следующие байты:

$$s'_{0c} = (\{0e\} \cdot s_{0c}) \oplus (\{0b\} \cdot s_{1c}) \oplus (\{0d\} \cdot s_{2c}) \oplus (\{09\} \cdot s_{3c}),$$

$$s'_{1c} = (\{09\} \cdot s_{0c}) \oplus (\{0e\} \cdot s_{1c}) \oplus (\{0b\} \cdot s_{2c}) \oplus (\{0d\} \cdot s_{3c}),$$

$$s'_{2c} = (\{0d\} \cdot s_{0c}) \oplus (\{09\} \cdot s_{1c}) \oplus (\{0e\} \cdot s_{2c}) \oplus (\{0b\} \cdot s_{3c}),$$

$$s'_{3c} = (\{0b\} \cdot s_{0c}) \oplus (\{0d\} \cdot s_{1c}) \oplus (\{09\} \cdot s_{2c}) \oplus (\{0e\} \cdot s_{3c}).$$

Преобразование `AddRoundKey` является собственной инверсией, т.к. использует только операцию XOR.

В функции обратного расшифрования (рисунок 4а) последовательность преобразований отличается от последовательности преобразований при зашифровании, в то время как форма расширенных ключей для зашифрования и расшифрования остается той же самой. Однако, несколько свойств алгоритма AES позволяют функции расшифрования иметь ту же самую последовательность преобразований как и при зашифровании (с преобразованиями, замененными их инверсиями) (рисунок 2.37б). Это достигается при внесении изменений в расширенный ключ.

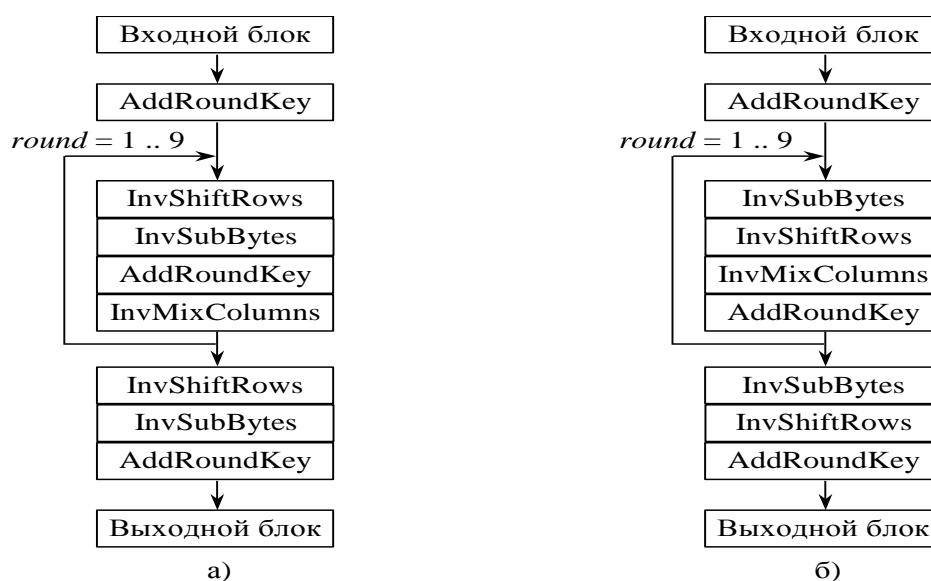


Рис. 7.37. Схемы функций обратного (а) и прямого (б) расшифрования

Существует два свойства, которые делают возможным использование функции прямого расшифрования:

1. Преобразования *SubBytes* и *ShiftRows* можно переставлять местами, т.е. можно сначала применить преобразование *SubBytes*, а затем *ShiftRows*, и наоборот – сначала *ShiftRows*, а затем *SubBytes*. То же самое верно и для их инверсий *InvSubBytes* и *InvShiftRows*.

2. Операции смешивания столбцов *MixColumns* и *InvMixColumns* являются линейными относительно входных данных, т.е.

$$\begin{aligned} \text{InvMixColumns}(\text{State XOR Round Key}) &= \\ &= \text{InvMixColumns}(\text{State}) \text{ XOR } \text{InvMixColumns}(\text{Round Key}). \end{aligned}$$

Эти свойства позволяют изменить порядок преобразований *InvSubBytes* и *InvShiftRows*. Порядок преобразований *AddRoundKey* и *InvMixColumns* также может быть полностью

изменен, при условии, что предварительно, используя преобразование *InvMixColumns*, будут изменены столбцы (слова) развернутого ключа расшифрования.

Режим обратной связи по шифртексту (CFB)

В режиме обратной связи по шифртексту (CFB) выходные блоки зашифрованного текста за счет обратной связи становятся входными блоками для последующего шифрования. Чтобы получить зашифрованный текст, открытый текст складывается, с помощью операции XOR, со сгенерированными выходными блоками. В качестве начального входного блока в режиме CFB используется значение вектора инициализации IV. Значение IV может быть открытым, но при этом должно быть непредсказуемым.

В режиме CFB используется целочисленный параметр s , $1 \leq s \leq b$, где b – размер блока (в байтах) открытого и зашифрованного текстов. Значение s иногда включается в название режима, например, 1-битовый режим CFB, 8-битовый режим CFB, 64-битовый режим CFB, или 128-битовый режим CFB.

При зашифровании в режиме CFB первым входным блоком является значение IV (рисунок 2.38). Чтобы произвести первый входной блок к значению IV применяется операция шифрования. Первая доля зашифрованного текста производится операцией XOR между первой долей открытого текста и s старшими битами первого выходного блока. Оставшиеся $(b-s)$ бит первого выходного блока отбрасываются. Затем $(b-s)$ младших бита значения IV соединяются операцией конкатенации с s битами первой доли зашифрованного текста, для формирования второго входного блока. Другими словами для формирования второго входного блока биты первого входного блока циклически сдвигаются на s позиции влево, и затем доля зашифрованного текста заменяет s младших бит.

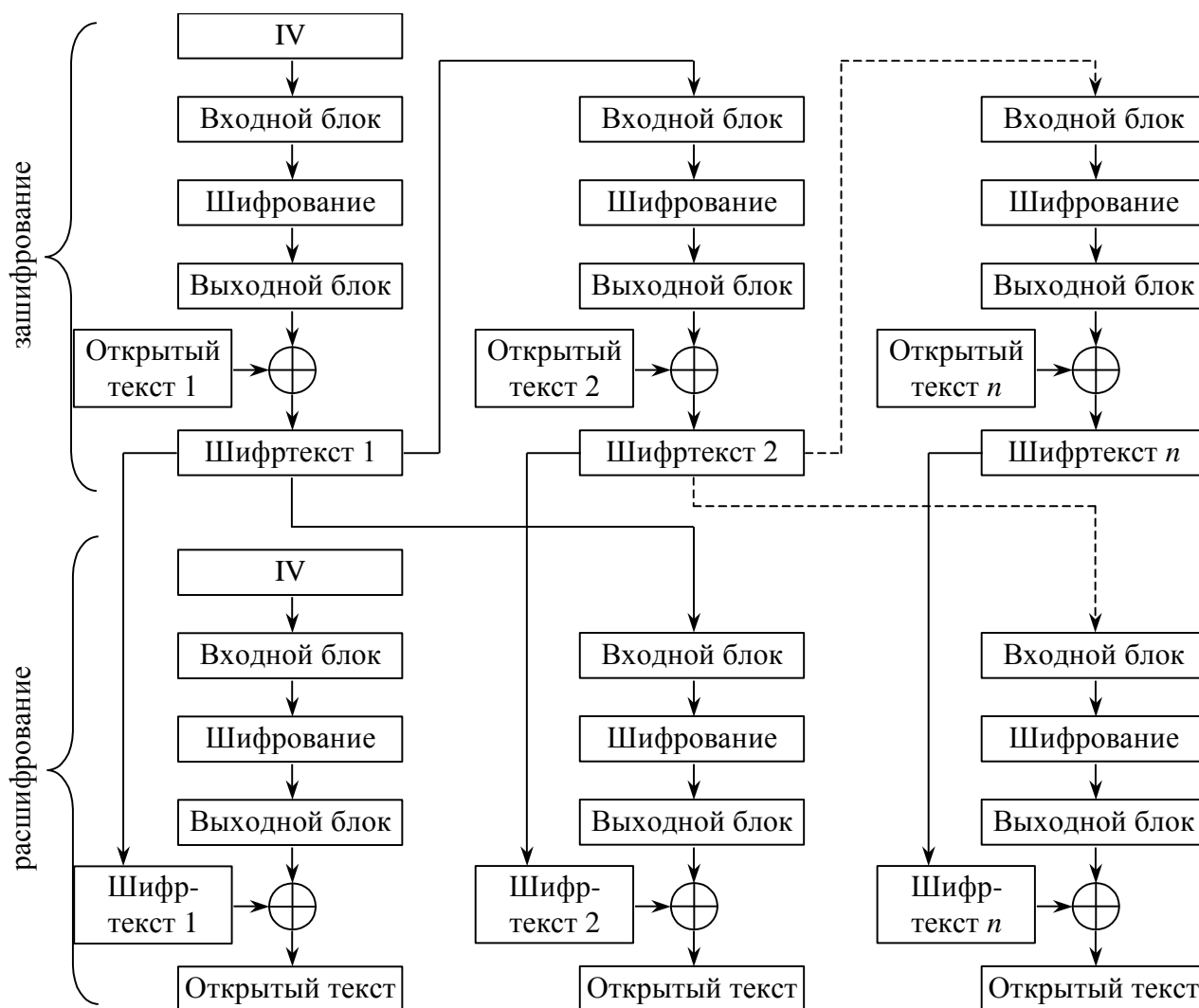


Рис. 7.38. Режим CFB

Процесс повторяется до тех пор, пока от каждой доли открытого текста не будет получена доля зашифрованного текста. Чтобы сформировать долю зашифрованного текста s старших бит каждого выходного блока складываются, с помощью операции XOR, с соответствующей долей открытого текста. Каждая доля зашифрованного текста (кроме последней) подается назад, как описано выше, чтобы сформировать новый входной блок. Обратная связь может быть описана в терминах отдельных бит следующим образом: если $i_1 i_2 \dots i_b - j$ -й входной блок, а $c_1 c_2 \dots c_s - j$ -ая доля зашифрованного текста, то $(j+1)$ -й входной блок: $i_1 i_2 \dots i_b c_1 c_2 \dots c_s$.

При расшифровании в режиме CFB значение IV является первым входным блоком. Каждый последующий входной блок формируется как при CFB зашифровании, соединяя операцией конкатенации $(b-s)$ младших бит предыдущего входного блока с s старшими битами предыдущего зашифрованного текста. Чтобы произвести выходные блоки к каждому входному блоку применяется функция шифрования. Чтобы получить доли открытого текста

с старших бит выходного блока складывается, с помощью операции XOR, с соответствующими долями зашифрованного текста.

При зашифровании в режиме CFB входной блок для каждой функции шифрования (кроме первого) зависит от результата предыдущей функции шифрования, поэтому многократные операции шифрования не могут быть выполнены параллельно. При расшифровании в режиме CFB, необходимые операции шифрования могут быть выполнены параллельно, если из значения IV и зашифрованного текста сначала создать (последовательно) входные блоки.

Режим обратной связи по выходу (OFB)

В режиме обратной связи по выходу (OFB) последовательность выходных блоков генерируется путем последовательного применения операции шифрования к значению IV. Чтобы получить зашифрованный текст открытый текст складывается с помощью операции XOR со сгенерированными выходными блоками и наоборот. Для данного режима шифрования значения IV должны быть случайными, т.е. IV должны быть уникальными для каждого применения режима OFB на данном ключе.

Для зашифрования в режиме OFB, чтобы произвести первый выходной блок, значение IV преобразовывается функцией шифрования. Чтобы произвести первый блок зашифрованного текста первый выходной блок складывается, с помощью операции XOR, с первым блоком открытого текста. Затем, чтобы произвести второй выходной блок, функция шифрования применяется к первому выходному блоку. Второй выходной блок складывается с помощью операции XOR со вторым блоком открытого текста, чтобы произвести второй блок зашифрованного текста. Затем ко второму выходному блоку применяется функция шифрования, чтобы произвести третий выходной блок. Таким образом, последовательные выходные блоки производятся путем применения функции шифрования к предыдущим выходным блокам, а затем выходные блоки складывается с помощью операции XOR с соответствующими блоками открытого текста, чтобы произвести блоки зашифрованного текста. В последнем блоке (который может быть неполным блоком из u бит) для операции XOR используются старшие u бит последнего выходного блока. Оставшиеся $(b-u)$ бит последнего выходного блока отбрасываются.

При расшифровании в режиме OFB чтобы произвести первый выходной блок значение IV преобразовывается функцией шифрования (рисунок 2.39). Чтобы получить первый блок открытого текста первый выходной блок складывается, с помощью операции XOR, с первым блоком зашифрованного текста. Затем первый выходной блок преобразовывается функцией шифрования, чтобы произвести второй выходной блок. Второй выходной блок складывается с помощью операции XOR со вторым блоком зашифрованного текста, чтобы произвести

второй блок открытого текста. Второй выходной блок также преобразовывается функцией шифрования, чтобы произвести третий выходной блок. Таким образом, последовательные выходные блоки производятся путем применения функции шифрования к предыдущим выходным блокам, затем выходные блоки складываются, с помощью операции XOR, с соответствующими блоками зашифрованного текста, чтобы получить блоки открытого текста. В последнем блоке (который может быть неполным блоком из u бит) для операции XOR используются старшие u бит последнего выходного блока. Оставшиеся $(b-u)$ бит последнего выходного блока отбрасываются.

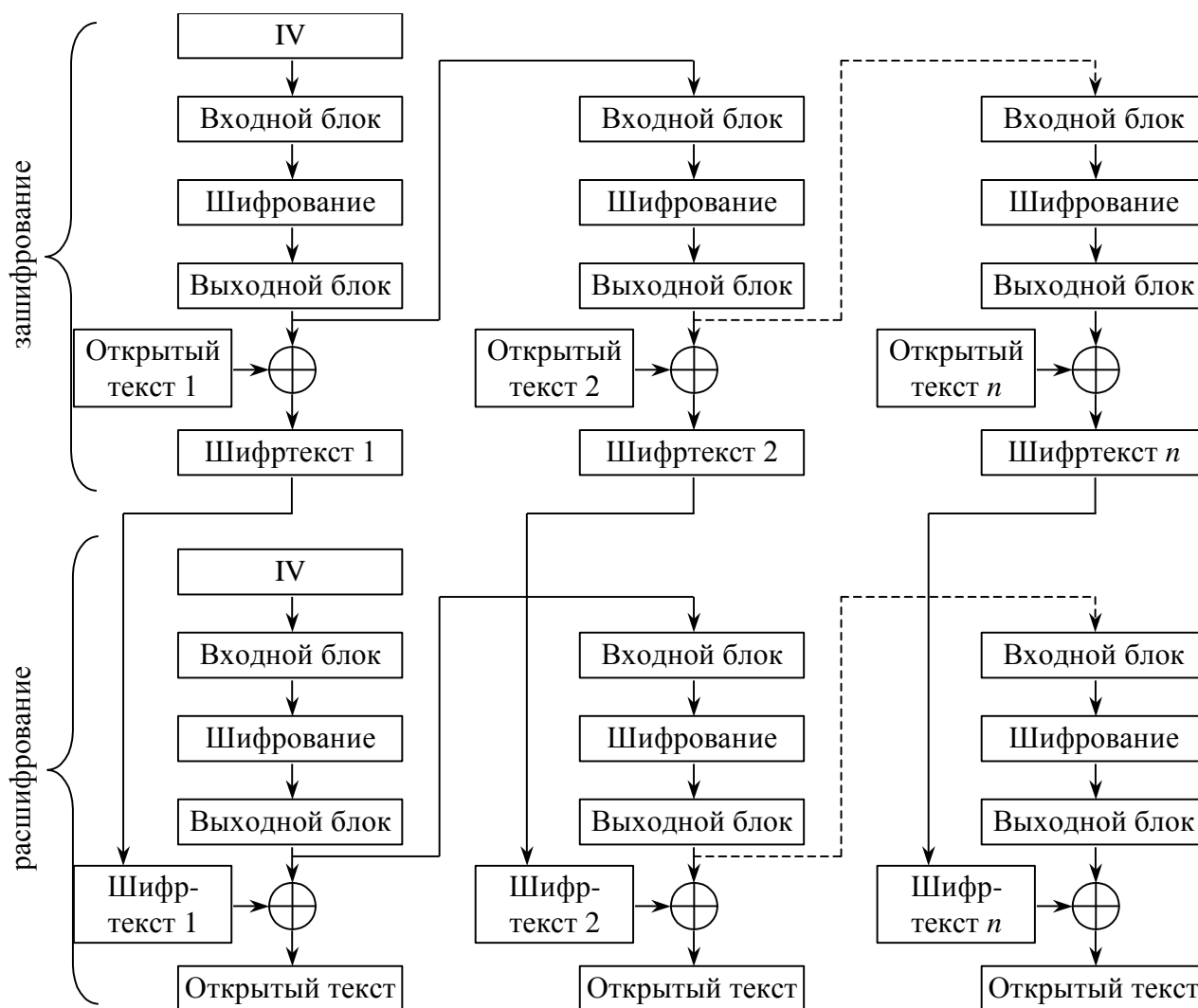


Рис. 7.39. Режим OFB

В режиме OFB и при зашифровании и при расшифровании каждая функция шифрования (кроме первой) зависит от результатов предыдущей функции шифрования, поэтому многократные функции шифрования не могут быть выполнены параллельно. Однако, если значение IV известно, выходные блоки могут быть сгенерированы прежде, чем станет доступным зашифрованный или открытый текст.

Для режима OFB требует уникальное значение IV для каждого сообщения, которое когда-либо будет зашифровано на данном ключе. Если, вопреки этому требованию, используется то же самое значение IV для зашифровывания более, чем одного сообщения, то конфиденциальность этих сообщений может быть поставлена под угрозу. В частности если известен блок открытого текста любого из этих сообщений, скажем, j -й блок открытого текста, то выход j -й функции шифрования может быть определен легко из j -го блока зашифрованного текста сообщения. Эта информация позволяет легко восстановить j -й блок открытого текста любого другого сообщения, зашифрованного с использованием того же самого значения IV , из соответствующего j -го блока зашифрованного текста этого сообщения.

Конфиденциальность может быть аналогичным образом поставлена под угрозу, если любой из входных блоков функции шифрования определяется как IV для зашифрования другого сообщения на данном ключе.

Режим счетчика (Counter mode)

В режиме счетчика (CTR) для производства последовательности выходных блоков операция шифрования применяется к набору выходных блоков, называемых счетчиками. Чтобы получить зашифрованный текст открытый текст складывается с помощью операции XOR со сгенерированными выходными блоками и наоборот. Все блоки в последовательности счетчика должны отличаться друг от друга. Это условие не ограничивается одним сообщением, т.е. для всех сообщений, которые зашифровываются на данном ключе, все счетчики должны быть различными.

При зашифровании в режиме CRT функция шифрования применяется к каждому блоку счетчика (рисунок 2.40). Полученные выходные блоки складываются с помощью операции XOR с соответствующими блоками открытого текста, для того чтобы произвести блоки зашифрованного текста. В последнем блоке (который может быть неполным блоком из u битов) для операции XOR используются старшие u бит последнего выходного блока. Оставшиеся $(b-u)$ бит последнего выходного блока отбрасываются.

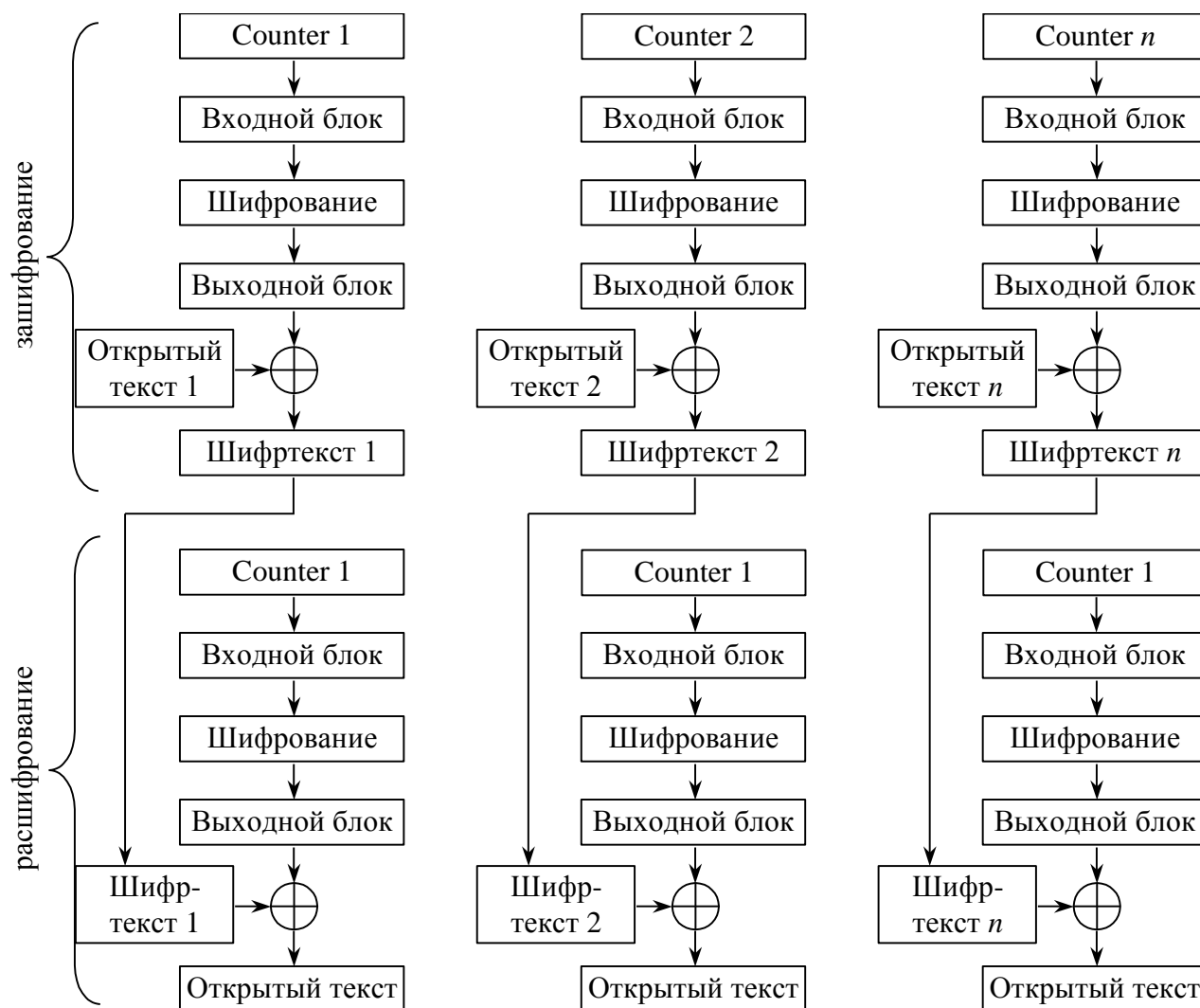


Рис. 7.40. Режим CRT

При расшифровании в режиме CTR выполняются те же самые операции, что и при зашифровании.

В режиме CTR и при зашифровании и при расшифровании функции шифрования могут быть выполнены параллельно. Блок открытого текста, который соответствует какому-то отдельному блоку зашифрованного текста, может быть получен независимо от других блоков открытого текста, если может быть определен соответствующий блок счетчика. Кроме того, функции шифрования могут быть применены к счетчикам прежде, чем станет доступным зашифрованный или открытый текст.

7.2. Компьютерный практикум по изучению шифрования с секретным ключом

7.2.1 Изучение стандарта криптографической защиты ГОСТ 28147-89

Цель работы Изучить криптографический стандарт шифрования ГОСТ 28147-89 и его особенности, познакомиться с различными режимами блочного шифрования.

ГОСТ 28147-89 — советский и российский стандарт симметричного шифрования, введённый в 1990 году, также является стандартом СНГ. Полное название — «ГОСТ 28147-89 Системы обработки информации. Защита криптографическая. Алгоритм криптографического преобразования». С момента опубликования ГОСТа на нём стоял ограничительный гриф «Для служебного пользования», и формально шифр был объявлен «полностью открытым» только в мае 1994 года. К сожалению, история создания шифра и критерии разработчиков до сих пор не опубликованы.

Алгоритм криптографического преобразования предназначен для аппаратной или программной реализации, удовлетворяет криптографическим требованиям и по своим возможностям не накладывает ограничений на степень секретности защищаемой информации.

Стандарт обязателен для организаций, предприятий и учреждений, применяющих криптографическую защиту данных, хранимых и передаваемых в сетях, в отдельных вычислительных комплексах или на персональных компьютерах.

То, что в его названии вместо термина «шифрование» фигурирует более общее понятие «криптографическое преобразование», вовсе не случайно. Помимо нескольких тесно связанных между собой процедур шифрования, в документе описан один построенный на общих принципах с ними алгоритм выработки имитовставки. Последняя является не чем иным, как криптографической контрольной комбинацией, то есть кодом, вырабатываемым из исходных данных с использованием секретного ключа с целью имитозащиты, или защиты данных от внесения в них несанкционированных изменений.

Математические операции

Сложение по модулю 2

Операция поразрядного XOR (обозначается как \oplus) — булева функция и логическая операция. Результат выполнения операции является истинным только при условии, если является истинным в точности один из аргументов. Пример выполнения операции сложения:

$$(x^6 + x^4 + x^2 + x + 1) \oplus (x^7 + x + 1) = x^7 + x^6 + x^4 + x^2 \text{ (в виде многочленов)} \quad (10)$$

$$\{01010111\} \oplus \{10000011\} = \{11010100\} \text{ (двоичное представление)} \quad (11)$$

Правила суммирования по модулю 2^{32} и по модулю $(2^{32}-1)$

1. Два целых числа a, b , где $0 \leq a, b \leq 2^{32}-1$, представленные в двоичном виде

$$a = (a_{32}, a_{31}, \dots, a_2, a_1), \quad b = (b_{32}, b_{31}, \dots, b_2, b_1), \quad (12)$$

$$\text{т.е. } a = a_{32} \cdot 2^{31} + a_{31} \cdot 2^{30} + \dots + a_2 \cdot 2 + a_1, \quad (13)$$

$$b = b_{32} \cdot 2^{31} + b_{31} \cdot 2^{30} + \dots + b_2 \cdot 2 + b_1, \quad (14)$$

суммируются по модулю 2^{32} (операция \boxplus) по следующему правилу:

$$a \boxplus b = a + b, \text{ если } a + b < 2^{32}, \quad (15)$$

$$a \boxplus b = a + b - 2^{32} \text{ если } a + b \geq 2^{32}, \quad (16)$$

где операция $+(-)$ есть арифметическая сумма (разность) двух целых чисел.

2. Два целых числа a, b , где $0 \leq a, b \leq 2^{32}-1$, представленные в двоичном виде

$$a = (a_{32}, a_{31}, \dots, a_2, a_1), \quad (17)$$

$$b = (b_{32}, b_{31}, \dots, b_2, b_1), \quad (18)$$

суммируются по модулю $(2^{32}-1)$ (операция \boxplus') по следующему правилу:

$$a \boxplus' b = a + b, \text{ если } a + b < 2^{32}, \quad (19)$$

$$a \boxplus' b = a + b - 2^{32} + 1, \text{ если } a + b \geq 2^{32}. \quad (20)$$

Структура алгоритма

Симметричное шифрование — способ шифрования, в котором для зашифровки и расшифровывания применяется один и тот же криптографический ключ. До изобретения схемы асимметричного шифрования единственным существовавшим способом являлось симметричное шифрование. Ключ алгоритма должен сохраняться в секрете обеими сторонами. Алгоритм шифрования выбирается сторонами до начала обмена сообщениями. К симметричному шифрованию предъявляются следующие требования:

- Отсутствие линейности (то есть условия $f(a) \text{ xor } f(b) == f(a \text{ xor } b)$), в противном случае облегчается применение дифференциального криптоанализа к шифру. (Функция *xor* – «сложение по модулю 2», «исключающее или» – результат выполнения операции является истинным только при условии, если является истинным в точности один из аргументов).

- Полная утрата всех статистических закономерностей исходного сообщения.

Алгоритм ГОСТ 28147-89 является *блочным шифром* – разновидность симметричного шифра. Особенностью блочного шифра является обработка блока нескольких байт за одну

итерацию (в нашем случае 8). Как и большинство современных блочных шифров, ГОСТ основан на *сети Фейстеля* (рисунок 7.1). Сеть представляет собой определённую 7многократно повторяющуюся (итерированную) структуру, называемую ячейкой Фейстеля. При переходе от одной ячейки к другой меняется ключ, причём выбор ключа зависит от конкретного алгоритма. Операции шифрования и расшифрования на каждом этапе очень просты, и при определённой доработке совпадают, требуя только обратного порядка используемых ключей. Шифрование при помощи данной конструкции легко реализуется как на программном уровне, так и на аппаратном, что обеспечивает широкие возможности применения.

Если внимательно изучить оригинал ГОСТ 28147–89, можно заметить, что в нем содержится описание алгоритмов нескольких уровней. На самом верхнем находятся практические алгоритмы, предназначенные для шифрования массивов данных и выработки для них имитовставки. Все они опираются на три алгоритма низшего уровня, называемые в тексте ГОСТа циклами. Эти фундаментальные алгоритмы будут называться «базовые циклы», чтобы отличать их от всех прочих циклов. Они имеют следующие названия и обозначения: (последние приведены в скобках)

- цикл зашифрования (32-3);
- цикл расшифрования (32-Р);
- цикл выработки имитовставки (16-3).

В свою очередь, каждый из базовых циклов представляет собой многократное повторение одной единственной процедуры, называемой «основным шагом криптопреобразования».

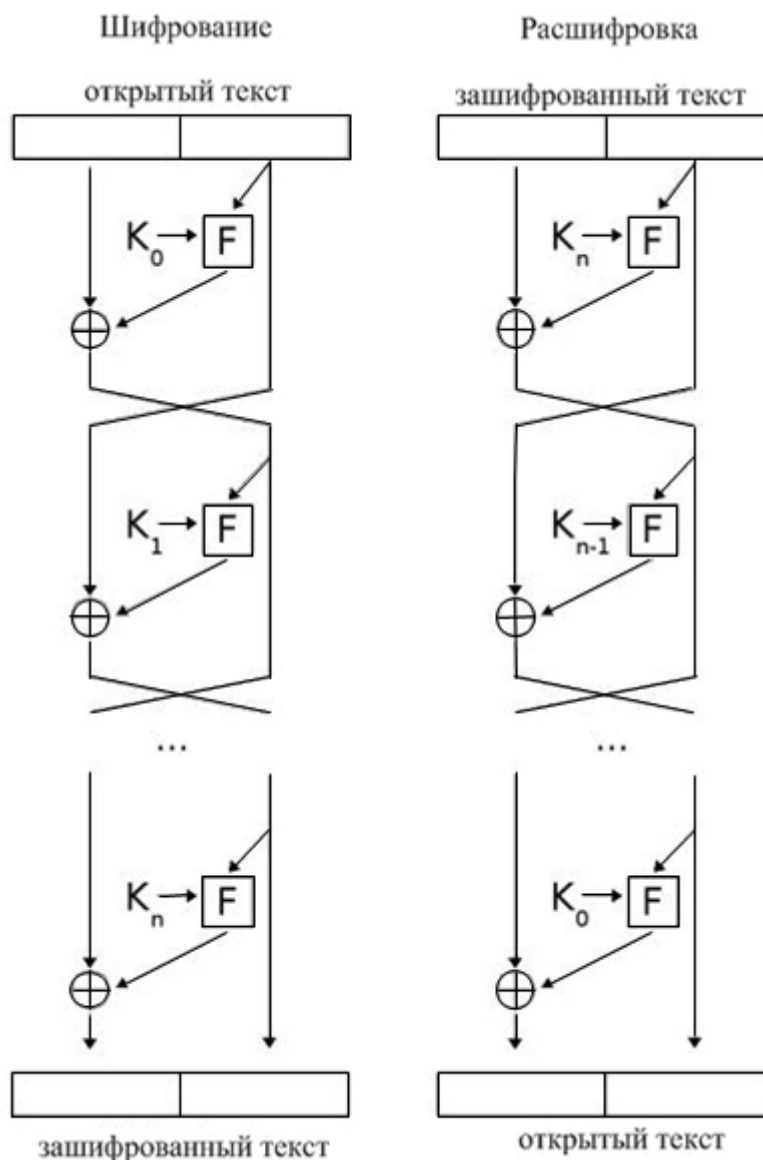


Рис. 7.41. Сеть Фейстеля

- *Ключ* является массивом из восьми 32-битовых элементов кода В ГОСТе элементы ключа используются как 32-разрядные целые числа без знака. Таким образом, размер ключа составляет 256 бит (32 байта). Ключ должен являться массивом статистически независимых битов, принимающих с равной вероятностью значения 0 и 1. При этом некоторые конкретные значения ключа могут оказаться «слабыми», то есть шифр может не обеспечивать заданный уровень стойкости в случае их использования. Однако, предположительно, доля таких значений в общей массе всех возможных ключей ничтожно мала. Поэтому ключи, выработанные с помощью некоторого датчика истинно случайных чисел, будут качественными с вероятностью, отличающейся от единицы на ничтожно малую величину.

- *Таблица замен* является вектором, содержащим восемь узлов замены. Каждый узел замены, в свою очередь, является вектором, содержащим шестнадцать 4-битовых элементов замены, которые можно представить в виде целых чисел от 0 до 15, все элементы одного узла замены обязаны быть различными. Таким образом, таблица замен может быть представлена в виде матрицы размера 8x16 или 16x8, содержащей 4-битовые заменяющ Таким образом, общий объем таблицы замен равен 512 бит (64 байта).

Общее количество узлов замены (S-блоков) ГОСТа — восемь. Каждый S-блок представляет собой перестановку чисел от 0 до 15. Первая 4-битная подпоследовательность попадает на вход первого S-блока, вторая — на вход второго и т. д.

Если S-блок выглядит так: 1, 15, 13, 0, 5, 7, 10, 4, 9, 2, 3, 14, 6, 11, 8, 12

и на входе S-блока 0, то на выходе будет 1, если 4, то на выходе будет 5, если на входе 12, то на выходе 6 и т. д. (для таблицы 2.24)

Таблица 7.24. S-блоки, приведенные в ГОСТ Р 34.11-94 для целей тестирования

Номер S-блока	Значение															
	4	10	9	2	13	8	0	14	6	11	1	12	7	15	5	3
1	4	10	9	2	13	8	0	14	6	11	1	12	7	15	5	3
2	14	11	4	12	6	13	15	10	2	3	8	1	0	7	5	9
3	5	8	1	13	10	3	4	2	14	15	12	7	6	0	9	11
4	7	13	10	1	0	8	9	15	14	4	6	12	11	2	5	3
5	6	12	7	1	5	15	13	8	4	10	9	14	0	3	11	2
6	4	11	10	0	7	2	1	13	3	6	8	5	9	12	15	14
7	13	11	4	1	3	15	5	9	0	10	14	7	6	8	2	12
8	1	15	13	0	5	7	10	4	9	2	3	14	6	11	8	12

Основной шаг криптопреобразования

Основной шаг криптопреобразования по своей сути является оператором, определяющим преобразование 64-битового блока данных. Дополнительным параметром этого оператора является 32-битовый блок, в качестве которого используется какой-либо элемент ключа. Схема алгоритма основного шага приведена на рисунке 2.2. Рассмотрим подробнее этапы основного шага криптопреобразования:

Шаг 0 Определяет исходные данные для основного шага криптопреобразования.

N – преобразуемый 64-битовый блок данных, в ходе выполнения шага его младшая (N_1) и старшая (N_2) части обрабатываются как отдельные 32-битовые целые числа без знака. Таким образом, можно записать $N=(N_1, N_2)$.

X – 32-битовый элемент ключа;

Шаг 1 Сложение с ключом.

Младшая половина преобразуемого блока складывается по модулю 2^{32} с используемым на шаге элементом ключа, результат передается на следующий шаг;

Шаг 2 Поблочная замена.

32-битовое значение, полученное на предыдущем шаге, интерпретируется как массив из восьми 4-битовых блоков кода:

$$S=(S_0, S_1, S_2, S_3, S_4, S_5, S_6, S_7) \quad (21)$$

причем S_0 содержит 4 самых младших, а S_7 – 4 самых старших бита S . Далее значение каждого из восьми блоков заменяется новым, которое выбирается по таблице замен следующим образом: значение блока S_i меняется на S_i -тый по порядку элемент (нумерация с нуля) i -того узла замены (т.е. i -той строки таблицы замен, нумерация также с нуля). Другими словами, в качестве замены для значения блока выбирается элемент из таблицы замен с номером строки, равным номеру заменяемого блока, и номером столбца, равным значению заменяемого блока как 4-битового целого неотрицательного числа. Отсюда становится понятным размер таблицы замен: число строк в ней равно числу 4-битовых элементов в 32-битовом блоке данных, то есть восьми, а число столбцов равно числу различных значений 4-битового блока данных, равному, как известно 2^4 (т. е. шестнадцати).

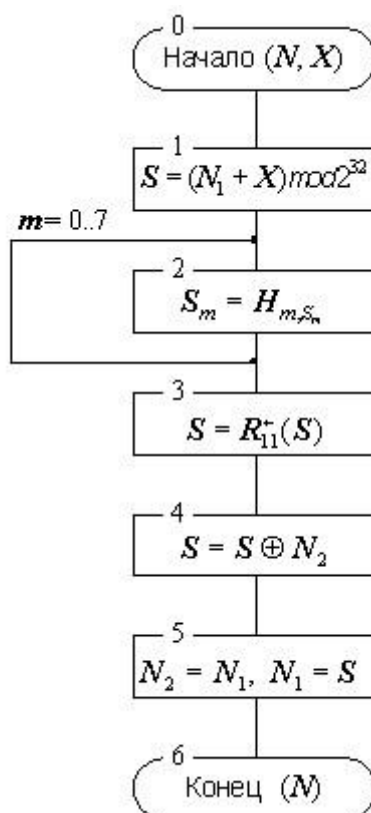


Рис. 7.42. Схема основного шага криптопреобразования алгоритма ГОСТ 28147-89

Шаг 3 Циклический сдвиг на 11 бит влево.

Результат предыдущего шага сдвигается циклически на 11 бит в сторону старших разрядов и передается на следующий шаг. На схеме алгоритма символом « R_{11}^{\leftarrow} » обозначена функция циклического сдвига своего аргумента на 11 бит влево, т.е. в сторону старших разрядов.

Шаг 4 Побитовое сложение

Значение, полученное на шаге 3, побитно складывается по модулю 2 со старшей половиной преобразуемого блока.

Шаг 5 Сдвиг по цепочке

Младшая часть преобразуемого блока сдвигается на место старшей, а на ее место помещается результат выполнения предыдущего шага.

Шаг 6 Завершение криптопреобразования.

Полученное значение преобразуемого блока возвращается как результат выполнения алгоритма основного шага криптопреобразования.

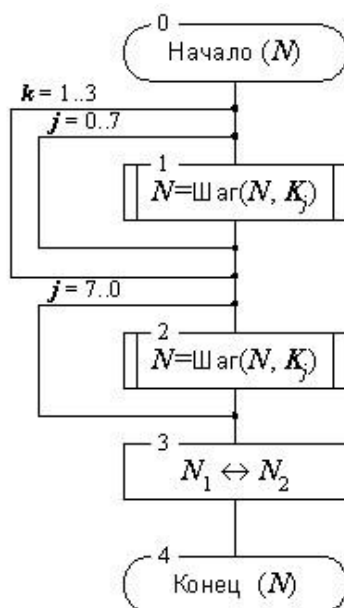


Рис. 7.43. Схема цикла зашифрования 32-3

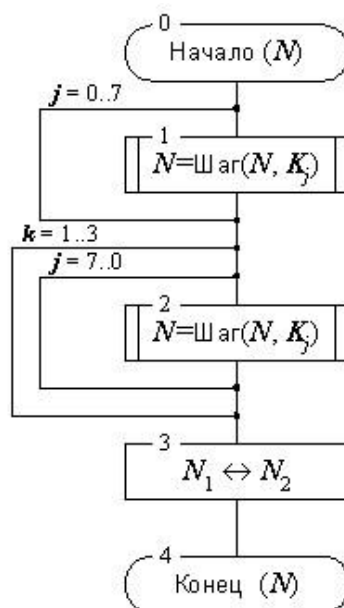


Рис. 7.44. Схема цикла расшифрования 32-Р

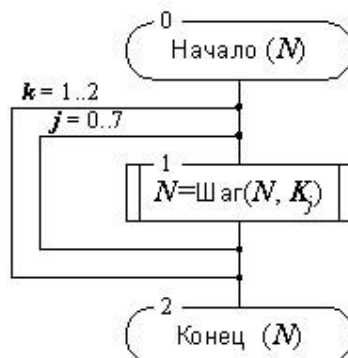


Рис. 7.45. Схема цикла выработки имитовставки 16-3.

Порядок использования ключевых элементов:

- Цикл зашифрования 32-3:
 $K_0, K_1, K_2, K_3, K_4, K_5, K_6, K_7, K_0, K_1, K_2, K_3, K_4, K_5, K_6, K_7, K_0, K_1, K_2, K_3, K_4, K_5, K_6, K_7, K_7, K_6,$
 $K_5, K_4, K_3, K_2, K_1, K_0$
- Цикл расшифрования 32-Р:
 $K_0, K_1, K_2, K_3, K_4, K_5, K_6, K_7, K_7, K_6, K_5, K_4, K_3, K_2, K_1, K_0, K_7, K_6, K_5, K_4, K_3, K_2, K_1, K_0, K_7, K_6, K_5,$
 K_4, K_3, K_2, K_1, K_0
- Цикл выработки имитовставки 16-3:
 $K_0, K_1, K_2, K_3, K_4, K_5, K_6, K_7, K_0, K_1, K_2, K_3, K_4, K_5, K_6, K_7$

Режимы шифрования

ГОСТ 28147-89 предусматривает три следующих режима шифрования данных:

- простая замена,
- гаммирование,

- гаммирование с обратной связью,
- дополнительный режим выработки имитовставки

Простая замена

Зашифрование в данном режиме заключается в применении цикла 32-3 к блокам открытых данных, расшифрование – цикла 32-Р к блокам зашифрованных данных. Это наиболее простой из режимов, 64-битовые блоки данных обрабатываются в нем независимо друг от друга. Схемы алгоритмов зашифрования и расшифрования в режиме простой замены приведены на рисунках 2.43 и 2.44 соответственно. Размер массива открытых или зашифрованных данных, подвергающийся соответственно зашифрованию или расшифрованию, должен быть кратен 64 битам:

$$|T_o|=|T_{ш}|=64 \cdot n \quad (21)$$

где n – целое положительное число

после выполнения операции размер полученного массива данных не изменяется.

Режим шифрования простой заменой имеет следующие особенности:

- Так как блоки данных шифруются независимо друг от друга и от их позиции в массиве данных, при зашифровании двух одинаковых блоков открытого текста получаются одинаковые блоки шифртекста и наоборот. Отмеченное свойство позволит криптоаналитику сделать заключение о тождественности блоков исходных данных, если в массиве зашифрованных данных ему встретились идентичные блоки, что является недопустимым для серьезного шифра.
- Если длина шифруемого массива данных не кратна 8 байтам (64 битам), возникает проблема, чем и как дополнять последний неполный блок данных массива до полных 64 бит. Эта задача не так проста, как кажется на первый взгляд. Очевидные решения вроде «дополнить неполный блок нулевыми битами» или, более общее, «дополнить неполный блок фиксированной комбинацией нулевых и единичных битов» могут при определенных условиях дать в руки криптоаналитика возможность методами перебора определить содержимое этого самого неполного блока, и этот факт означает снижение стойкости шифра. Кроме того, длина шифртекста при этом изменится, увеличившись до ближайшего целого, кратного 64 битам, что часто бывает нежелательным.

На первый взгляд, перечисленные выше особенности делают практически невозможным использование режима простой замены, ведь он может применяться только для шифрования массивов данных с размером кратным 64 битам, не содержащим повторяющихся 64-битовых блоков. Кажется, что для любых реальных данных гарантировать выполнение указанных

условий невозможно. Это почти так, но есть одно очень важное исключение: вспомните, что размер ключа составляет 32 байта, а размер таблицы замен – 64 байта. Кроме того, наличие повторяющихся 8-байтовых блоков в ключе или таблице замен будет говорить об их весьма плохом качестве, поэтому в реальных ключевых элементах такого повторения быть не может. Таким образом, мы выяснили, что режим простой замены вполне подходит для шифрования ключевой информации, тем более, что прочие режимы для этой цели менее удобны, поскольку требуют наличия дополнительного синхронизирующего элемента данных – синхропосылки (см. следующий подпункт). И поэтому ГОСТ предписывает использовать режим простой замены исключительно для шифрования ключевых данных.

Гаммирование

Как же можно избавиться от недостатков режима простой замены? Для этого необходимо сделать возможным шифрование блоков с размером менее 64 бит и обеспечить зависимость блока шифртекста от его номера, иными словами, рандомизировать процесс шифрования. В ГОСТе это достигается двумя различными способами в двух режимах шифрования, предусматривающих гаммирование.

Гаммирование – это наложение (снятие) на открытые (зашифрованные) данные криптографической гаммы, то есть последовательности элементов данных, вырабатываемых с помощью некоторого криптографического алгоритма, для получения зашифрованных (открытых) данных.

Для наложения гаммы при зашифровании и ее снятия при расшифровании должны использоваться взаимно обратные бинарные операции. В ГОСТе для этой цели используется операция побитового сложения по модулю 2, поскольку она является обратной самой себе и, к тому же, наиболее просто реализуется аппаратно. Гаммирование решает обе упомянутые проблемы: во-первых, все элементы гаммы различны для реальных шифруемых массивов и, следовательно, результат зашифрования даже двух одинаковых блоков в одном массиве данных будет различным. Во-вторых, хотя элементы гаммы и вырабатываются одинаковыми порциями в 64 бита, использоваться может и часть такого блока с размером, равным размеру шифруемого блока.

Гамма для этого режима получается следующим образом: с помощью некоторого алгоритмического рекуррентного генератора последовательности чисел (РГПЧ) вырабатываются 64-битовые блоки данных, которые далее подвергаются преобразованию по циклу 32-3, то есть зашифрованию в режиме простой замены, в результате получают блоки гаммы. Благодаря тому, что наложение и снятие гаммы осуществляется при помощи одной и

той же операции побитового исключающего или, алгоритмы зашифрования и расшифрования в режиме гаммирования идентичны.

РГПЧ, используемый для выработки гаммы, является рекуррентной функцией:

$$\Omega_{i+1} = f(\Omega_i), \quad (22)$$

где Ω_i - элементы рекуррентной последовательности,

f – функция преобразования

Элемент данных « Ω_0 » является параметром алгоритма для режимов гаммирования, на схемах он обозначен как S , и называется в криптографии *синхросылкой*, а в ГОСТе – *начальным заполнением* одного из регистров шифрователя.

Синхросылка (вектор инициализации) – случайное число, которое регулярно обновляется, передается по каналу управления и используется для инициализации алгоритма шифрования.

Из соображений увеличения стойкости алгоритма, разработчики ГОСТа решили использовать для инициализации РГПЧ не непосредственно синхросылку, а результат ее преобразования по циклу 32-3.

Последовательность элементов, вырабатываемых РГПЧ, целиком зависит от его начального заполнения, то есть элементы этой последовательности являются функцией своего номера и начального заполнения РГПЧ:

$$\Omega_i = f_i(\Omega_0), \quad (23)$$

где $f_i(X) = f(f_{i-1}(X))$,

$f_0(X) = X$

С учетом преобразования по алгоритму простой замены добавляется еще и зависимость от ключа:

$$\Gamma_i = \Pi_{32-3}(\Omega_i) = \Pi_{32-3}(f_i(\Omega_0)) = \Pi_{32-3}(f_i(\Pi_{32-3}(S))) = \varphi_i(S, K),$$

где Γ_i – i -тый элемент гаммы,

K – ключ

Естественно, для обратимости процедуры шифрования в процессах за- и расшифрования должна использоваться одна и та же синхросылка. Из требования уникальности гаммы, невыполнение которого приводит к катастрофическому снижению стойкости шифра, следует, что для шифрования двух различных массивов данных на одном ключе необходимо обеспечить использование различных синхросылок. Это приводит к необходимости хранить или передавать синхросылку по каналам связи вместе с зашифрованными

данными, хотя в отдельных особых случаях она может быть предопределена или вычисляться особым образом, если исключается шифрование двух массивов на одном ключе.

Подробнее рассмотрим РГПЧ, используемый в ГОСТе для генерации элементов гаммы. Прежде всего, надо отметить, что к нему не предъявляются требования обеспечения каких-либо статистических характеристик вырабатываемой последовательности чисел. РГПЧ спроектирован разработчиками ГОСТа исходя из необходимости выполнения следующих условий:

- период повторения последовательности чисел, вырабатываемой РГПЧ, не должен сильно (в процентном отношении) отличаться от максимально возможного при заданном размере блока значения 2^{64} ;
- соседние значения, вырабатываемые РГПЧ, должны отличаться друг от друга в каждом байте, иначе задача криптоаналитика будет упрощена;
- РГПЧ должен быть достаточно просто реализуем как аппаратно, так и программно на наиболее распространенных типах процессоров, большинство из которых, как известно, имеют разрядность 32 бита

Исходя из перечисленных принципов, создатели ГОСТа спроектировали весьма удачный РГПЧ, имеющий следующие характеристики:

- в 64-битовом блоке старшая и младшая части обрабатываются независимо друг от друга, фактически, существуют два независимых РГПЧ для старшей и младшей частей блока:

$$\Omega_i = (\Omega_i^0, \Omega_i^1), \quad (23)$$

$$|\Omega_i^0| = |\Omega_i^1| = 32, \quad (24)$$

$$\Omega_{i+1}^0 = \hat{f}(\Omega_i^0), \quad (25)$$

$$\Omega_{i+1}^1 = \tilde{f}(\Omega_i^1) \quad (26)$$

- рекуррентные соотношения для старшей и младшей частей следующие (нижний индекс в записи числа означает его систему счисления, таким образом, константы, используемые на данном шаге, записаны в 16-ричной системе счисления):

$$\Omega_{i+1}^0 = (\Omega_i^0 + C_0) \bmod 2^{32} \quad (27)$$

где $C_0 = 1010101_{16}$

$$\Omega_{i+1}^1 = (\Omega_i^1 + C_1 - 1) \bmod (2^{32} - 1) + 1 \quad (28)$$

где $C_1 = 1010104_{16}$

Второе выражение нуждается в комментариях, так как в тексте ГОСТа приведено нечто другое:

$$\Omega_{i+1}^1 = (\Omega_i^1 + C_1) \bmod (2^{32} - 1) \quad (29)$$

тем же значением константы C_1 . Но далее в тексте стандарта дается комментарий, что, оказывается, под операцией взятия остатка по модулю $(2^{32}-1)$ там понимается не то же самое, что и в математике. Отличие заключается в том, что согласно ГОСТу:

$$(2^{32}-1) \bmod (2^{32}-1) = (2^{32}-1), a \neq 0 \quad (30)$$

На самом деле, это упрощает реализацию формулы, а математически корректное выражение для нее приведено в формуле (30).

Период повторения последовательности для младшей части составляет 2^{32} , для старшей части $(2^{32}-1)$, для всей последовательности период составляет $2^{32} * (2^{32}-1)$. Первая формула из двух реализуется за одну команду, вторая, несмотря на ее кажущуюся громоздкость, за две команды на всех современных 32-разрядных процессорах – первой командой идет обычное сложение по модулю 2^{32} с запоминанием бита переноса, а вторая команда прибавляет бит переноса к полученному значению.

Рассмотрим шаги алгоритма гаммирования, представленного на рисунке 2.46:

Шаг 0 Определяет исходные данные для основного шага криптопреобразования.

$T_{o(ш)}$ – массив открытых (зашифрованных) данных произвольного размера, подвергаемый процедуре зашифрования (расшифрования), по ходу процедуры массив подвергается преобразованию порциями по 64 бита;

S – синхропосылка, 64-битовый элемент данных, необходимый для инициализации генератора гаммы;

Шаг 1 Начальное преобразование синхропосылки.

Выполняется для ее «рандомизации», то есть для устранения статистических закономерностей, присутствующих в ней, результат используется как начальное заполнение РГПЧ;

Шаг 2 Один шаг работы РГПЧ, реализующий его рекуррентный алгоритм.

В ходе данного шага старшая (S_1) и младшая (S_0) части последовательности данных вырабатываются независимо друг от друга;

Шаг 3 Гаммирование.

Очередной 64-битовый элемент, выработанный РГПЧ, подвергается процедуре зашифрования по циклу 32–3, результат используется как элемент гаммы для зашифрования (расшифрования) очередного блока открытых (зашифрованных) данных того же размера;

Шаг 4 Завершение гаммирования.

Результат работы алгоритма – зашифрованный (расшифрованный) массив данных.

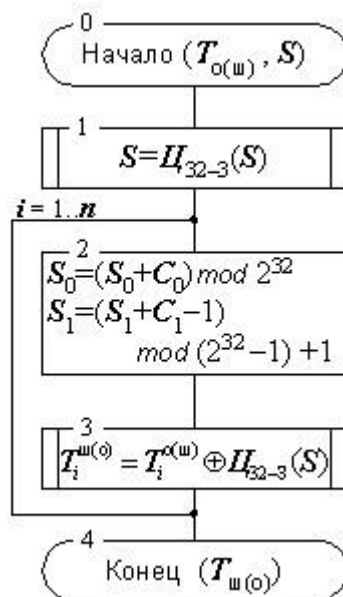


Рис. 2.46. Алгоритм зашифрования (расшифрования) данных в режиме гаммирования

Особенности гаммирования как режима шифрования:

1. Одинаковые блоки в открытом массиве данных дадут при зашифровании различные блоки шифртекста, что позволит скрыть факт их идентичности.

2. Поскольку наложение гаммы выполняется побитно, шифрование неполного блока данных легко выполнимо как шифрование битов этого неполного блока, для чего используется соответствующие биты блока гаммы. Так, для зашифрования неполного блока в 1 бит согласно стандарту следует использовать самый младший бит из блока гаммы.

3. Синхропосылка, использованная при зашифровании, каким-то образом должна быть передана для использования при расшифровании. Это может быть достигнуто следующими путями:

- хранить или передавать синхропосылку вместе с зашифрованным массивом данных, что приведет к увеличению размера массива данных при зашифровании на размер синхропосылки, то есть на 8 байт;

- использовать predetermined значение синхропосылки или вырабатывать ее синхронно источником и приемником по определенному закону, в этом случае изменение размера передаваемого или хранимого массива данных отсутствует

4. Биты массива данных шифруются независимо друг от друга. Таким образом, каждый бит шифртекста зависит от соответствующего бита открытого текста и, естественно, порядкового номера бита в массиве. Из этого вытекает, что изменение

бита шифртекста на противоположное значение приведет к аналогичному изменению бита открытого текста на противоположный. Данное свойство дает злоумышленнику возможность, воздействуя на биты шифртекста, вносить предсказуемые и даже целенаправленные изменения в соответствующий открытый текст, получаемый после его расшифрования, не обладая при этом секретным ключом. Это иллюстрирует хорошо известный в криптологии факт, что *секретность* и *аутентичность* различные свойства криптографических систем.

Гаммирование с обратной связью

Данный режим очень похож на режим гаммирования и отличается от него только способом выработки элементов гаммы – очередной элемент гаммы вырабатывается как результат преобразования по циклу 32-3 предыдущего блока зашифрованных данных, а для зашифрования первого блока массива данных элемент гаммы вырабатывается как результат преобразования по тому же циклу синхропосылки (рисунок 2.7). Этим достигается сцепление блоков – каждый блок шифртекста в этом режиме зависит от соответствующего и всех предыдущих блоков открытого текста, это видно и в уравнении режима шифрования (расшифрования) гаммирования с обратной связью (формула 2.25). Поэтому данный режим иногда называется *гаммированием с сцеплением блоков*. На стойкость шифра факт сцепления блоков не оказывает никакого влияния.

Шифрование в режиме гаммирования с обратной связью обладает теми же особенностями, что и шифрование в режиме обычного гаммирования, за исключением влияния искажений шифртекста на соответствующий открытый текст (свойство 4).

$$T_i^o = T_i^w \oplus C_{32-3}(T_{i-1}^w) \quad (31)$$

Если в режиме обычного гаммирования изменения в определенных битах шифртекста влияют только на соответствующие биты открытого текста, то в режиме гаммирования с обратной связью картина несколько сложнее. Как видно из соответствующего уравнения, при расшифровании блока данных в режиме гаммирования с обратной связью, блок открытых данных зависит от соответствующего и предыдущего блоков зашифрованных данных. Поэтому, если внести искажения в зашифрованный блок, то после расшифрования искаженными окажутся два блока открытых данных – соответствующий и следующий за ним, причем искажения в первом случае будут носить тот же характер, что и в режиме гаммирования, а во втором случае – как в режиме простой замены. Другими словами, в соответствующем блоке открытых данных искаженными окажутся те же самые биты, что и в

блоке зашифрованных данных, а в следующем блоке открытых данных все биты независимо друг от друга с вероятностью 1/2 изменят свои значения.

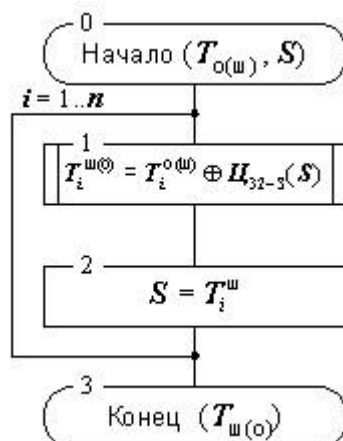


Рис. 2.47. Алгоритм зашифрования (расшифрования) данных в режиме гаммирования с обратной связью.

Выработка имитовставки к массиву данных

Для решения задачи обнаружения искажений в зашифрованном массиве данных с заданной вероятностью в ГОСТе предусмотрен дополнительный режим криптографического преобразования – выработка имитовставки (рисунок 2.48).

Целью использования имитовставки является обнаружение всех случайных или преднамеренных изменений в массиве информации. Для потенциального злоумышленника две следующие задачи практически неразрешимы, если он не владеет ключом:

- вычисление имитовставки для заданного открытого массива информации;
- подбор открытых данных под заданную имитовставку

В качестве имитовставки берется часть блока, полученного на выходе, обычно – 32 его младших бита. При выборе размера имитовставки надо принимать во внимание, что вероятность успешного навязывания ложных данных равна величине 2^{-11} на одну попытку подбора, если в распоряжении злоумышленника нет более эффективного метода подбора, чем простое угадывание. При использовании имитовставки размером 32 бита эта вероятность равна $2^{-32} \approx 0.23 \cdot 10^{-9}$.

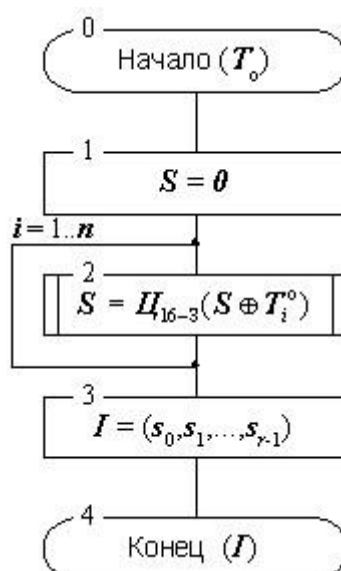


Рис. 2.48. Алгоритм выработки имитовставки для массива данных

Криптоанализ алгоритма

В 1994 г. описание алгоритма ГОСТ 28147-89 было переведено на английский язык и опубликовано в статье Волонгонгского университета (Австралия) «Советский алгоритм шифрования». Именно после этого стали появляться результаты его анализа, выполненного зарубежными специалистами; однако, в течение значительного времени не было найдено каких-либо атак, приближающихся к практически осуществимым.

Анализ таблиц замен

Поскольку таблицы замен в стандарте не приведены, в ряде работ высказывается предположение, что «компетентная организация» может выдать как «хорошие», так и «плохие» таблицы замен. Ясно, что криптостойкость алгоритма во многом зависит от свойств используемых таблиц замен, соответственно, существуют слабые таблицы замен, применение которых может упростить вскрытие алгоритма. Тем не менее, возможность использования различных таблиц замен кажется весьма достойной идеей, в пользу которой можно привести два следующих факта из истории стандарта шифрования США DES:

- Атаки с помощью как линейного, так и дифференциального криптоанализа алгоритма DES используют конкретные особенности таблиц замен. При использовании других таблиц криптоанализ придется начинать сначала
- Были попытки усилить DES против линейного и дифференциального криптоанализа путем использования более стойких таблиц замен. Такие таблицы, действительно более стойкие, были предложены, например, в алгоритме s^5 DES. Но, увы, заменить DES на s^5 DES было невозможно, поскольку таблицы замен жестко

определены в стандарте, соответственно, реализации алгоритма наверняка не поддерживают возможность смены таблиц на другие

Однако, в работе доказано, что секретные таблицы замен могут быть вычислены с помощью следующей атаки, которая может быть применена практически:

1. Устанавливается нулевой ключ и выполняется поиск «нулевого вектора». Этот этап занимает порядка 2^{32} операций шифрования.
2. С помощью нулевого вектора вычисляются значения таблиц замен, что занимает не более 2^{11} операций.

Модификации алгоритма и их анализ

Алгоритмы:

- GOST-H, в котором, относительно оригинального алгоритма, изменен порядок использования подключей, а именно, в раундах с 25-го по 32-й подключи используются в прямом порядке, т.е. точно так же, как и в предыдущих раундах алгоритма;
- 20-раундовый алгоритм GOST \tilde{A} , в раунде которого для наложения ключа используется операция XOR вместо сложения по модулю 2^{32} .

По результатам анализа сделан вывод о том, что GOST-H и GOST \tilde{A} слабее исходного алгоритма ГОСТ 28147-89, поскольку оба имеют классы слабых ключей.

Анализ полнораундового алгоритма

Существуют атаки и на полнораундовый ГОСТ 28147-89 без каких-либо модификаций. Одна из первых открытых работ, в которых был проведен анализ алгоритма посвящена атакам, использующим слабости процедур расширения ключа ряда известных алгоритмов шифрования. В частности, полнораундовый алгоритм ГОСТ 28147-89 может быть вскрыт с помощью дифференциального криптоанализа на связанных ключах, но только в случае использования слабых таблиц замен. 24-раундовый вариант алгоритма (в котором отсутствуют первые 8 раундов) вскрывается аналогичным образом при любых таблицах замен, однако, сильные таблицы замен делают такую атаку абсолютно непрактичной.

Отечественные ученые А.Г. Ростовцев и Е.Б. Маховенко в 2001 г. предложили принципиально новый метод криптоанализа (по мнению авторов, существенно более эффективный, чем линейный и дифференциальный криптоанализ) путем формирования целевой функции от известного открытого текста, соответствующего ему шифртекста и искомого значения ключа и нахождения ее экстремума, соответствующего истинному значению ключа. Они же нашли большой класс слабых ключей алгоритма ГОСТ 28147-89,

которые позволяют вскрыть алгоритм с помощью всего 4-х выбранных открытых текстов и соответствующих им шифртекстов с достаточно низкой сложностью.

В 2004 г. группа специалистов из Кореи предложила атаку, с помощью которой, используя дифференциальный криптоанализ на связанных ключах, можно получить с вероятностью 91,7% 12 бит секретного ключа. Для атаки требуется 2^{35} выбранных открытых текстов и 2^{36} операций шифрования. Как видно, данная атака, практически, бесполезна для реального вскрытия алгоритма.

Достоинства ГОСТа

- бесперспективность силовой атаки, т.е. полным перебором (XSL-атаки в учёт не берутся, т.к. их эффективность на данный момент полностью не доказана);
- эффективность реализации и соответственно высокое быстродействие на современных компьютерах.
- наличие защиты от навязывания ложных данных (выработка имитовставки) и одинаковый цикл шифрования во всех четырех алгоритмах ГОСТа

Недостатки ГОСТа

Основные проблемы ГОСТа связаны с неполнотой стандарта в части генерации ключей и таблиц замен. Тривиально доказывается, что у ГОСТа существуют «слабые» ключи и таблицы замен, но в стандарте не описываются критерии выбора и отсева «слабых». Также стандарт не специфицирует алгоритм генерации таблицы замен (S-блоков). С одной стороны, это может являться дополнительной секретной информацией (помимо ключа), а с другой, поднимает ряд проблем:

- нельзя определить криптостойкость алгоритма, не зная заранее таблицы замен;
- реализации алгоритма от различных производителей могут использовать разные таблицы замен и могут быть несовместимы между собой;
- возможность преднамеренного предоставления слабых таблиц замен лицензирующими органами;
- потенциальная возможность (отсутствие запрета в стандарте) использования таблиц замены, в которых узлы не являются перестановками, что может привести к чрезвычайному снижению стойкости шифра.

Аппаратные шифраторы серии Криптон

Устройства криптографической защиты данных серии КРИПТОН — это аппаратные шифраторы для IBM PC-совместимых компьютеров. Устройства применяются в составе средств и систем криптографической защиты данных для обеспечения информационной безопасности (в том числе защиты с высоким уровнем секретности) в государственных и коммерческих структурах.

Устройства КРИПТОН гарантируют защиту информации, обрабатываемой на персональном компьютере и/или передаваемой по открытым каналам связи.

Устройства КРИПТОН выполнены в виде плат расширения ISA и PCI персонального компьютера с процессором i386 и выше.

Устройства КРИПТОН разработаны, производятся и реализуются фирмой АНКАД. Они построены на разработанных фирмой АНКАД специализированных 32-разрядных шифрпроцессорах серии БЛЮМИНГ.

За 10 лет работы Фирма АНКАД поставила более 15 тысяч устройств КРИПТОН заказчикам в Центральном Банке, Федеральном агентстве правительственной связи и информации при Президенте РФ, министерствах обороны и внутренних дел, Министерстве по налогам и сборам, Федеральном казначействе, коммерческих банках, финансовых и страховых компаниях, многим корпоративным клиентам.

Сеть кооперационного производства устройств КРИПТОН охватывает наиболее известные предприятия российской электроники (ОАО “Ангстрем” и др.).

Устройства серии КРИПТОН имеют сертификаты соответствия требованиям ФСБ и ФСТЭК (в том числе в составе абонентских пунктов и автоматизированных рабочих мест для защиты информации, содержащей сведения, составляющие государственную тайну).

Таблица 2.24. Основные технические данные и характеристики серии КРИПТОН
(данные за 2004-2005г.г.)

Алгоритм шифрования	ГОСТ 28147-89
Размерность секретного ключа шифрования, бит	256 (количество возможных комбинаций ключей — 10^{77})
Размерность открытого ключа, бит	512 или 1024
Количество уровней ключевой системы	3 (главный ключ — пользовательский/сетевой ключ — сеансовый ключ)
Датчик случайных чисел	аппаратный (аттестован экспертной организацией)
Отклонение распределения значения	не более 0,0005

случайных чисел от равновероятного распределения	
Поддерживаемые операционные системы	MS-DOS, Windows 95(98)/ME/NT 4.0/2000/XP/2003 UNIX (Solaris/Intel) (возможно создание оригинальных программных драйверов для работы устройств)
Скорость передачи данных по сети	10/100 Мбит/с
Скорость шифрования	до 9 Мбайт/с – сетевые шифраторы до 240 Мбит/с – шифраторы жесткого диска
Класс защиты по классификации ФСБ	до КС1
Алгоритм ЭЦП	ГОСТ Р 34.10-94 (с длиной секретного ключа 256 бит)

Описание лабораторной установки и методики измерений

Интерфейс учебно-программного комплекса

Лабораторный комплекс был разработан на языке высокого уровня Visual Basic 6.

Комплекс позволяет работать только с введенным текстом (обработка ведется в двоичном коде, преобразование к которому производится посредством ASCII кодов).

Для удобства работы комплекс оснащен всплывающими подсказками на все входные и выходные параметры, рабочие поля, активные элементы и элементы управления.

Оптимизация кода программы не производилась, поэтому при работе с большими объемами текста возможны некорректное отображение результата или зависание программы.

Основные элементы комплекса:

1. Поле для ввода ключа – максимальная длина ключа 32 символа, при его отсутствии будет использован нулевой вектор
2. Выбор режима шифрования:
 - При выборе простой замены есть возможность проверить кратность текста 8 байтовым блокам;
 - При выборе режимов гаммирования есть возможность ввести вектор инициализации (синхропосылку), при ее отсутствии будет использован нулевой вектор

3. Поле для ввода текста для зашифрования
4. Возможность очистить поле шифрования
5. Возможность узнать об'ем текста в поле шифрования
6. Кнопка для зашифрования текста
7. Время, затраченное на зашифрования – напротив поля текста
8. Поле зашифрованного текста
9. Возможность очистить поле зашифрованного текста
10. Возможность узнать об'ем зашифрованного текста
11. Кнопка для расшифрования текста
12. Время, затраченное на расшифровку – напротив поля шифротекста
13. Системное время вашего компьютера
14. Возможность использовать материалы в электронном виде (рисунок 2.49)

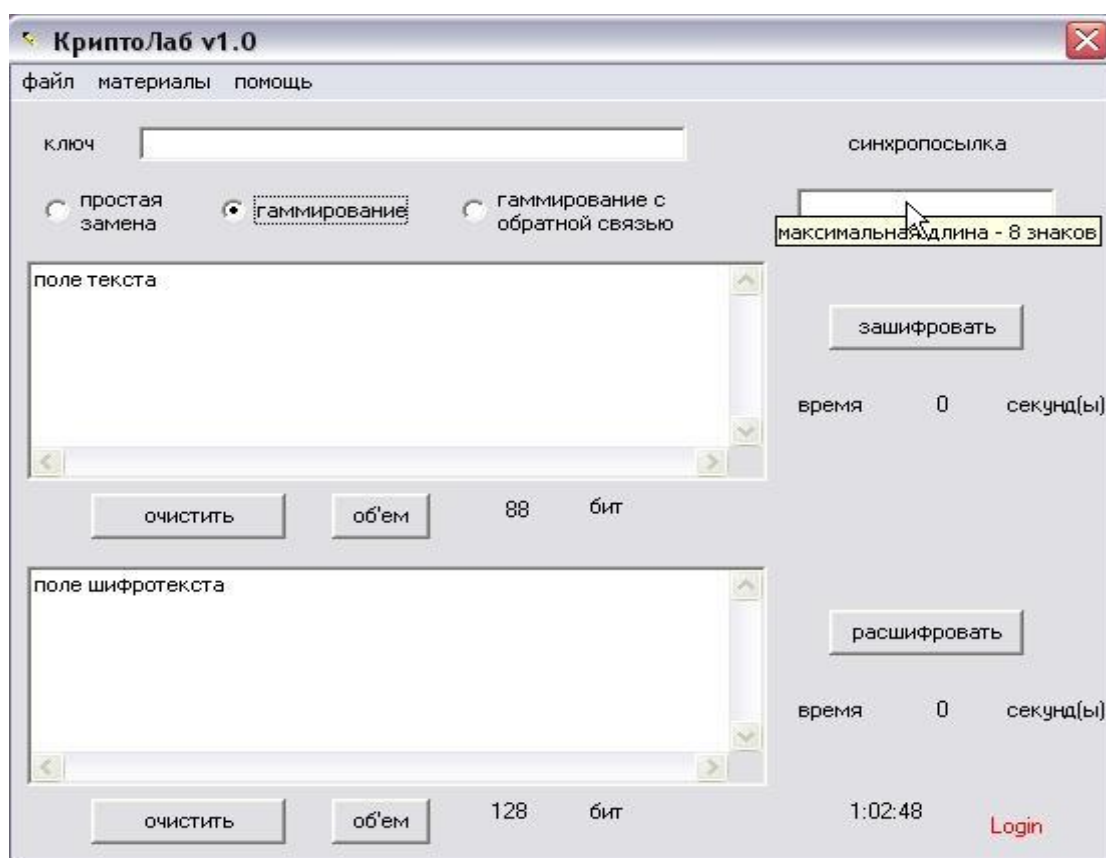


Рис. 2.49. Лабораторный комплекс

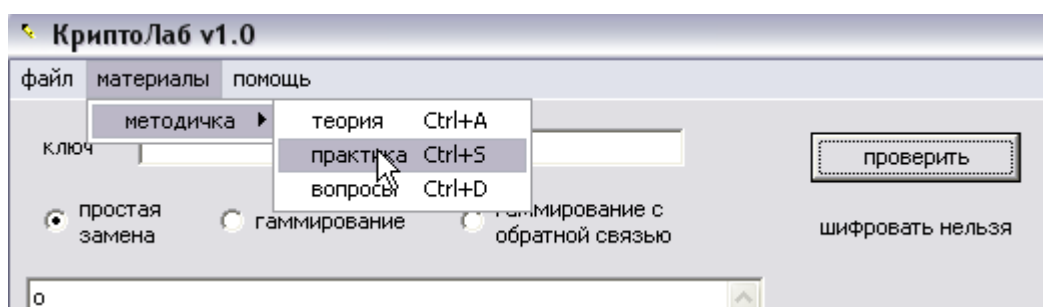


Рис. 7.50. Предоставленные материалы

Порядок выполнения работы

1. Ознакомится с теорией по стандарту ГОСТ 28147-89 (режимы шифрования, строение основного шага криптопреобразования, особенности использования таблицы замен).

2. Взять у преподавателя задание, состоящие из:

уникального ключа для шифрования/дешифрования текстовых сообщений, синхропосылки

3. Исследование стандарта шифрования

3.1 выбрать режим шифрования «простая замена»

3.2 ввести значение ключа

3.3 ввести тестовый текст (взять у преподавателя или пройти **материалы** → **методичка**

→ **практика** → **тексты** → **простая замена**

3.4 проверить кратность 8 байтовых блоков нажатием кнопки «проверить»

3.5 зашифровать текст нажатием соответствующей кнопки

3.6 очистить текстовое поле

3.7 расшифровать текст

3.8 проверить наличие ошибок

3.9 зафиксировать об'емы текстов и время за/расшифрования в таблицу

3.10 Прodelать пункты 3.3-3.9 для оставшихся текстов

3.11 выбрать режим шифрования «гаммирование»

3.12 ввести тестовый текст (взять у преподавателя или пройти **материалы** → **методичка**

→ **практика** → **тексты** → **гаммирования**

3.13 ввести значения ключа и синхропосылки

3.14 зашифровать текст нажатием соответствующей кнопки

3.15 очистить текстовое поле

3.16 расшифровать текст

3.17 проверить наличие ошибок

- 3.18 зафиксировать объемы текстов и время за/расшифрования в таблицу
- 3.19 Прodelать пункты 3.14-3.18 для оставшихся текстов
- 3.20 прodelать пункты 3.11-3.19 для режима «гаммирование с обратной связью»
- 3.21 Построить графики зависимости размера файла от времени его шифрования /расшифрования для всех режимов.
- 3.22 проверить изменение шифротекста при изменении ключа и синхропосылки
- 3.23 посмотреть как изменится текст после расшифрования если в шифротекст внести изменения (для всех режимов шифрования)

Таблица 7.25. Результаты измерения

Размер файла				
Время зашифрования				
Время расшифрования				

Содержание отчета

- 1 Цель работы.
- 2 Описание действий по каждому пункту.
- 3 Результаты прodelанной работы (таблицы, графики, скриншоты).
- 4 Выводы.

Контрольные вопросы

1. Что такое симметричное шифрование, блочные коды?
2. Что представляет собой стандарт ГОСТ 28147-89 (длина ключа, размер входного блока, ключа, таблица замен)?
3. Что собой представляет архитектура данного стандарта?
4. Как устроен основной шаг криптопреобразования?
5. Сколько циклов шифрования предусмотрено стандартом?
6. Что быстрее шифрование или расшифрование? Почему?
7. Какие режимы шифрования применяются в стандарте ГОСТ 28147-89?
8. Какие режимы быстрее при расшифровании? Почему?
9. Какие режимы лучше восстанавливают зашифрованную информацию при ошибке в одном символе? Почему?
10. С какой целью используется синхропосылка или вектор инициализации?

Описание лабораторной установки шифрование AES и методики измерений

Интерфейс учебно-программного комплекса

Учебно-программный комплекс (в дальнейшем просто комплекс) был разработан на языке высокого уровня Visual Basic.

Главное окно

На рисунке 2.51 приведен интерфейс главного окна комплекса.

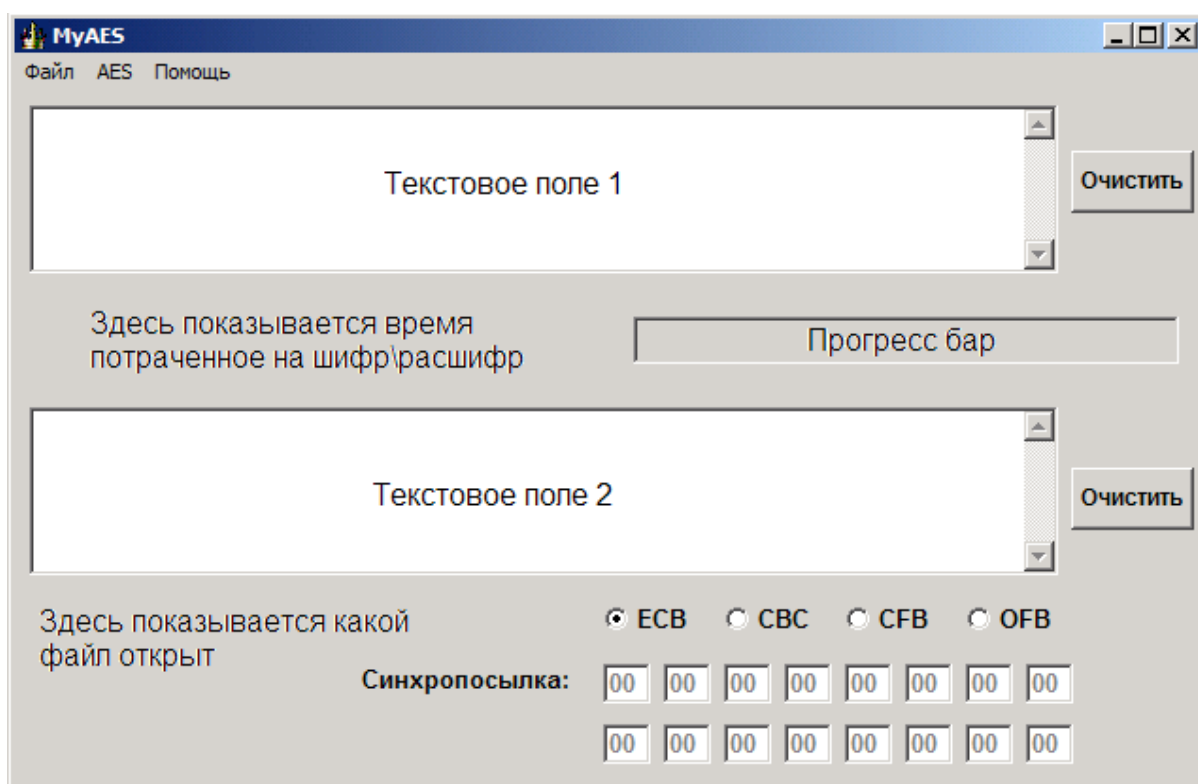


Рис. 7.51. Интерфейс главного окна комплекса

Главное окно состоит из:

1. Текстовое поле 1 – сюда вводится открытый текст, также если открывается файл с расширением txt, то его содержимое помещается сюда же.
2. Текстовое поле 2 – сюда помещается зашифрованный текст, также если открывается файл с расширением aes, то его содержимое помещается сюда же.

3. Прогресс бар – наглядно показывает время требуемое для шифрования \ расшифрования.

4. Кнопки “Очистить” – удаляют весь текст из текстовых полей.

5. Переключатели ECB, CBC, CFB, OFB – показывают, какой режим блочного шифрования используется.

6. Синхропосылка – задает синхропосылку для CBC, CFB и OFB.

В главном окне и происходит шифрование и расшифрование. Также в нем можно выбрать режимы симметричного шифрования (ECB, CBC, CFB, OFB). Про них сказано ниже.

Пункт меню “Файл”

На рисунке 2.52 приведен пункт меню “Файл”:

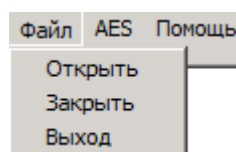


Рис. 7.52. Пункт меню “Файл”

Пункт меню “Файл” содержит следующие опции:

1. Открыть – открывает файл (*.txt, *.aes), но не больше 32кВ для текстовых файлов и 64кВ для файлов типа aes.

2. Закреть – закрывает выбранный файл.

3. Выход – завершает работу комплекса.

Пункт меню “AES”

На рисунке 13 приведен пункт меню “AES”:

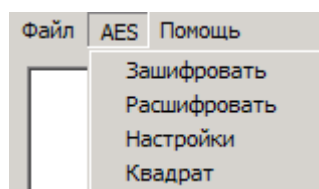


Рис. 7.53. Пункт меню “AES”

Пункт меню “AES” содержит следующие опции:

1. Зашифровать – шифрует текст, расположенный в текстовом поле 1, и помещает его в текстовое поле 2.
2. Расшифровать - расшифровывает текст, расположенный в текстовом поле 2, и помещает его в текстовое поле 1.
3. Настройки – позволяет выбрать свой ключ шифрования, а также включить \ выключить режим логирования. (рисунок 7.54)

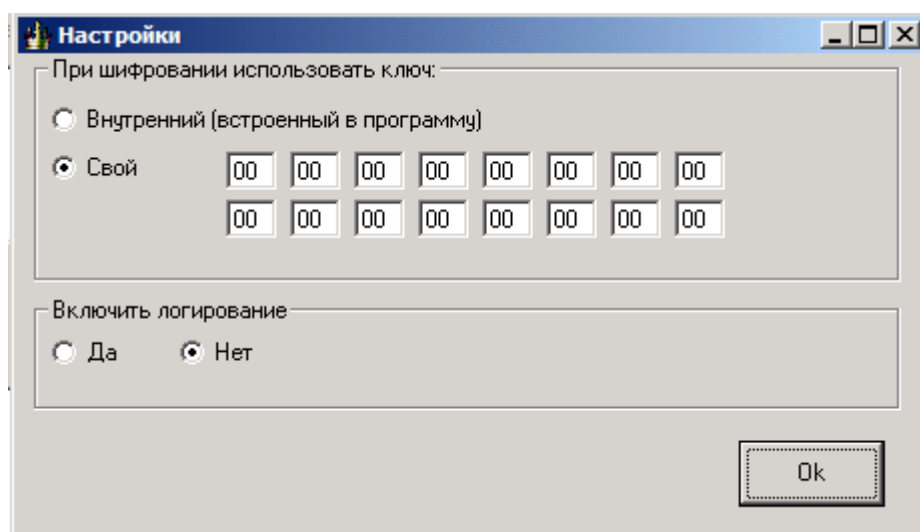


Рис. 7.54. Окно “Настройки”

В этом окне вы можете ввести свой ключ шифрования (например из задания по лабораторной работе) или же использовать внутренний. При использовании внутреннего ключа шифрования вы можете шифровать свои сообщения и отправлять их кому-либо. Причем получателю не обязательно знать ключ, главное чтобы у него была данная программа.

Здесь же вы можете включить логирование. При включенном логировании в папке программы создается Encrypt.txt (Decrypt.txt), в зависимости от того, что вы делаете шифруете или расшифровываете. В этом файле содержатся все преобразования, происходящие с открытым текстом. Пример такого файла приведен на рисунке 2.55:

```

Зашифрование
Входной блок CFF0EEE2E5F0EAE00101010101010101
Раундовый ключ: 00000000000000000000000000000000
Состояние после сложения с ключом: CFF0EEE2E5F0EAE0010101010101010101
Раунд №1
Состояние после SubBytes: 8A8C2898D98C87E17C7C7C7C7C7C7C7C
Состояние после ShiftRows: 8A8C7C7CD97C7C987C7C28E17C8C877C
Состояние после MixColumns: 80717A8DC93DEE5BB51D68098C916177
Раундовый ключ: 62636363626363636263636362636363
Состояние после сложения с ключом: E21219EEAB5E8D38D77E0B6AEEF20214
Раунд №2
Состояние после SubBytes: 98C9D42862585D070EF32B02288977FA
Состояние после ShiftRows: 98582BFA62F377280E89D40728C95D02
Состояние после MixColumns: 12AF832F952E07724F673D414F445DE8
Раундовый ключ: 9B9898C9F9FBFBA9B9898C9F9FBFBA9
Состояние после сложения с ключом: 89371BE66CD5FCD8D4FFA588B6BFA642
Раунд №3
Состояние после SubBytes: A79AAF8E5003B061481606C44E08242C
Состояние после ShiftRows: A703062C5016248E4808AF614E9AB0C4
Состояние после MixColumns: 7A87DCAF309E87C546D3A6BD5D6EF86B
Раундовый ключ: 90973450696CCFFAF2F457330B0FAC99
Состояние после сложения с ключом: EA10E8FF59F2483FB427F18E566154F2

```

Рис. 2.55. Зашифрование слова “Проверка”

Логирование работает только в режиме ECB, и автоматически отключается при открытии файла или использовании другого режима.

4. Квадрат – простенькая реализация единственной атаки на AES, под названием квадрат. Реализована для сокращенной версии AES, для 4 раундов. В стандарте же их 10.

Окно Атака «Квадрат»

Смысл атаки “Квадрат” описан выше.

В этом окне (рисунок 16) можно просмотреть ключи для любого раунда. При открытии этого окна в поля “Реальный ключ” загружается ключ, заданный в настройках. При нажатии кнопок “Up” или “Down” будет меняться номер раунда, а также поля “Реальный ключ”. Они будут равны тому раундовому ключу, который указан в строке “Номер раунда”. Нас будет интересовать ключ 4 раунда поэтому, лучше поставить значение поля “Номер раунда” в 4.

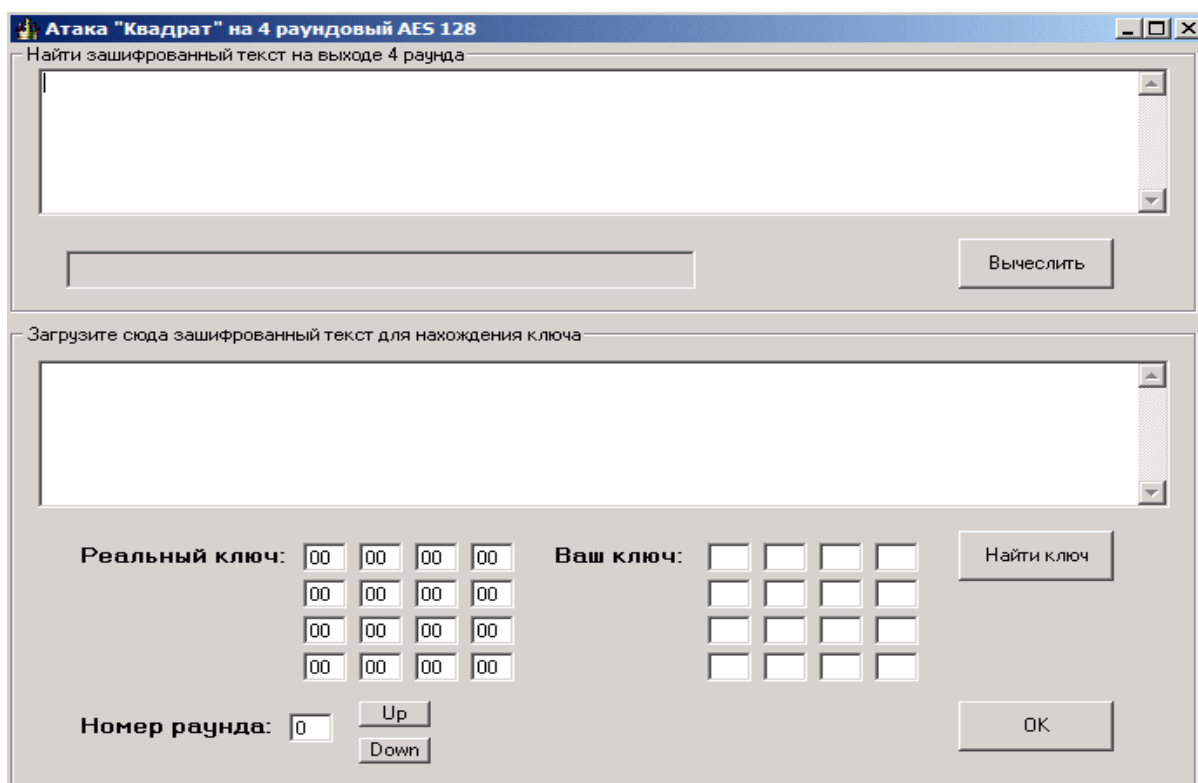


Рис. 7.56. Окно “Квадрат”

При нажатии кнопки “Вычислить” в текстовом поле 1 появляется зашифрованный L набор после 4 раунда. Прогресс бар наглядно показывает время требуемое для его вычисления. Пример L набора, зашифрованного ключом “00” $\{32\}$, показан на рисунке 2.57:

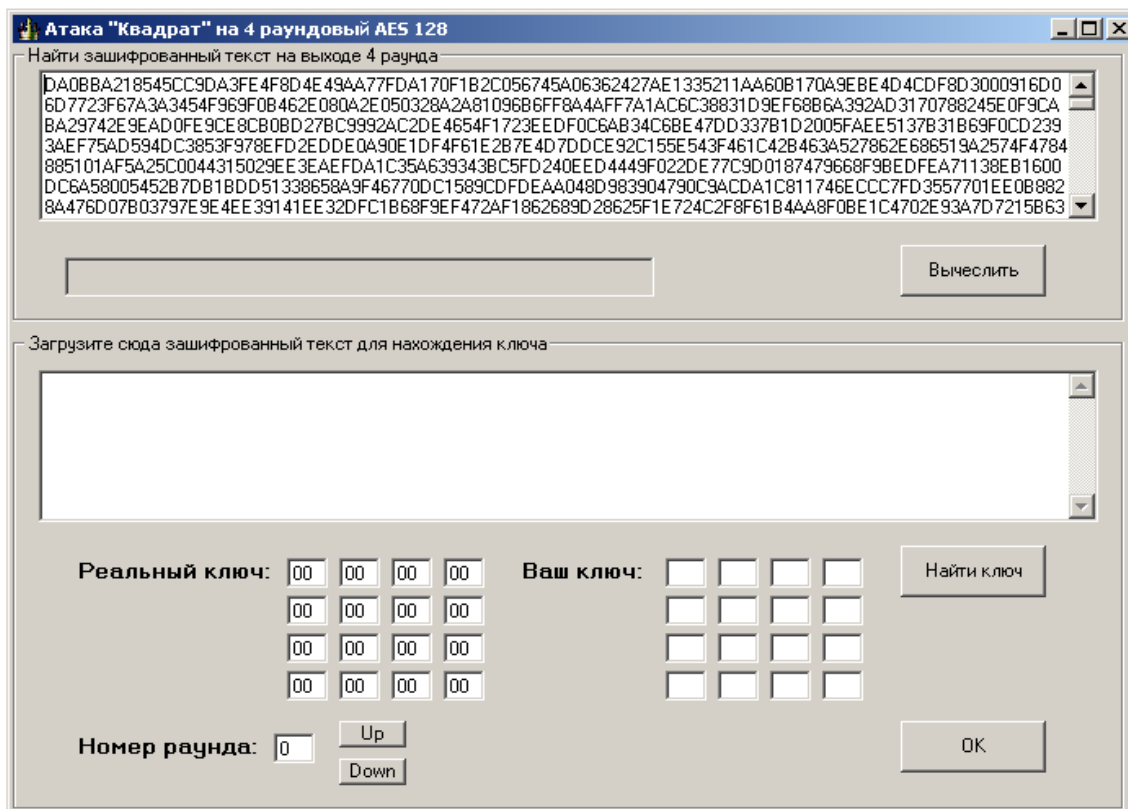


Рис. 7.57. Пример L набора, зашифрованного ключом “00” $\{32\}$

Этот текст нужно скопировать и вставить в текстовое поле 2.

При нажатии на кнопку “Найти ключ” начнется вычисление ключа 4 раунда на основе текста в текстовом поле 2. Прогресс бар наглядно показывает время, требуемое для его вычисления.

После вычисления ключа он появится в окошках “Ваш ключ”. Можно сравнить окна “Ваш ключ” и “Реальный ключ” и увидеть, что они практически совпадают (рисунок 2.58).

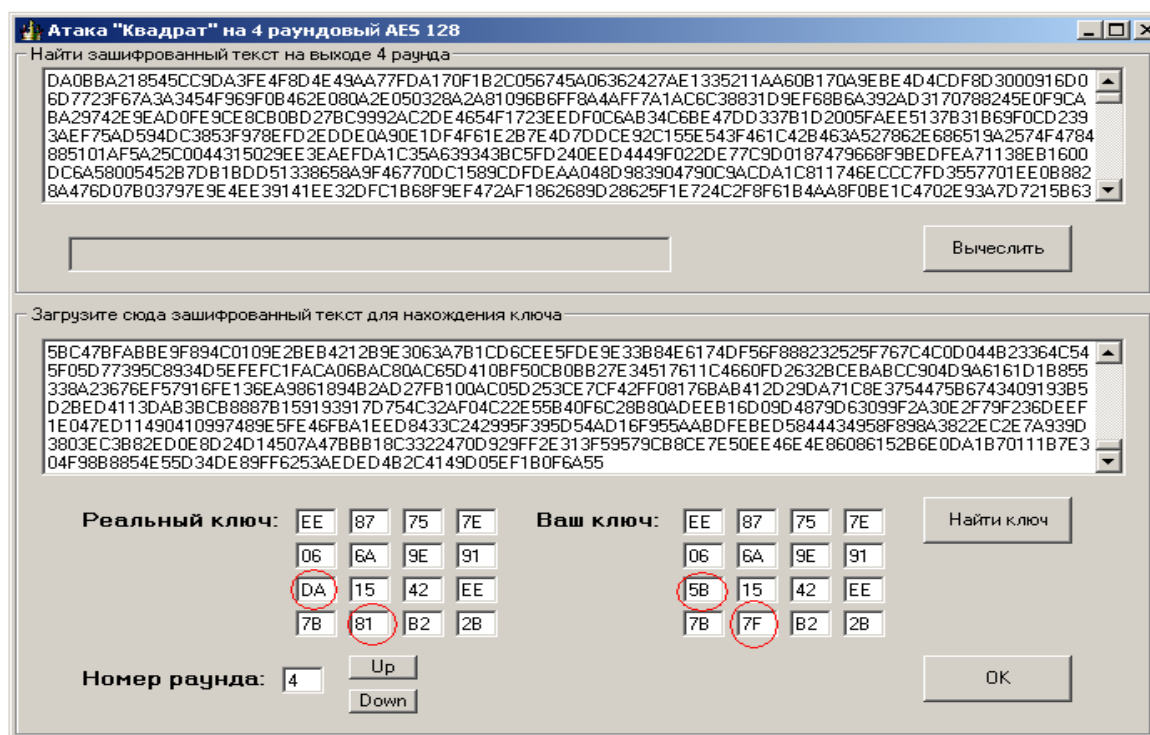


Рис. 7.58. Пример вычисления раундового ключа

Режимы ECB, CBC, CFB, OFB

Для различных ситуаций, встречающихся на практике, разработано значительное количество режимов шифрования.

Режим ECB (Electronic Code Book – режим электронной кодовой книги)

В режиме ECB каждый блок открытого текста заменяется блоком шифротекста. А так как один и тот же блок открытого текста заменяется одним и тем же блоком шифротекста, теоретически возможно создать кодовую книгу блоков открытого текста и соответствующих шифротекстов. Но если размер блока составляет n бит, кодовая книга будет состоять из 2^n записей.

Режим ECB - простейший режим шифрования. Все блоки открытого текста шифруются независимо друг от друга. Это важно для зашифрованных файлов с произвольным доступом, например, файлов баз данных. Если база данных зашифрована в режиме ECB, любая запись может быть добавлена, удалена, зашифрована или расшифрована независимо от любой другой записи (при условии, что каждая запись состоит из целого числа блоков шифрования). Кроме того, обработка может быть параллельной: если используются несколько шифровальных процессоров, они могут шифровать или расшифровывать различные блоки независимо друг от друга. Режим ECB показан на рисунке 2.59:

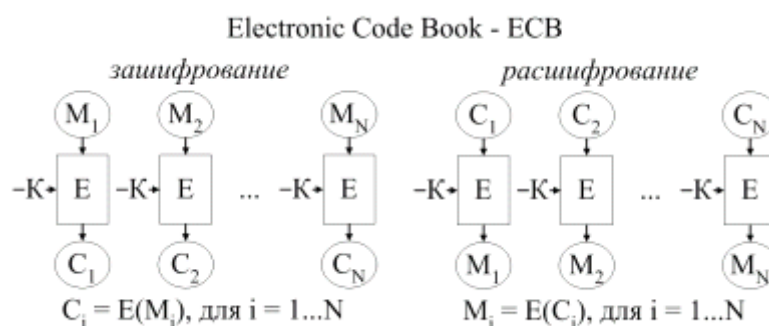


Рис. 7.59. Режим ECB

К недостаткам режима ECB можно отнести то обстоятельство, что если у криптоаналитика есть открытый текст и шифротекст нескольких сообщений, он может, не зная ключа, начать составлять шифровальную книгу. В большинстве реальных ситуаций фрагменты сообщений имеют тенденцию повторяться. В различных сообщениях могут быть одинаковые битовые последовательности. В сообщениях, которые, подобно электронной почте, создаются компьютером, могут быть периодически повторяющиеся структуры. Сообщения могут быть высоко избыточными или содержать длинные строки нулей или пробелов.

К достоинствам режима ECB можно отнести возможность шифрования нескольких сообщений одним ключом без снижения надежности. По существу, каждый блок можно рассматривать как отдельное сообщение, шифрованное тем же самым ключом. При расшифровании ошибки в символах шифротекста ведут к некорректному расшифрованию соответствующего блока открытого текста, однако не затрагивают остальной открытый текст. Но если бит шифротекста случайно потерян или добавлен, весь последующий шифротекст будет дешифрован некорректно, если только для выравнивания границ блоков не используется какое-нибудь выравнивание по границам блока. Большинство сообщений не делятся точно на n - битовые блоки шифрования – в конце обычно оказывается укороченный блок. Однако режим ECB требует использовать строго n - битовые блоки. Для решения этой проблемы используют дополнение (padding). Чтобы создать полный блок, последний блок дополняют некоторым стандартным шаблоном - нулями, единицами, чередующимися нулями и единицами.

Режим CBC (Ciphertext Block Chaining – режим сцепления блоков шифротекста)

Сцепление добавляет в блочный шифр механизм обратной связи: результаты шифрования предыдущих блоков влияют на шифрование текущего блока, т.е. каждый блок используется для модифицирования шифрования следующего блока. Каждый блок шифротекста зависит не только от шифруемого блока открытого текста, но и от всех предыдущих блоков открытого текста.

В режиме сцепления блоков шифротекста перед шифрованием над открытым текстом и предыдущим блоком шифротекста выполняется операция XOR. Процесс шифрования в режиме CBC. Когда блок открытого текста зашифрован, полученный шифротекст сохраняется в регистре обратной связи. Следующий блок открытого текста перед шифрованием подвергается операции XOR с содержимым регистра обратной связи. Результат операции XOR используется как входные данные для следующего этапа процедуры шифрования. Полученный шифротекст снова сохраняется в регистре обратной связи, чтобы подвергнуться операции XOR вместе со следующим блоком открытого текста, и так до конца сообщения. Шифрование каждого блока зависит от всех предыдущих блоков.

Расшифрование выполняется в обратном порядке. Блок шифротекста расшифровывается обычным путем, но сохраняется в регистре обратной связи. Затем следующий блок расшифровывается и подвергается операции XOR с содержимым регистра обратной связи. Теперь следующий блок шифротекста сохраняется в регистре обратной связи и т.д. до конца сообщения.

На рисунке 7.60 показан режим CBC (IV – синхроросылка):

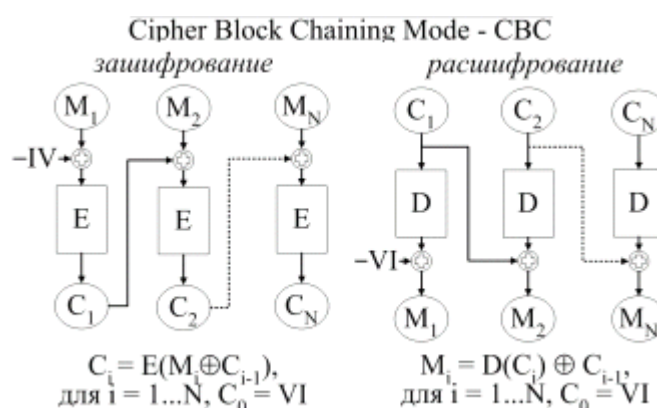


Рис. 7.60. Режим CBC

При шифровании в режиме CBC одинаковые блоки открытого текста превращаются в различающиеся друг от друга блоки шифротекста только в том случае, если различались какие-то предшествующие блоки открытого текста. Однако при шифровании двух 48 идентичных сообщений создается один и тот же шифротекст. Хуже того, два одинаково начинающихся сообщения будут шифроваться одинаково вплоть до первого различия.

Чтобы избежать этого, можно зашифровать в первом блоке какие-то произвольные данные. Этот блок случайных данных называют вектором инициализации (ВИ) (Initialization Vector, IV, русский термин - синхроросылка), инициализирующей переменной или начальным значением сцепления. Вектор ВИ не имеет какого-то смыслового значения, он используется только для того, чтобы сделать каждое сообщение уникальным. Когда получатель расшифровывает этот блок, он использует его только для заполнения регистра

обратной связи. В качестве вектора ВИ удобно использовать метку времени, либо какие-то случайные биты.

Если используется вектор инициализации, сообщения с идентичным открытым текстом после шифрования превращаются в сообщения с разными шифротекстами. Следовательно, злоумышленник не может попытаться повторить блок, и создание шифровальной книги затруднится. Хотя для каждого сообщения, шифруемого одним и тем же ключом, рекомендуется выбирать уникальный вектор ВИ, это требование необязательное. Вектор ВИ не обязательно хранить в секрете, его можно передавать открыто - вместе с шифротекстом.

Режим CFB (Ciphertext Feedback – обратная связь по шифротексту)

В режиме CBC начать шифрование до поступления полного блока данных невозможно. Для некоторых сетевых приложений это создает проблемы. Например, в защищенном сетевом окружении терминал должен иметь возможность передавать хосту каждый символ сразу после ввода. Если же данные нужно обрабатывать блоками в несколько байт, режим CBC просто не работает.

На рисунке 7.61 показан режим CFB (IV – синхропосылка):

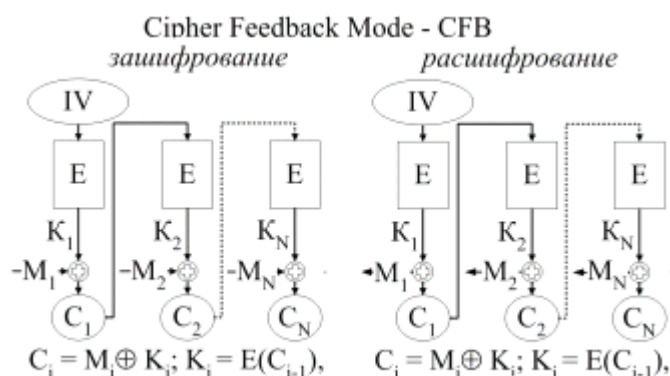


Рис. 7.61. Режим CFB

Как и в режиме CBC, первоначально очередь заполнена вектором инициализации ВИ. Очередь шифруется, затем выполняется операция XOR над n старшими (крайними левыми) битами результата и первым n -битовым символом открытого текста. В результате появляется первый n -битовый символ шифротекста. Теперь этот символ можно передать. Кроме того, полученные n битов попадают в очередь на место n младших битов, а все остальные биты сдвигаются на n позиций влево. Предыдущие n старших битов отбрасываются. Затем точно также шифруется следующие n битов открытого текста. Расшифрование выполняется в обратном порядке. Обе стороны - шифрующая и расшифровывающая - использует блочный алгоритм в режиме шифрования. Как и режим CBC, режим CFB сцепляет символы открытого текста с тем, чтобы шифротекст зависел от всего предыдущего открытого текста.

Для инициализации процесса шифрования в режиме CFB в качестве входного блока алгоритма можно использовать вектор инициализации ВИ. Как и в режиме CBC, хранить в тайне вектор ВИ не нужно. Однако вектор ВИ должен быть уникальным. (В отличие от режима CBC, где уникальность вектора ВИ необязательна, хотя и желательна). Если вектор ВИ в режиме CFB не уникален, криптоаналитик может восстановить соответствующий открытый текст. Вектор инициализации должен меняться в каждом сообщении. Например, вектором ВИ может служить порядковый номер, возрастающий в каждом новом сообщении и не повторяющийся все время жизни ключа. Если данные шифруются с целью последующего хранения, вектор ВИ может быть функцией индекса, используемого для поиска данных.

Режим OFB (Output Feedback – режим обратной связи по выходу)

Режим OFB представляет собой метод использования блочного шифра в качестве синхронного потокового шифра. Этот режим подобен режиму CFB.

Блочный алгоритм работает в режиме шифрования как на шифрующей, так и на расшифровывающей сторонах. Такую обратную связь иногда называют внутренней, поскольку механизм обратной связи не зависит ни от потока открытого текста, ни от потока шифротекста.

К достоинствам режима OFB относится то, что большую часть работы можно выполнить оффлайново, даже когда открытого текста сообщения еще вовсе не существует. Когда, наконец, сообщение поступает, для создания шифротекста над сообщением и выходом алгоритма необходимо выполнить операцию XOR.

В сдвиговый регистр OFB сначала надо загрузить вектор ВИ. Вектор должен быть уникальным, но сохранять его в тайне не обязательно.

На рисунке 7.62 показан режим OFB (IV – синхроросылка):

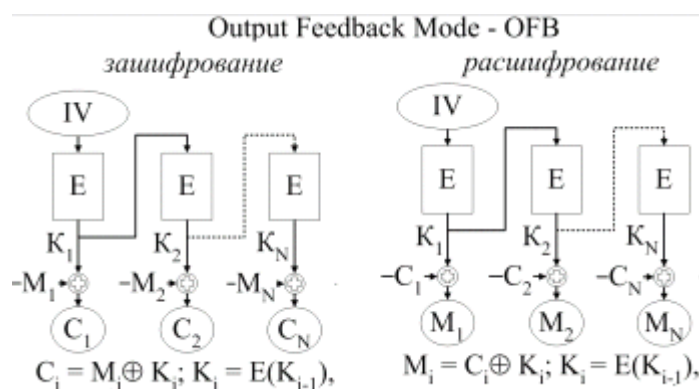


Рис. 7.62. Режим OFB

Анализ режима OFB показывает, что OFB целесообразно использовать только если разрядность обратной связи совпадает с размером блока.

В режиме OFB над гаммой и текстом выполняется операция XOR. Эта гамма, в конце концов, повторяется. Существенно, чтобы она не повторялась для одного и того же ключа, в противном случае секретность не обеспечивается ничем.

Порядок выполнения работы

Ознакомится с теорией по стандарту AES (из чего состоит раунд шифрования, как работают раундовые преобразования, с необходимым математическим аппаратом).

Взять у преподавателя задание, состоящие из:

- уникального ключа для шифрования/дешифрования текстовых сообщений
- набора файлов разного размера для исследования стандарта
- файла для исследования режимов шифрования
- вектора инициализации
- файла для проведения атаки «Квадрат»

Исследование стандарт шифрования

1. Открыть окно «Настройки» (AES>Настройки>), установить переключатель в положение «Свой ключ» и ввести посимвольное значение ключа (32 символа, 16 ячеек). Нажать «ОК».
2. В главном окне поставить переключатель в положение ECB.
3. Открыть файл из папки программы (Файл >Открыть)
4. Зашифровать и расшифровать его. (AES > Зашифровать (Расшифровать)).
5. Прodelать пункты 3.2-3.4 для оставшихся трех файлов.
6. По полученным результатам заполнить таблицу

Размер файла				
Время зашифрования				
Время расшифрования				

7. Построить графики зависимости шифрования \ расшифрования от времени.
8. По полученным данным оценить примерное время взлома стандарта AES с помощью всего перебора ключей. Построить график.

Пример:

На расшифрование 5кВ (5120 байт) текста на Athlon 2600+ тратится 9с. Поэтому один блок (16 байт) программа расшифровывает за $9/(5120/16) = (9*16)/5120 = 28$ мс. Таким образом, на перебор 1 ключа, а это 128 бит, потребуется 28 мс.

За 1 секунду программа способна перебрать 4571 бит или примерно 35 вариантов ключей.

За год непрерывной работы программа переберет $(35*60*60*24*365) = 1\,103\,760\,000$ ключей.

Если поделить 2^{128} на $1\,103\,760\,000$, то получим $3.08*10^{29}$ лет непрерывной работы программы.

Исследование быстродействия различных режимов шифрования

1. Поставить переключатель в режим CBC.
2. Открыть файл из папки программы (Файл >Открыть)
3. Зашифровать и расшифровать его (AES > Зашифровать (Расшифровать))
4. Повторить пункт 4.2-4.3 для оставшихся трех файлов (10, 15, 30 кБ)
5. Поставить переключатель в режим CFB.
6. Повторить пункты 4.2 – 4.3.
7. Поставить переключатель в режим OFB.
8. Повторить пункты 4.2 – 4.3
9. По полученным результатам заполнить таблицу:

Размер файла				
ЕСВ время заш/расш				
СВС время заш/расш				
CFB время заш/расш				
OFB время заш/расш				

10. Построить на одном графике зависимость шифрования \ расшифрования от времени.

11. Для каждого режима загрузить еще раз файл (любой). Зашифровать его. В текстовом поле 2 изменить какой-нибудь символ. Нажать AES > Расшифровать.

Исследование свойств различных режимов шифрования

Загрузить файл «Свойства режимов».

Поочередно зашифровать текстовое сообщение во всех четырех режимах.

Сравнить полученные шифротексты и сделать выводы.

Исследование пораундовой работы алгоритма

1. Открыть любой из предоставленных для работы файлов
2. Включить функцию логирования (AES>Настройки>Включить логирование>Да).
3. В главном окне поставить переключатель в положение ECB.
4. Зашифровать и расшифровать.
5. Открыть с помощью текстового редактора файлы Encrypt.txt и Decrypt.txt из папки программы
6. Проанализировать полученный результат.

Знакомство с атакой «Квадрат»

1. В главном окне в текстовое поле 1 загрузить файл «Атака Квадрат» (Файл>Открыть)
2. Открыть окно Атака «Квадрат» (AES > Квадрат)
3. С помощью переключателя “Номер раунда” установить 4 раундовый ключ.
4. Нажать кнопку “вычислить”.
5. Поместить текст из поля 1 в поле 2.
6. Нажать кнопку найти ключ.
7. Сравнить значения текстовых полей “Ваш ключ” и значения текстовых полей “Реальный ключ”.
8. Сделать выводы.

Содержание отчета

1. Цель работы.
2. Описание действий по каждому пункту.
3. Результаты проделанной работы (таблицы, графики, расчетная часть).
4. Выводы.

Контрольные вопросы

1. Что такое симметричное шифрование?
2. Что представляет собой стандарт AES (длина ключа, размер входного блока)?
3. Какой алгоритм выбран в качестве стандарта AES?
4. Что собой представляет архитектура данного стандарта?
5. Из чего состоит один раунд?
6. Сколько раундов шифрования предусмотрено стандартом?
7. Что быстрее шифрование или дешифрование? Почему?
8. Какие режимы шифрования применяются в стандарте AES?
9. Какие режимы быстрее при дешифровании? Почему?
10. Какие режимы лучше восстанавливают зашифрованную информацию при ошибке в одном символе? Почему?
11. С какой целью используется синхропосылка или вектор инициализации?
12. Что представляет собой атака Квадрат? Какие ее особенности?

1. Ознакомится с теорией по стандарту AES (из чего состоит раунд шифрования, как работают раундовые преобразования, с необходимым математическим аппаратом).

2. Взять у преподавателя задание, состоящие из:

- уникального ключа для шифрования/дешифрования текстовых сообщений
- набора файлов разного размера для исследования стандарта
- файла для исследования режимов шифрования
- вектора инициализации
- файла для проведения атаки «Квадрат»

3. Исследование стандарт шифрования

3.1. Открыть окно «Настройки» (AES>Настройки>), установить переключатель в положение «Свой ключ» и ввести посимвольное значение ключа (32 символа, 16 ячеек). Нажать «ОК».

3.2. В главном окне поставить переключатель в положение ECB.

3.3. Открыть файл из папки программы (Файл >Открыть)

3.4. Зашифровать и расшифровать его. (AES > Зашифровать (Расшифровать)).

3.5. Прodelать пункты 3.2-3.4 для оставшихся трех файлов.

3.5. По полученным результатам заполнить таблицу

Размер файла				
Время зашифрования				
Время расшифрования				

3.6. Построить графики зависимости шифрования \ расшифрования от времени.

3.7. По полученным данным оценить примерное время взлома стандарта AES с помощью всего перебора ключей. Построить график.

Пример:

На расшифрование 5кВ (5120 байт) текста на Athlon 2600+ тратится 9с. Поэтому один блок (16 байт) программа расшифровывает за $9/(5120/16) = (9*16)/5120 = 28$ мс. Таким образом, на перебор 1 ключа, а это 128 бит, потребуется 28 мс.

За 1 секунду программа способна перебрать 4571 бит или примерно 35 вариантов ключей.

За год непрерывной работы программа переберет $(35*60*60*24*365) = 1\,103\,760\,000$ ключей.

Если поделить 2^{128} на $1\,103\,760\,000$, то получим $3.08*10^{29}$ лет непрерывной работы программы.

3.8. Сделать выводы

Исследование быстродействия различных режимов шифрования

1. Поставить переключатель в режим CBC.
2. Открыть файл из папки программы (Файл >Открыть)
3. Зашифровать и расшифровать его (AES > Зашифровать (Расшифровать))
4. Повторить пункт 4.2-4.3 для оставшихся трех файлов (10, 15, 30 кБ)
5. Поставить переключатель в режим CFB.
6. Повторить пункты 4.2 – 4.3.
7. Поставить переключатель в режим OFB.
8. Повторить пункты 4.2 – 4.3
9. По полученным результатам заполнить таблицу:

Размер файла				
ECB время заш/расш				

СВС заш/расш	время				
СFB заш/расш	время				
OFB заш/расш	время				

10. Построить на одном графике зависимость шифрования \ расшифрования от времени.

11. Для каждого режима загрузить еще раз файл (любой). Зашифровать его. В текстовом поле 2 изменить какой-нибудь символ. Нажать AES > Расшифровать.

12. Сделать выводы.

Исследование свойств различных режимов шифрования

Загрузить файл «Свойства режимов».

Поочередно зашифровать текстовое сообщение во всех четырех режимах.

Сравнить полученные шифротексты и сделать выводы.

Исследование пораундовой работы алгоритма

1 Открыть любой из предоставленных для работы файлов

2. Включить функцию логирования (AES>Настройки>Включить логирование>Да).

3. В главном окне поставить переключатель в положение ECB.

4. Зашифровать и расшифровать.

5. Открыть с помощью текстового редактора файлы Encrypt.txt и Decrypt.txt из папки программы

6. Проанализировать полученный результат.

Знакомство с атакой «Квадрат»

1. В главном окне в текстовое поле 1 загрузить файл «Атака Квадрат» (Файл>Открыть)

2. Открыть окно Атака «Квадрат» (AES > Квадрат)

3. С помощью переключателя “Номер раунда” установить 4 раундовый ключ.

4. Нажать кнопку “вычислить”.

5. Поместить текст из поля 1 в поле 2.

6. Нажать кнопку найти ключ.

7. Сравнить значения текстовых полей “Ваш ключ” и значения текстовых полей “Реальный ключ”.

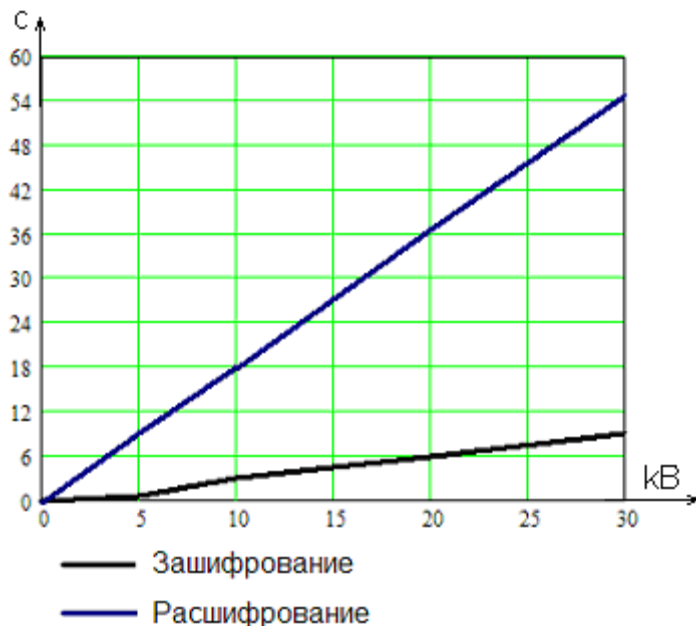
8. Сделать выводы.

Содержание отчета

1. Цель работы.
2. Описание действий по каждому пункту.
3. Результаты проделанной работы (таблицы, графики, расчетная часть).
4. Выводы.

Исследование стандарт шифрования

Размер файла	5kB	10kB	20kB	30kB
Время зашифрования, с	1	3	6	9
Время расшифрования, с	9	18	37	55



Зависимости размера файла от времени его шифрования \ расшифрования при проведении эксперимента на Athlon 2600+

Исследование быстродействия различных режимов шифрования

Размер файла	5kB	10kB	20kB	30kB
ECB время зашифрования	1	3	6	9
ECB время расшифрования	9	18	37	55
CBC время зашифрования	1	3	6	9
CBC время расшифрования	8	17	37	56
CFB время зашифрования	1	3	5	9
CFB время расшифрования	2	3	6	10
OFB время зашифрования	1	3	6	9
OFB время расшифрования	2	3	7	9

Зашифруем разными режимами слово пример, затем в шифротексте изменим 1 символ, получим:

	Исходный текст	Исходный шифротекст	Измененный шифротекст	Результат
ECB	пример	63891BCDBE14D6E7 B827C0AF68EF541 9	63891BCDBE14D6E7 B827C0AF68EF541 0	ЛМЎ!ня«m;5J9 w
CBC	пример	000000000000000000 000000000000000006389 1BCDBE14D6E7B82 7C0AF68EF5419	100000000000000000 000000000000000006389 1BCDBE14D6E7B82 7C0AF68EF5419	япример
CFB	пример	000000000000000000 0000000000000000 0 8919 A3380A7A2D3A894D	000000000000000000 0000000000000000 6 8919 A3380A7A2D3A894	ЗЮSŷ”[ж<Яс-(s

		FB58CB352A2F	D	
			FB58CB352A2F	
OFB	пример	000000000000000000 00000000000000008919 A3380A7A2D3A894D FB58CB352A2F	D D	μZi'alj:jOŃC& Ú
			FB58CB352A2F	

Исследование свойств различных режимов шифрования

Файл «Свойства режимов» содержит несколько раз повторяющееся слово. Например, *задание задание задание задание задание задание задание задание задание*.

Каждый режим при этом покажет свои особенности:

1. ECB

задание задание задание задание задание задание задание задание задание

D3D5451C8296E7E653A82652AED4E56A

D3D5451C8296E7E653A82652AED4E56A

D3D5451C8296E7E653A82652AED4E56A

A8F11E11A02FF140F0936291695512CA

2. CBC

00000000000000000000000000000000

D3D5451C8296E7E653A82652AED4E56A

F84E11957A1A02F67B97ED2A9DE04DD2

00D20F40BEE5D0C7EEF499531BEF8B8C

9BB124631613C483AE57E40BCD63C2AC

3. CFB

00000000000000000000000000000000

8109AF340262C91B6FAC1EB927DCCE0E

6D376DFF127D70059697E293142F5DA6

DBC4EAC4CA79A6139E944EB9B374DDAD

4128D7EF70BE71BB9F57B9D64DF4EC68

4. OFB

00000000000000000000000000000000
8109AF340262C91B6FAC1EB927DCCE0E
107559AABF0A7BF7F433F71ACD01689C
46EC128DE2351660B490505F601D5A93
A02767750D304B01733AEE7046216D8B

В данном примере синхропосыла равна 00000000000000000000000000000000

Исследование пораундовой работы алгоритма

Используем слово «пример»

Зашифрование:

Зашифрование

Входной блок EFF0E8ECE5F001010101010101010101

Раундовый ключ: 00000000000000000000000000000000

Состояние после сложения с ключом: EFF0E8ECE5F00101010101010101010101

Раунд №1

Состояние после SubBytes: DF8C9BCED98C7C7C7C7C7C7C7C7C7C

Состояние после ShiftRows: DF8C7C7CD97C7CCE7C7C9B7C7C8C7C7C

Состояние после MixColumns: 2A242F729F6B14F79B4EA99B77878C8C

Раундовый ключ: 62636363626363636263636362636363

Состояние после сложения с ключом: 48474C11FD087794F92DCAF815E4EFEF

Раунд №2

Состояние после SubBytes: 52A029825430F52299D874415969DFDF

Состояние после ShiftRows: 523074DF54D8DF829969292259A0F541

Состояние после MixColumns: 5F71F0178607B4E49912C4B4FD47CB3C

Раундовый ключ: 9B9898C9F9FBFBA9B9898C9F9FBFBA9

Состояние после сложения с ключом: C4E968DE7FFC4F4E028A5C7D04BC3096

Раунд №3

Состояние после SubBytes: 1C1E451DD2B0842F777E4AFF2650490

Состояние после ShiftRows: 1CB04A90D27E041D7765452FF21E84FF

Состояние после MixColumns: 292993E5243F832D2B5DE9E7A6A6E572

Раундовый ключ: 90973450696CCFFAF2F457330B0FAC99

Состояние после сложения с ключом: B9BEA7B54D534CD7D9A9BED4ADA949EB

Раунд №4

Состояние после SubBytes: 56AE5CD5E3ED290E35D3AE4895D33BE9

Состояние после ShiftRows: 56EDAEE9E3D33BD535D35C0E95AE2948

Состояние после MixColumns: C797DC705DC6226756624CCCB9E1B1B3

Раундовый ключ: EE06DA7B876A1581759E42B27E91EE2B

Состояние после сложения с ключом: 2991060BDAAC37E623FC0E7EC7705F98

Раунд №5

Состояние после SubBytes: A5816F2B57919A8E26B0ABF3C651CF46

Состояние после ShiftRows: A591AB4657B0CF2B26516F8EC6819AF3

Состояние после MixColumns: 143CB342814D1FD05EBB2053669966B7

Раундовый ключ: 7F2E2B88F8443E098DDA7CBBF34B9290

Состояние после сложения с ключом: 6B1298CA790921D9D3615CE895D2F427

Раунд №6

Состояние после SubBytes: 7FC94674B601FD3566EF4A9B2AB5BFCC

Состояние после ShiftRows: 7F014ACCB6EFBF7466B546352AC9FD9B

Состояние после MixColumns: 7B6FA54996DDA0797BE800337224B467

Раундовый ключ: EC614B851425758C99FF09376AB49BA7

Состояние после сложения с ключом: 970EEEC82F8D5F5E217090418902FC0

Раунд №7

Состояние после SubBytes: 88AB284B134103E698F001F2AD6015BA

Состояние после ShiftRows: 884101BA13F0154B986028E6ADAB03F2

Состояние после MixColumns: 73B31EAC739C144645C6992C56170DBB

Раундовый ключ: 217517873550620BACAF6B3CC61BF09B

Состояние после сложения с ключом: 52C6092B46CC764DE969F210900CFD20

Раунд №8

Состояние после SubBytes: 00B401F15A4B38E31EF989CA60FE54B7

Состояние после ShiftRows: 004B89B75AF954F11EFE01E360B438CA

Состояние после MixColumns: E3A180B701BE03BAC719DC00F591E1A3

Раундовый ключ: 0EF903333BA9613897060A04511DFA9F

Состояние после сложения с ключом: ED5883843A176282501FD604A48C1B3C

Раунд №9

Состояние после SubBytes: 556AEC5F80F0AA1353C0F6F24964AFEB

Состояние после ShiftRows: 55F0F6EB80C0AF5F5364EC13496AAAF2

Состояние после MixColumns: BC447434B0AEE44AF5A7C15B748A61E4

Раундовый ключ: B1D4D8E28A7DB9DA1D7BB3DE4C664941

Состояние после сложения с ключом: 0D90ACD63AD35D90E8DC728538EC28A5
 Раунд №10
 Состояние после SubBytes: D76091F680664C609B86409707CE3406
 Состояние после ShiftRows: D7664006808634F69BCE916007604C97
 Раундовый ключ: B4EF5BCB3E92E21123E951CF6F8F188E
 Состояние после сложения с ключом: 63891BCDBE14D6E7B827C0AF68EF5419

Расшифрование

Разшифрование

Входной блок 63891BCDBE14D6E7B827C0AF68EF5419
 Раундовый ключ: B4EF5BCB3E92E21123E951CF6F8F188E
 Состояние после сложения с ключом: D7664006808634F69BCE916007604C97
 Раунд №1
 Состояние после invShiftRows: D76091F680664C609B86409707CE3406
 Состояние после invSubBytes: 0D90ACD63AD35D90E8DC728538EC28A5
 Раундовый ключ: B1D4D8E28A7DB9DA1D7BB3DE4C664941
 Состояние после сложения с ключом: BC447434B0AEE44AF5A7C15B748A61E4
 Состояние после invMixColumns: 55F0F6EB80C0AF5F5364EC13496AAAF2
 Раунд №2
 Состояние после invShiftRows: 556AEC5F80F0AA1353C0F6F24964AFEB
 Состояние после invSubBytes: ED5883843A176282501FD604A48C1B3C
 Раундовый ключ: 0EF903333BA9613897060A04511DFA9F
 Состояние после сложения с ключом: E3A180B701BE03BAC719DC00F591E1A3
 Состояние после invMixColumns: 004B89B75AF954F11EFE01E360B438CA
 Раунд №3
 Состояние после invShiftRows: 00B401F15A4B38E31EF989CA60FE54B7
 Состояние после invSubBytes: 52C6092B46CC764DE969F210900CFD20
 Раундовый ключ: 217517873550620BACAF6B3CC61BF09B
 Состояние после сложения с ключом: 73B31EAC739C144645C6992C56170DBB
 Состояние после invMixColumns: 884101BA13F0154B986028E6ADAB03F2
 Раунд №4
 Состояние после invShiftRows: 88AB284B134103E698F001F2AD6015BA
 Состояние после invSubBytes: 970EEECC82F8D5F5E217090418902FC0
 Раундовый ключ: EC614B851425758C99FF09376AB49BA7

Состояние после сложения с ключом: 7B6FA54996DDA0797BE800337224B467

Состояние после invMixColumns: 7F014ACCB6EFBF7466B546352AC9FD9B

Раунд №5

Состояние после invShiftRows: 7FC94674B601FD3566EF4A9B2AB5BFCC

Состояние после invSubBytes: 6B1298CA790921D9D3615CE895D2F427

Раундовый ключ: 7F2E2B88F8443E098DDA7CBBF34B9290

Состояние после сложения с ключом: 143CB342814D1FD05EBB2053669966B7

Состояние после invMixColumns: A591AB4657B0CF2B26516F8EC6819AF3

Раунд №6

Состояние после invShiftRows: A5816F2B57919A8E26B0ABF3C651CF46

Состояние после invSubBytes: 2991060BDAAC37E623FC0E7EC7705F98

Раундовый ключ: EE06DA7B876A1581759E42B27E91EE2B

Состояние после сложения с ключом: C797DC705DC6226756624CCCB9E1B1B3

Состояние после invMixColumns: 56EDAEE9E3D33BD535D35C0E95AE2948

Раунд №7

Состояние после invShiftRows: 56AE5CD5E3ED290E35D3AE4895D33BE9

Состояние после invSubBytes: B9BEA7B54D534CD7D9A9BED4ADA949EB

Раундовый ключ: 90973450696CCFFAF2F457330B0FAC99

Состояние после сложения с ключом: 292993E5243F832D2B5DE9E7A6A6E572

Состояние после invMixColumns: 1CB04A90D27E041D7765452FF21E84FF

Раунд №8

Состояние после invShiftRows: 1C1E451DD2B0842F777E4AFF2650490

Состояние после invSubBytes: C4E968DE7FFC4F4E028A5C7D04BC3096

Раундовый ключ: 9B9898C9F9FBFBAA9B9898C9F9FBFBAA

Состояние после сложения с ключом: 5F71F0178607B4E49912C4B4FD47CB3C

Состояние после invMixColumns: 523074DF54D8DF829969292259A0F541

Раунд №9

Состояние после invShiftRows: 52A029825430F52299D874415969DFDF

Состояние после invSubBytes: 48474C11FD087794F92DCAF815E4EFEF

Раундовый ключ: 62636363626363636263636362636363

Состояние после сложения с ключом: 2A242F729F6B14F79B4EA99B77878C8C

Состояние после invMixColumns: DF8C7C7CD97C7CCE7C7C9B7C7C8C7C7C

Раунд №10

Состояние после invShiftRows: DF8C9BCED98C7C7C7C7C7C7C7C7C7C

Состояние после invSubBytes: EFF0E8ECE5F0010101010101010101

Раундовый ключ: 00000000000000000000000000000000

Состояние после сложения с ключом: EFF0E8ECE5F001010101010101010101

2.2.3. Практикум по поточным шифрам

Потоковый шифр RC4

Целью данной лабораторной работы является изучение алгоритма RC4, рассмотрение его слабых и сильных сторон, определение в каких продуктах и каким образом он используется, а так же знакомство с программной реализацией на языке C++, с использованием программного обеспечения Borland C++ Builder 6.

RC4 (англ. Rivest Cipher 4 или англ. Ron's Code, также известен как ARCFOUR или ARC4 (англ. Alleged RC4)) — это потоковый шифр, широко применяющийся в различных системах защиты информации в компьютерных сетях (например, в протоколах SSL и TLS, алгоритме безопасности беспроводных сетей WEP, для шифрования паролей).

Шифр разработан компанией RSA Security и для его использования требуется лицензия.

Алгоритм RC4 строится, как и любой потоковый шифр на основе параметризованного ключом генератора псевдослучайных битов с равномерным распределением. Длина ключа обычно составляет от 5 до 64 байт. Максимальная длина ключа 256 байт.

Основные преимущества шифра — высокая скорость работы и переменный размер ключа. RC4 довольно уязвим, если используются не случайные или связанные ключи, один ключевой поток используется дважды. Эти факторы, а также способ использования могут сделать криптосистему небезопасной (например, WEP).

Потоковый шифр RC4 был создан Роном Ривестом из RSA Security в 1987 году. Хотя официально сокращение обозначает Rivest Cipher 4, его часто считают сокращением от Ron's Code.

Шифр являлся коммерческой тайной, но в сентябре 1994 года его описание было анонимно отправлено в рассылку Cypherpunks. Вскоре описание RC4 было опубликовано в ньюс-группе sci.crypt. Именно оттуда исходный код попал на множество сайтов в сети Интернет. Опубликованный шифр давал те же шифротексты на выходе, какие давал подлинный RC4. По-видимому, данный текст был получен в результате анализа исполняемого кода. Опубликованный шифр совместим с имеющимися продуктами, использующими RC4, а некоторые участники телеконференции, имевшие, по их словам, доступ к исходному коду RC4, подтвердили идентичность алгоритмов при различиях в обозначениях и структуре программы.

Поскольку данный алгоритм известен, он более не является коммерческой тайной. Однако, название «RC4» является торговой маркой компании RSA. Поэтому иногда шифр называют «ARCFOUR» или «ARC4» (имея ввиду Alleged RC4 — предполагаемый RC4, поскольку RSA официально не опубликовала алгоритм), чтобы избежать возможных претензий со стороны владельца торговой марки.

Шифр RC4 применяется в некоторых широко распространённых стандартах и протоколах шифрования таких, как WEP, WPA и TLS.

Главными факторами, способствовавшими широкому применению RC4, были простота его аппаратной и программной реализации, а также высокая скорость работы алгоритма в обоих случаях.

В США длина ключа для использования внутри страны рекомендуется равной 128 битов, но соглашение, заключённое между Software Publishers Association (SPA) и правительством США даёт RC4 специальный статус, который означает, что разрешено экспортировать шифры длиной ключа до 40 бит. 56-битные ключи разрешено использовать заграничным отделениям американских компаний.

Описание алгоритма

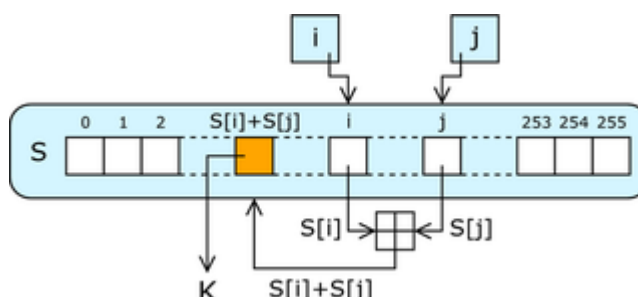


Рис. 7.62. Генератор ключевого потока RC4

Ядро алгоритма состоит из функции генерации ключевого потока. Эта функция генерирует последовательность битов (k_i), которая затем объединяется с открытым текстом (m_i) посредством суммирования по модулю два. Так получается шифrogramма (c_i):

$$c_i = m_i \oplus k_i.$$

Расшифровка заключается в регенерации этого ключевого потока (k_i) и сложении его и шифrogramмы (c_i) по модулю два. В силу свойств суммирования по модулю два на выходе мы получим исходный незашифрованный текст (m_i):

$$m_i = c_i \oplus k_i = (m_i \oplus k_i) \oplus k_i.$$

Другая главная часть алгоритма — функция инициализации, которая использует ключ переменной длины для создания начального состояния генератора ключевого потока.

RC4 — фактически класс алгоритмов, определяемых размером его блока. Этот параметр n является размером слова для алгоритма. Обычно, $n = 8$, но в целях анализа можно уменьшить его. Однако для повышения безопасности необходимо увеличить эту величину. Внутреннее состояние RC4 представляется в виде массива слов размером $2n$ и двух счетчиков, каждый размером в одно слово. Массив известен как S-блок, и далее будет обозначаться как S . Он всегда содержит перестановку $2n$ возможных значений слова. Два счетчика обозначены через i и j .

Алгоритм инициализации RC4 приведен ниже. Этот алгоритм также называется алгоритмом ключевого расписания (Key-Scheduling Algorithm or KSA). Этот алгоритм использует ключ, сохраненный в Key , и имеющий длину l байт. Инициализация начинается с заполнения массива S , далее этот массив перемешивается путем перестановок определяемых ключом. Так как только одно действие выполняется над S , то должно выполняться утверждение, что S всегда содержит все значения кодового слова.

Начальное заполнение массива:

for $i = 0$ to $2n - 1$

$S[i] = i$

Скремблирование:

$j = 0$

for $i = 0$ to $2n - 1$

$j = (j + S[i] + Key[i \bmod l]) \bmod 2n$

Перестановка ($S[i], S[j]$)

Генератор ключевого потока RC4 переставляет значения, хранящиеся в S , и каждый раз выбирает различное значение из S в качестве результата. В одном цикле RC4 определяется одно n -битное слово K из ключевого потока, которое в последующем суммируется с исходным текстом для получения зашифрованного текста. Эта часть алгоритма называется генератором псевдослучайной последовательности (Pseudo-Random Generation Algorithm or PRGA).

Инициализация:

$i = 0$

$j = 0$

Цикл генерации:

$i = (i + 1) \bmod 2n$

$j = (j + S[i]) \bmod 2n$

Перестановка ($S[i], S[j]$)

Результат: $K = S[(S[i] + S[j]) \bmod 2n]$

Безопасность

В отличие от современных шифров (таких, как в eSTREAM), RC4 не использует отдельной okazji (nonce) наряду с ключом. Это значит, что если один ключ должен использоваться в течение долгого времени для шифрования нескольких потоков, сама криптосистема, использующая RC4, должна комбинировать okaziu и долгосрочный ключ для получения потокового ключа для RC4. Один из возможных выходов — генерировать новый ключ для RC4 с помощью хэш-функции от долгосрочного ключа и okazji. Однако, многие приложения, использующие RC4, просто конкатенируют ключ и okaziu. Из-за этого и слабого расписания ключей, используемого в RC4, приложение может стать уязвимым.

Здесь будут рассмотрены некоторые атаки на шифр и методы защиты от них.

Манипуляция битами

Шифр RC4 крайне уязвим к манипуляции битами, если он не реализован верным образом. И поэтому он был признан устаревшим многими софтверными компаниями, такими как Microsoft. Например, в .NET Framework от Microsoft отсутствует реализация RC4.

Исследования Руза и восстановление ключа из перестановки

В 1995 году Андроу Руз (Andrew Roos) экспериментально пронаблюдал, что первый байт ключевого потока коррелирован с первыми тремя байтами ключа, а первые несколько байт перестановки после алгоритма расписания ключей (KSA) коррелированы с некоторой линейной комбинацией байт ключа. Эти смещения не были доказаны до 2007 года, когда Пол, Рафи и Мэйтрэ доказали коррелированность ключа и ключевого потока. Также Пол и Мэйтрэ доказали коррелированность перестановки и ключа. Последняя работа также использует коррелированность ключа и перестановки для того, чтобы создать первый алгоритм полного восстановления ключа из последней перестановки после KSA, не делая предположений о ключе и векторе инициализации (V or Initial Vector). Этот алгоритм имеет постоянную вероятность успеха в зависимости от времени, которая соответствует квадратному корню из сложности полного перебора. Позднее было сделано много работ о восстановлении ключа из внутреннего состояния RC4.

Атака Флурера, Мантина и Шамира (ФМШ)

В 2001 году, Флурер, Мантин и Шамир опубликовали работу об уязвимости ключевого расписания RC4. Они показали, что среди всех возможных ключей, первые несколько байт ключевого потока являются совсем неслучайными. Из этих байт можно с высокой вероятностью получить информацию о используемом шифром ключе. И если долговременный ключ и оказия (nonce) просто конкатенируются для создания ключа шифра RC4, то этот долговременный ключ может быть получен с помощью анализа достаточно большого количества сообщений, зашифрованных с использованием данного ключа. Эта уязвимость и некоторые связанные с ней эффекты были использованы при взломе шифрования WEP в беспроводных сетях стандарта IEEE 802.11. Это показало необходимость скорейшей замены WEP, что повлекло за собой разработку нового стандарта безопасности беспроводных сетей WPA.

Криптосистему можно сделать невосприимчивой к этой атаке, если отбрасывать начало ключевого потока. Таким образом, модифицированный алгоритм называется «RC4-drop[n]», где n — количество байт из начала ключевого потока, которые следует отбросить. Рекомендовано использовать $n = 768$, консервативная оценка составляет $n = 3072$.

Атака Кляйна

В 2005 году Андреас Кляйн представил анализ шифра RC4, в котором он указал на сильную коррелированность ключа и ключевого потока RC4. Кляйн проанализировал атаки на первом раунде (подобные атаке ФМШ), на втором раунде и возможные их улучшения. Он также предложил некоторые изменения алгоритма для усиления стойкости шифра. В частности, он утверждает, что если поменять направление цикла на обратное в алгоритме ключевого расписания, то можно сделать шифр более стойким к атакам типа ФМШ.

Комбинаторная проблема

В 2001 году Ади Шамир и Ицхак Мантин первыми поставили комбинаторную проблему, связанную с количеством всевозможных входных и выходных данных шифра RC4. Если из всевозможных 256 элементов внутреннего состояния шифра известно x элементов из состояния ($x \leq 256$), то, если предположить, что остальные элементы нулевые, максимальное

количество элементов, которые могут быть получены детерминированным алгоритмом за следующие 256 раундов также равно x . В 2004 году это предположение было доказано Сорадюти Полом (Souradyuti Paul) и Бартом Прэнилом (Bart Preneel).

Программная реализация

Работа многих поточных шифров основана на линейных регистрах сдвига с обратной связью (LFSR). Это позволяет достичь высокой эффективности реализаций шифра в виде ИС. Но затрудняет программную реализацию таких шифров. Поскольку шифр RC4 не использует LFSR и основан на байтовых операциях, его удобно реализовывать программно. Типичная реализация выполняет от 8 до 16 машинных команд на каждый байт текста, поэтому программная реализация шифра должна работать очень быстро.

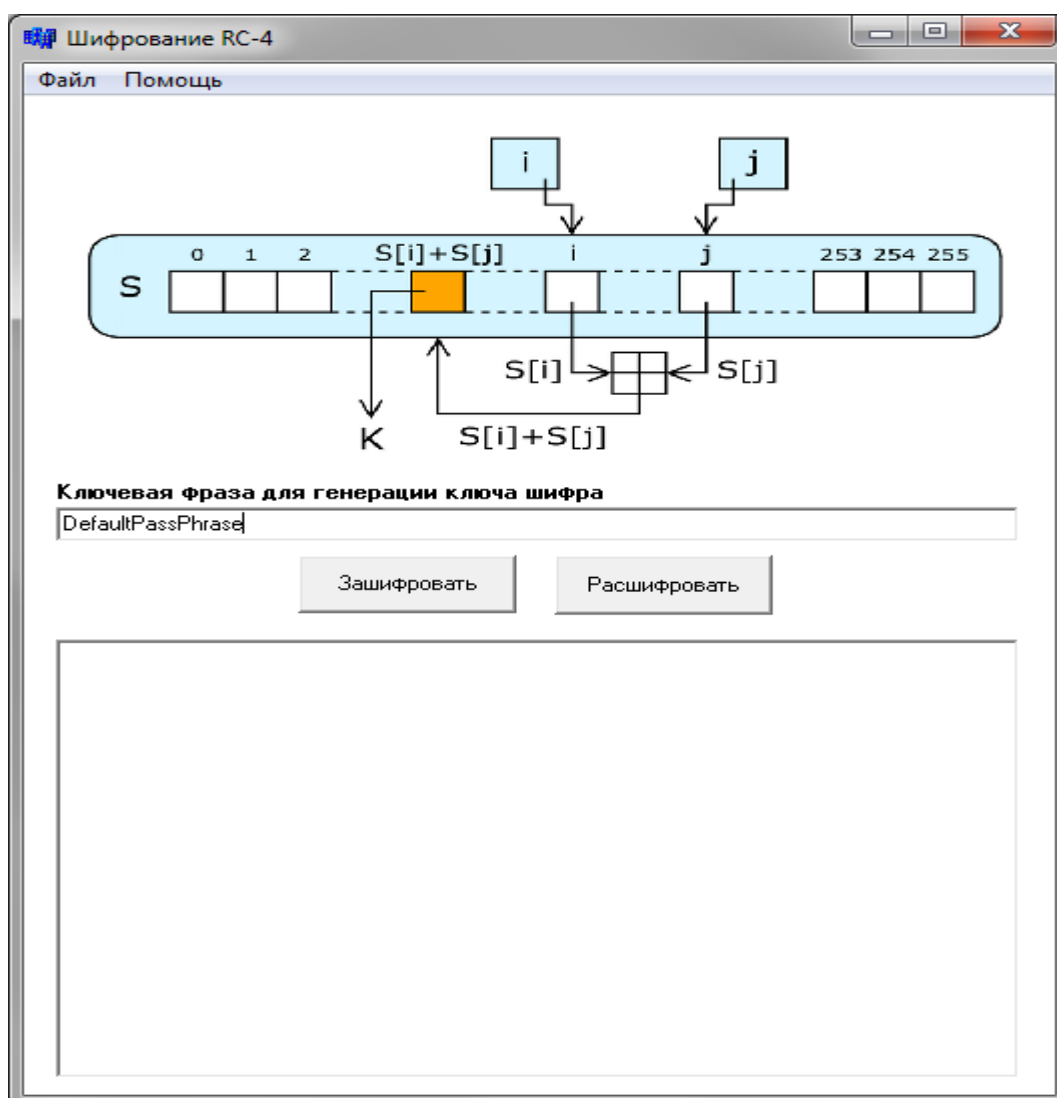


Рис. 7.63. Общий вид программы

Файл

Содержит:

Зашифровать – выбор файла, который необходимо зашифровать. При нажатии открывается интерфейс, позволяющий выбрать файл в проводнике.

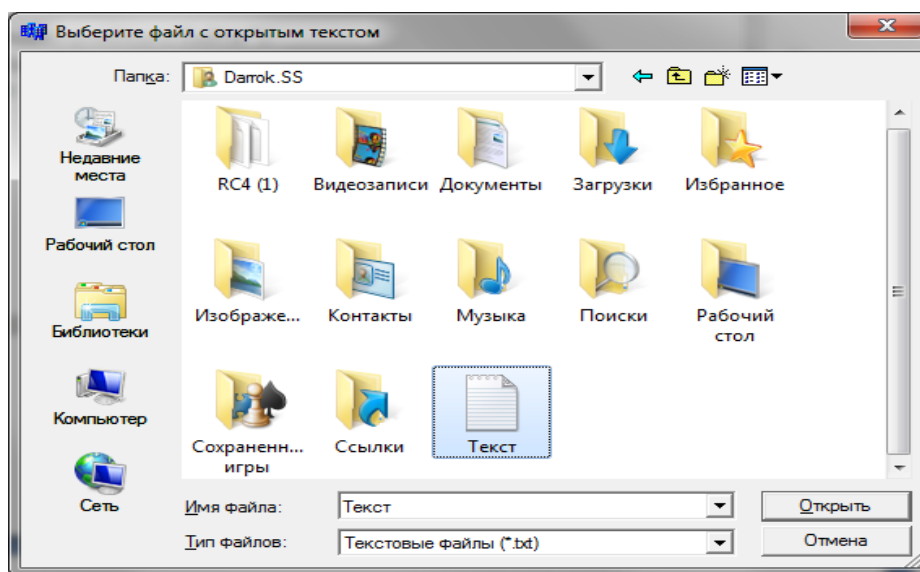


Рис. 7.64. Выбор текстового файла для шифрования

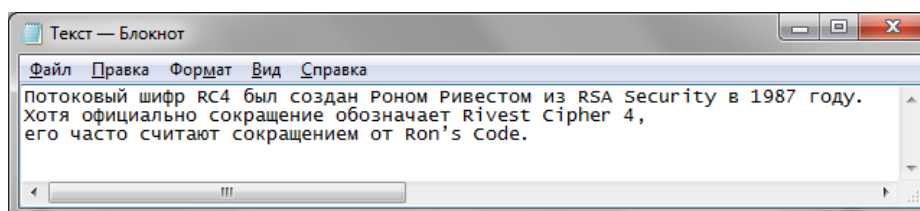


Рис. 7.65. Шифруемый текст

Расшифровать – выбор файла, который необходимо расшифровать. При нажатии открывается интерфейс, позволяющий выбрать файл в проводнике.

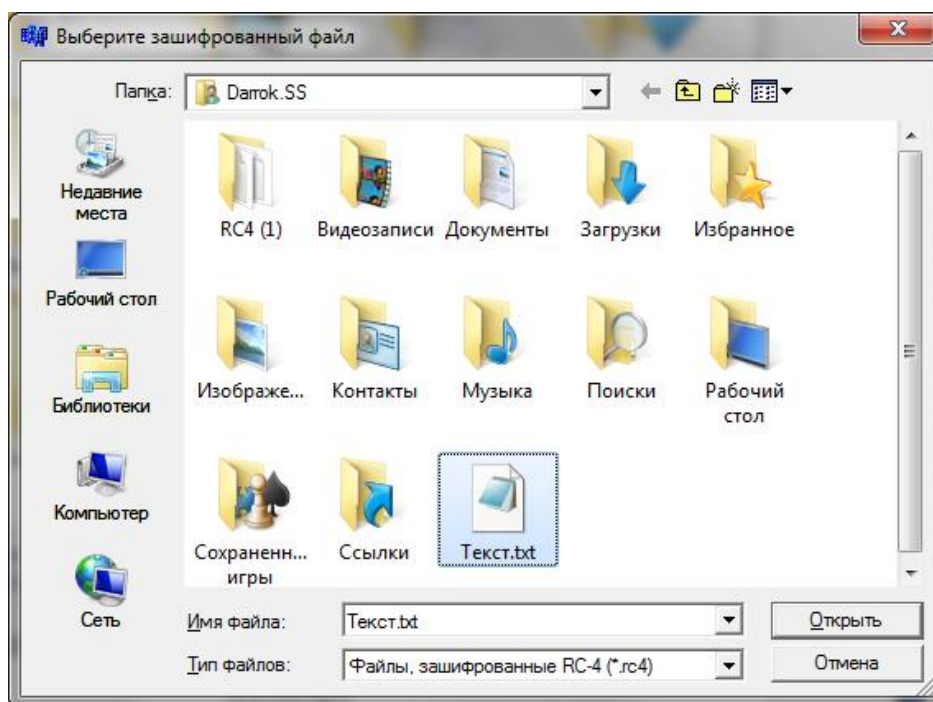


Рис. 7.66. Выбор текстового файла для расшифрования

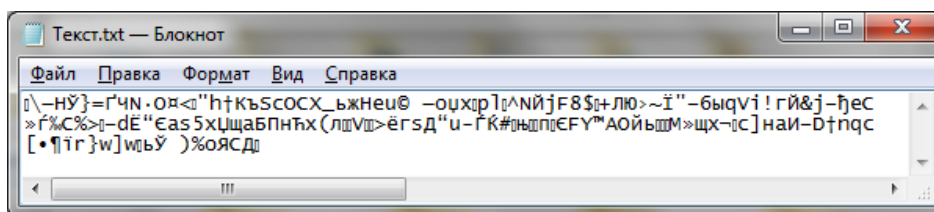


Рис. 7.67. Зашифрованный текст

Выход – выход из программы.

Помощь

Содержит:

Алгоритм – отображает теоретическую информацию об алгоритме, реализуемом в лабораторной работе.

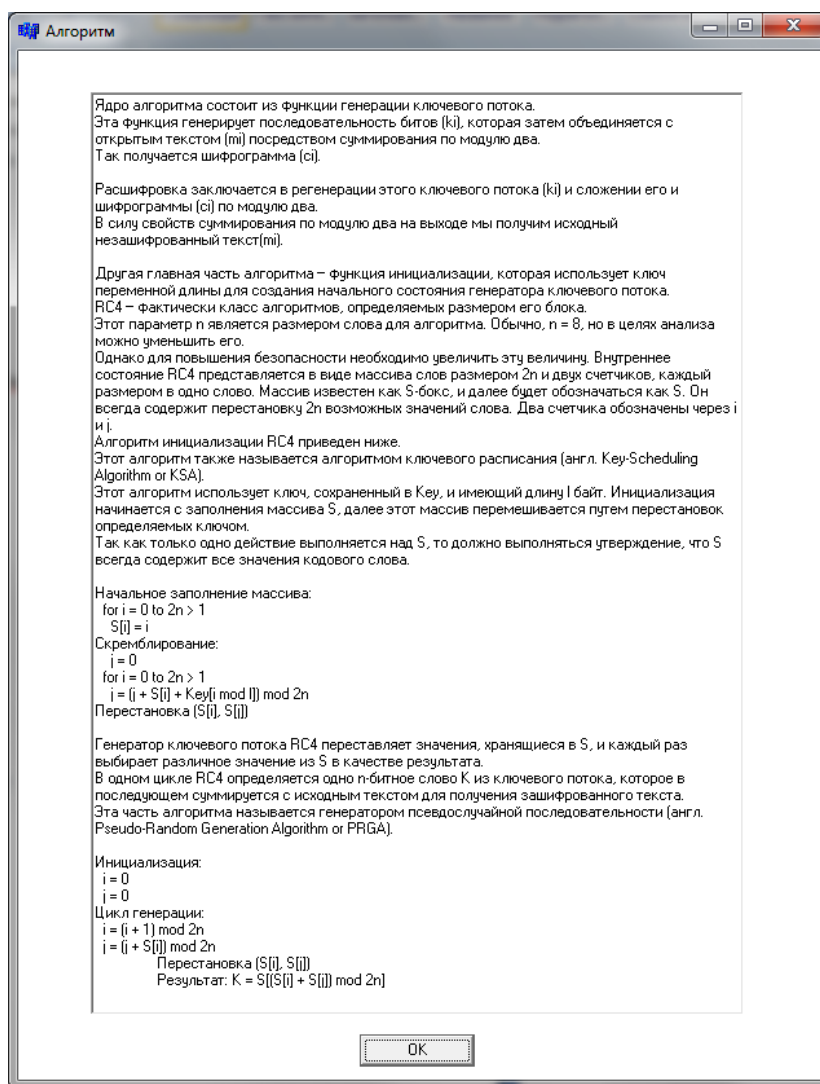


Рис. 7.68. Описание работы алгоритма

Автор – отображает информацию об авторе лабораторной работы и его научного руководителя.

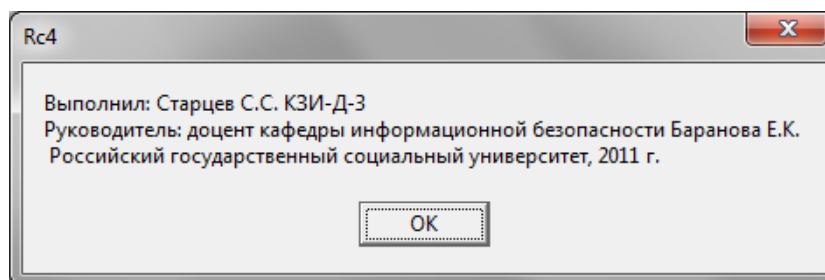


Рис. 7.69. Информация об авторе и руководителе

Поле “Ключ”

В это поле вводятся символы, являющиеся Ключом, который необходим для реализации шифрования методом, используемым в алгоритме.

Окно процесса шифрования

Отображает процесс шифрования и расшифрования, а именно: инициализацию матрицы, выполнение перестановок, считывание блока, количество обработанных байт и запись блока.

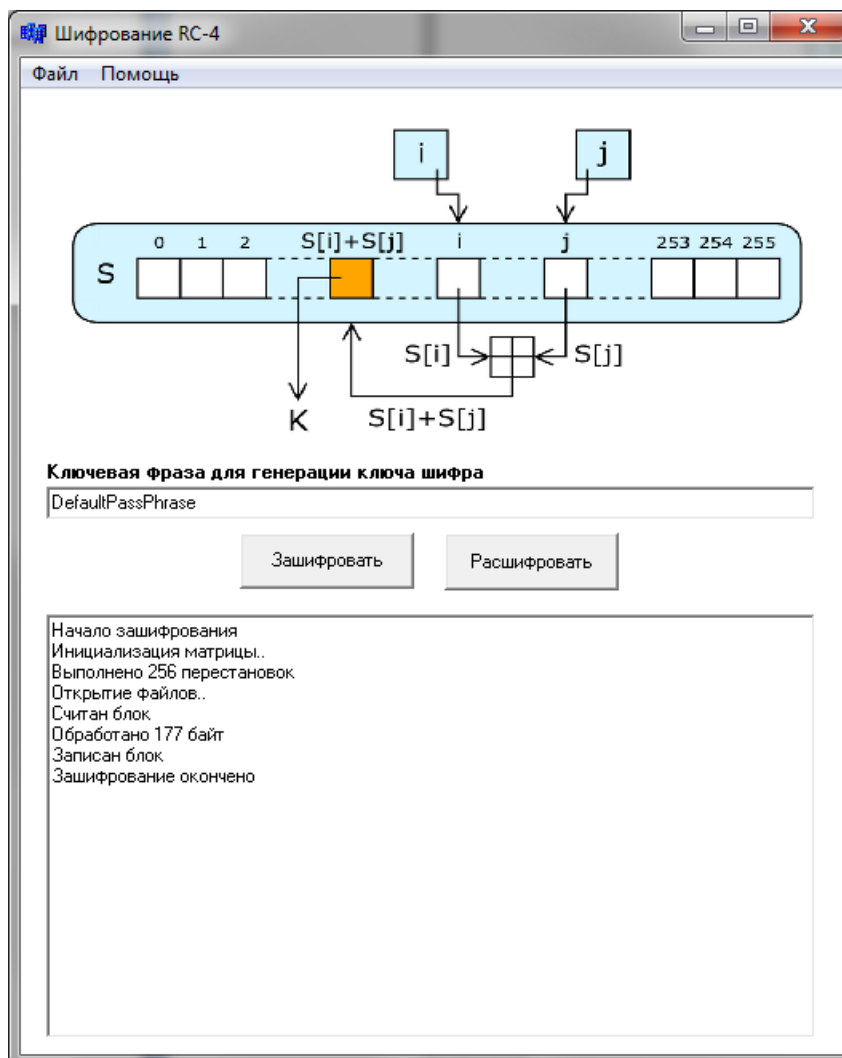


Рис. 7.70. Отображение процесса шифрования

Последовательность действий

Выбрать входной и выходной файлы, ввести ключ, нажать «Зашифровать/Расшифровать». Т.к. алгоритм симметричный, то процессы шифрования и расшифрования аналогичны.

Заключение

Была поставлена и, в ходе подготовки данной лабораторной работы, достигнута цель: разработка программы, реализующей алгоритм шифрования RC4, а так же рассмотрены и изучены особенности этого алгоритма.

Потоковый шифр RC4 был разработан Роном Ривестом в 1987 году. Этот шифр позволяет использовать ключи размером от 8 до 2048 бит (с шагом 8). В RC4 для зашифрования и расшифрования применяются одни и те же действия: генерируется гамма, которая накладывается на шифруемое сообщение путем сложения по модулю 2 (операция XOR).

RC4 применяется в таких продуктах, как Microsoft Office, Lotus Notes, Adobe Acrobat и др.

Алгоритм RC4 является собственностью компании RSA Data Security, Inc. Его описание никогда не было опубликовано и предоставлялось партнерам только после подписания соглашения о неразглашении. Однако в сентябре 1994 года в списке рассылки Cipherpunks (Шифропанки) кто-то анонимно опубликовал алгоритм шифрования, который на всех известных тестовых значениях совпадал с RC4. С тех пор сам алгоритм перестал быть секретом, но название RC4 остается торговой маркой. То есть, чтобы получить право заявлять, что в коммерческом программном продукте используется RC4, необходимо приобрести лицензию на этот алгоритм у RSA Data Security. А без лицензии можно утверждать лишь то, что "используется алгоритм, похожий на RC4 и совпадающий с ним на всем известном множестве тестов". Именно поэтому на языке ADA был реализован Alleged (предполагаемый) RC4.

8. ШИФРОВАНИЕ С ОТКРЫТЫМ КЛЮЧОМ

8.1. Теория шифров с открытым ключом

Асимметричные криптосистемы [33 -39]

Предпосылки появления асимметричных криптосистем

Появлению нового направления в криптологии - асимметричной криптографии с открытым ключом - способствовали две проблемы, которые не удавалось решить в рамках классической симметричной одноключевой криптографии.

Первая из этих проблем связана с *распространением секретных ключей*. Как передать участникам обмена информацией сменяемые секретные ключи, которые требуются им для осуществления этого обмена? В общем случае для передачи ключа опять же требуется использование какой-то криптосистемы, то есть задача в рамках симметричной криптографии неразрешима.

Вторая из этих проблем связана с распространением электронного документооборота. Возникла проблема обеспечения подлинности и авторства электронных документов. В

обычном, бумажном документообороте эта проблема решается с помощью подписи на бумаге. Подделать подпись человека на бумаге совсем не просто, а скопировать цепочку цифр на ЭВМ - несложная операция. Возникла проблема цифровой подписи, которая бы выполняла все те задачи, которые выполняет подпись, поставленная на документе рукой. Обе эти проблемы были успешно решены с помощью криптографии с открытыми ключами. В опубликованной в 1976 г. статье "Новые направления в криптографии" У.Диффи и М.Хеллман впервые показали, что секретная связь возможна без передачи секретного ключа между отправителем и получателем.

На основе результатов, полученных классической и современной алгеброй, были предложены *системы с открытым ключом*, называемые также *асимметричными криптосистемами*.

Суть их состоит в том, что каждым адресатом ИС генерируются два ключа, связанные между собой по определенному правилу. Один ключ объявляется *открытым*, а другой *закрытым*. Открытый ключ публикуется и доступен любому, кто желает послать сообщение

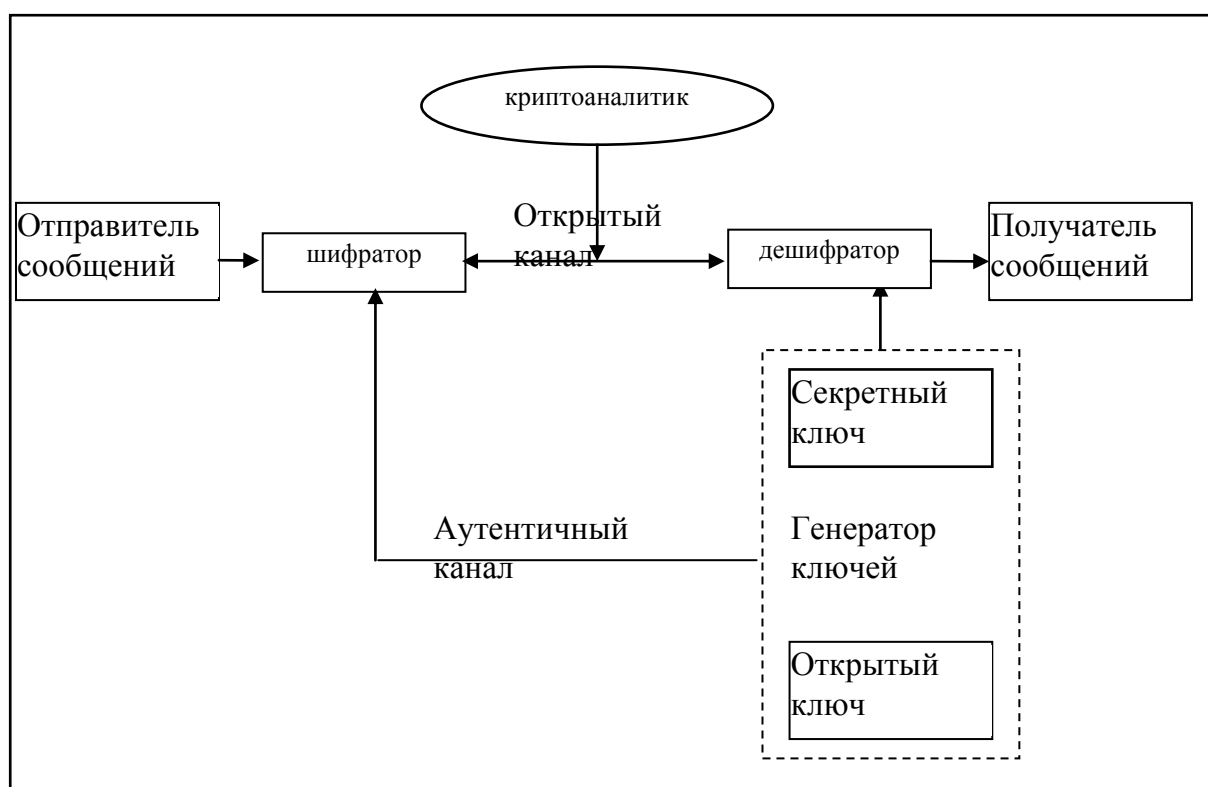


Рис. 8.1. Обобщенная схема асимметричной криптосистемы

адресату. Секретный ключ сохраняется в тайне. Если генератор ключей расположить на стороне получателя, то отпадает необходимость пересылки секретного ключа по каналу связи.

Исходный текст шифруется открытым ключом адресата и передается ему. Зашифрованный текст *в принципе не может быть расшифрован* тем же открытым ключом. Дешифрование сообщения возможно только с использованием закрытого ключа, который известен только самому адресату. Таким образом, не требуется секретный канал связи для передачи ключа, но необходимо обеспечить подлинность открытого ключа, так как его искажение или подмена не позволит расшифровать информацию на парном с ним закрытом ключе. Кроме того, замена злоумышленником законного открытого ключа на свой открытый ключ предоставляет ему полный доступ к шифруемой информации.

Обобщенная схема асимметричной крипосистемы

Ниже (рис.3.1) приведена *обобщенная схема асимметричной крипосистемы*.

Здесь для передачи ключа используется открытый канал связи, обеспечивающий аутентичность передаваемой информации.

Чтобы гарантировать надежную защиту информации, к системам с открытым ключом (СОК) предъявляются два важных и очевидных требования:

1. Преобразование исходного текста должно быть необратимым и исключать его восстановление на основе открытого ключа.
2. Определение закрытого ключа на основе открытого также должно быть невозможным на современном технологическом уровне. При этом желательна точная нижняя оценка сложности раскрытия шифра.

Алгоритмы шифрования с открытым ключом получили широкое распространение в современных информационных системах. Их используют в следующих основных направлениях:

1. Как самостоятельные средства защиты передаваемых и хранимых данных.
2. Как средства для распределения криптографических ключей. Алгоритмы асимметричных крипосистем более трудоемки, чем традиционные крипосистемы. Поэтому часто на практике имеет смысл с помощью асимметричных крипосистем распределять ключи, объем которых незначителен. А потом с помощью менее трудоемких симметричных алгоритмов осуществлять обмен большими информационными потоками.
3. Как средства аутентификации пользователей информационных систем, в том числе для решения проблемы электронной подписи.
4. Как средства для построения сложных криптографических протоколов для решения различных задач защиты информации в современных информационных системах.

Алгебраическая обобщенная модель шифра

Ранее мы рассматривали алгебраическую модель шифра К.Шеннона как трехосновную универсальную алгебру $A=(M,K,C,E)$, где

M - множество открытых текстов,

K - множество ключей,

C - множество криптограмм и

E – инъективная (взаимно однозначная) функция шифрования:

$$E_k: M \times K \rightarrow C \quad E_k(m)=c, \text{ где } m \in M, k \in K, c \in C.$$

Для того, чтобы эта модель могла быть применима к асимметричным криптографическим системам, необходимо ее расширение. Основная концептуальная идея построения такой модели, естественна и очевидна. Она состоит в отдельном описании моделей двух шифров: шифра шифрования и шифра расшифрования, совокупность которых и составляет обобщенную алгебраическую модель шифра.

Шифром зашифрования (алгеброй зашифрования) назовем алгебру

$$A_{\text{ш}}=(M, M_c, K_{\text{ш}}, C, C_c, E), \quad \text{где}$$

✧ множество $M_c \subseteq M$ трактуется как подмножество всех содержательных текстов из множества «открытых текстов» M .

✧ функция шифрования E осуществляет отображение $M \times K_{\text{ш}}$ на C :

$$E: M \times K_{\text{ш}} \rightarrow C, \quad E_k(m)=c,$$

то есть является сюръективной, причем $\forall k \in K_{\text{ш}}$ отображение $E_k(m)$ инъективно (образы двух различных элементов различны), а множество C_c состоит из шифрограмм, которые могут быть получены в результате шифрования содержательных текстов: $C_c = E(M_c \times K_{\text{ш}})$, то есть результатов шифрования тех открытых текстов m , для которых определено значение $E_k(m)$ для всех ключей шифра $k \in K_{\text{ш}}$.

Введение подмножества $M_c \subseteq M$ как множества содержательных текстов позволяет корректно вводить критерии на содержательные тексты.

Таким образом, шифр зашифрования есть некоторое уточнение модели шифра Шеннона $A=(M, K_{\text{ш}}, C, E)$.

Шифром расшифрования (алгеброй расшифрования) для $A_{\text{ш}}$ назовем алгебру

$$A_p=(M_p, K_p, C_p, D), \text{ где}$$

➤ $C \subseteq C_p$ и введение множества C_p — шифртекстов «правильных» сообщений (а точнее, $C_p \setminus C$ — множества искаженных шифртекстов) обеспечивает возможность

описания реакции приемной стороны на поступление искаженного шифрованного сообщения $c_p \notin C$;

➤ $M \subseteq M_p$, и введение множества $M_p \setminus M$ обеспечивает возможность описания результата расшифрования приемной стороной искаженного шифрованного сообщения $c_p \notin C$;

➤ функция дешифрования D - сюръективное отображение

$$D: C_p \times K_p \rightarrow M_p, D_k(c) = m,$$

для которого выполняются следующие условия:

1) существует биекция $f: K_{ш} \rightarrow K_p$;

2) для любых $m \in M, k \in K_{ш}$ из условия $E_k(m) = c$ вытекает

$$D(c, f(k)) = m.$$

При отсутствии искажений в канале связи функция расшифрования D полностью определена на всем множестве $S \times K_p$.

Отметим, что в определении шифра расшифрования не содержится требований инъективности функции f по переменной $k \in K_p$.

Алгебраической обобщенной моделью шифра назовем тройку

$$(A_{ш}, A_p, f).$$

К положительным свойствам этой модели относится возможность моделирования шифров как с симметричным, так и с асимметричным ключом.

При этом учитываются следующие соображения:

- ключ $k_{ш} \in K_{ш}$ несекретен, а ключ $k_p = f(k_{ш}) \in K_p$ является секретным;
- определение значения k связано с решением сложных проблем;
- синтез пар ключей $(k_{ш}, k_p)$ проводится достаточно просто.

Заметим, что здесь проявляется возможность классификации шифров по параметру сложности вычисления значения $f(k_{ш})$ ключа расшифрования, что определяет основной параметр криптографической стойкости шифров с асимметричным ключом.

Односторонние функции

Концепция асимметричных криптографических систем с открытым ключом основана на применении однонаправленных или односторонних функций. Последнее название было дано по ассоциации с односторонним движением, когда легко проехать в одну сторону и нельзя в другую. При этом в криптографии, как в жизни, «нельзя» не означает «невозможно ни при каких условиях», но говорит о том, что это сопряжено с серьезными трудностями.

Определение. Функция $f: X \rightarrow Y$ называется *односторонней (one-way function)*, если существует эффективный алгоритм для вычисления $f(x) \forall x$, но не существует эффективного алгоритма для вычисления хотя бы одного элемента прообраза $f^{-1}(y)$.

Никто не знает, существуют ли вообще односторонние функции. Основным критерием отнесения функции f к классу односторонних или необратимых является отсутствие эффективных с *вычислительной точки зрения* алгоритмов обратного преобразования $Y \rightarrow X$. В криптографии под *необратимостью* понимается не теоретическая необратимость функции, а практическая невозможность вычислить обратное значение, используя современные вычислительные средства за заданный интервал времени. Таким образом, проблемы построения односторонних функций связаны с теоретико-вероятностной сложностью алгоритмов и алгоритмическими вопросами теории чисел.

Множество классов необратимых функций порождает все разнообразие систем с открытым ключом. Большинство предлагаемых сегодня криптосистем с открытым ключом опираются на один из следующих типов необратимых преобразований:

1. Разложение больших целых чисел на простые множители.
2. Вычисление логарифма в конечном поле.
3. Вычисление корней алгебраических уравнений.

Факторизация

В качестве первого примера однонаправленной функции рассмотрим целочисленное умножение. Прямая задача — вычисление произведения двух очень больших целых чисел p и q , т.е. нахождение значения $n = p \cdot q$, является относительно несложной задачей.

Обратная задача, называемая задачей факторизации, — разложение на множители большого целого числа, т.е. нахождение делителей p и q большого целого числа

$$n = p \cdot q,$$

является практически неразрешимой задачей при достаточно больших значениях n . По современным оценкам теории чисел при целом $n \approx 2^{664}$ и $p \approx q$ для разложения числа n потребуется около 10^{23} операций, т.е. задача практически неразрешима на современных ЭВМ.

Если простые сомножители имеют специальный вид, известны более эффективные алгоритмы факторизации. Речь идет о сомножителях p , таких, у которых величины $p-1$ или $p+1$ являются «гладкими», т.е. имеют только малые простые делители.

Однако с появлением алгоритма факторизации с использованием эллиптических кривых класс чисел, допускающих быструю факторизацию, расширился и простые критерии проверки принадлежности данному классу утратили свою значимость. Поэтому, как правило,

единственным разумным критерием может служить размер простых множителей, поскольку с увеличением размера уменьшается вероятность выбрать число специального вида.

Дискретный логарифм

Другой пример однонаправленной функции — это модульная экспонента с фиксированными основанием и модулем. Пусть a и n — целые числа, такие, что $1 \leq a \leq n$. Тогда модульная экспонента с основанием a по модулю n представляет собой функцию

$$y = a^x \bmod n,$$

где x — целое число. Естественно записать $x = \log_a(y)$.

Задачу обращения этой функции в множестве целых чисел называют задачей нахождения дискретного логарифма.

Определение. Число x называют дискретным логарифмом числа y по основанию a и модулю n , если для всех $a \in \mathbb{Z}_n$ найдется такое целое y , что

$$y = a^x \bmod n.$$

Вычисление дискретных логарифмов (когда заданы a , y и n) примерно такая же труднорешаемая задача, как и разложение на множители.

Определение. Односторонняя функция $f: X \rightarrow Y$ называется *односторонней функцией с ловушкой*, если $f^{-1}(y)$ можно вычислить за полиномиальное время, имея некоторую дополнительную информацию, т. е. существует функция $g(y, t)$, вычисляемая за полиномиальное время и такая, что $g(y, t) = f^{-1}(y)$ для некоторой ловушки t .

Эффективное вычисление обратной функции возможно, если известен "потайной ход" (секретное число, строка или другая информация, ассоциирующаяся с данной функцией). В качестве примера однонаправленной функции с "потайным ходом" можно привести использование функции Эйлера в криптосистеме RSA.

Криптосистема RSA

Алгоритм RSA стал первым полноценным алгоритмом с открытым ключом, который может работать как в режиме шифрования данных, так и в режиме электронной цифровой подписи. Основанная на этом алгоритме популярная криптосистема RSA разработана в 1977 году и получила название в честь ее создателей: Рональда Ривеста (в настоящее время он возглавляет компанию RSA Data Security), Ади Шамира и Леонарда Эйдельмана.

В настоящее время RSA является наиболее распространенной криптосистемой с открытым ключом — стандартом де-факто для многих криптографических приложений. Статус де-факто послужил причиной включения криптосистемы RSA в принятые ранее криптографические стандарты, например в финансовые стандарты США и Франции, австралийский стандарт управления ключами и многие другие. Криптосистема RSA

применяется в различных протоколах Internet. В криптографические стандарты, действующие на территории России, RSA не входит, что осложняет ее применение с точки зрения правовых норм. Тем не менее, выбор этой криптосистемы признается оправданным отечественными авторами [8].

Основные определения и теоремы

Надежность алгоритма основывается на трудновычислимых задачах факторизации (разложения на множители) больших чисел и вычисления дискретных логарифмов. Определения и теоремы из алгебры, использованные при создании данной криптосистемы рассмотрены в приложении. Приведем здесь только основные утверждения.

1. Криптографические системы являются стойкими, если определенные их параметры являются *простыми числами*. Число a называется простым, если оно не имеет целых делителей, кроме единицы. Числа a и b называются взаимно простыми, если их наибольший общий делитель $\text{НОД}(a, b)=1$.

2. *Функцией Эйлера* $\varphi(n)$ называется число положительных целых чисел меньших n и взаимно простых с n .

Вычисление функции Эйлера $\varphi(n)$ для больших n в общем случае представляет собой трудоемкую процедуру перебора всех чисел меньших n и проверки для каждого взаимной простоты с n . Однако, эта функция обладает следующими свойствами:

- $\varphi(p) = p - 1 \quad \forall p$ – простого числа.
- $\varphi(a, b) = \varphi(a) \varphi(b)$ для любых натуральных взаимно простых a и b ,

которые позволяют легко вычислить значение функции Эйлера $\varphi(n)$ с помощью трех арифметических действий, если известно разложение числа n на простые сомножители p и q :

$$\varphi(n)=(p-1)(q-1).$$

3. В RSA используется теорема, которая носит название *китайской теоремы об остатках*, так как этот результат был известен еще в древнем Китае, где теорема была предложена китайским математиком первого века Сун Це. Она утверждает, что любое неотрицательное целое число, не превосходящее произведения модулей, можно однозначно восстановить, если известны его вычеты по этим модулям, и названа так потому, что результатом приведения числа a по модулю n является остаток от деления a на n .

Фактически в RSA используется следствие из этой теоремы, утверждающее, что если известно разложение числа n на простые множители $n=n_1n_2\dots n_k$, где все n_i попарно взаимно просты, и результат приведения числа x по модулю $n_i \quad \forall i=1, \dots, k$ одинаков, то результатом приведения числа x по модулю n будет то же число. То есть $\forall x, a$ – целых чисел

$$x \equiv a \pmod{n} \Leftrightarrow x \equiv a \pmod{n_i} \quad \forall i=1, \dots, k.$$

4. **Теорема Эйлера.** Если $n > 1$, то $\forall x \in Z_n^*$ (x взаимно простого с n), выполняется сравнение

$$x^{\varphi(n)} \equiv 1 \pmod{n}.$$

Следствие. $\forall x < n$ и взаимно простого с n можно легко вычислить обратный элемент x^{-1} в кольце вычетов Z_n из сравнения

$$x^{-1} \equiv x^{\varphi(n)-1} \pmod{n}.$$

Малая теорема Ферма. $\forall x \in GF(p)$, $x \neq 0$, выполняется сравнение

$$x^{p-1} \equiv 1 \pmod{p}.$$

Малая теорема Ферма является следствием из теоремы Эйлера, хотя исторически она была доказана раньше, затем Эйлер её обобщил.

Следствие 1: Если p — простое число, то $\forall x$, взаимно простого с p :

$$x^p = x \pmod{p}.$$

Следствие 2: если $\text{НОД}(e, \varphi(n)) = 1$ (e - простое относительно $\varphi(n)$)

то $\exists d$ -целое, такое, что

$$ed = 1 \pmod{n}.$$

На этих математических фактах основан алгоритм RSA.

Алгоритм RSA

В криптосистеме RSA сообщение m , криптограмма c , открытый ключ K_o , и секретный ключ K_c , принадлежат множеству целых чисел $Z_n = \{0, 1, 2, \dots, n-1\}$. Множество Z_n с операциями сложения и умножения по модулю n образует кольцо.

Модуль n определяется как составное число равное произведению $n = p \cdot q$ двух больших простых чисел p и q . Модуль n является открытым параметром алгоритма, а числа p и q — секретными параметрами. То есть множители p и q хранят в секрете, а их произведение n известно всем, кто пользуется данной криптосистемой. Здесь используется односторонняя функция с ловушкой. Зная секретные параметры алгоритма p и q можно легко вычислить функцию Эйлера $\varphi(n)$ по формуле

$$\varphi(n) = (p-1)(q-1),$$

тогда как вычисление $\varphi(n)$ только по большому числу n является трудновычислимой задачей.

Функция Эйлера используется в RSA при вычислении ключей.

Открытый ключ $K_o = e$ выбирают случайным образом из множества

$Z_{\varphi(n)}^*$ — чисел меньших $\varphi(n)$ и взаимно простых с $\varphi(n)$.

Иначе условие $e \in Z_{\varphi(n)}^*$ раносильно выполнению двух условий

$1 < K_o \leq \varphi(n)$, и $\text{НОД}(K_o, \varphi(n))=1$,

которым должно удовлетворять случайно выбранное число $K_o = e$, чтобы оно могло служить открытым ключом в RSA.

Секретный ключ $K_c = d$ вычисляют так, чтобы выполнялось условие

$$K_c \cdot K_o = e \cdot d \equiv 1 \pmod{\varphi(n)}. \quad (*)$$

То есть, секретный ключ d является обратным элементом к открытому ключу e в множестве $Z_{\varphi(n)}$: $d = e^{-1} \pmod{\varphi(n)}$. Решение сравнения (*) можно найти с помощью расширенного алгоритма Евклида.

Заметим, что d и n также взаимно простые числа. Вычисление секретного ключа по открытому является трудновычислимой задачей, если неизвестны секретные параметры алгоритма p и q , так как при этом трудно вычислить значение функции Эйлера, то есть модуля, по которому приводятся результаты операций при вычислении ключей.

При выполнении шифрования и дешифрования вычисления приводятся по модулю n . Открытый ключ K_o и модуль n сообщают всем, с кем предполагают обмениваться сообщениями и используют для шифрования данных, а секретный ключ K_c хранят в секрете на стороне получателя и используют для дешифрования.

Преобразование шифрования определяет криптограмму c через пару (открытый ключ K_o , сообщение m) в соответствии со следующей формулой:

$$c = E_{K_o}(m) = m^{K_o} \pmod{n}.$$

В качестве алгоритма быстрого вычисления значения c используют ряд последовательных возведений в квадрат целого m и умножений на m с приведением по модулю n .

Обращение функции $c = m^{K_o} \pmod{n}$, т.е. определение значения m по известным значениям c , K_o и n , является задачей дискретного логарифмирования и практически неосуществимо при $n \approx 2^{512}$.

Однако задачу расшифрования криптограммы c , можно легко решить, используя секретный ключ K_c , по следующей формуле:

$$m = D_{K_c}(c) = c^{K_c} \pmod{n}.$$

Докажем, что в результате возведения криптограммы c в степень секретного ключа K_c получается исходный текст m . Процесс расшифрования можно записать так:

$$D_{K_c}(E_{K_o}(m)) = D_{K_c}(m^e) \pmod{n} = (m^e)^d \pmod{n} = m^{ed} \pmod{n}.$$

По условию выбора ключей

$$e \cdot d \equiv 1 \pmod{\varphi(n)} \quad (*)$$

мы можем написать, что $\exists k$ такое что, с учетом свойств $\varphi(n)$:

$$e \cdot d = k \cdot \varphi(n) + 1 = k(p-1)(q-1) + 1.$$

Подставим это выражение в показатель степени:

$$m^{ed} \equiv m (m^{(p-1)})^{k(q-1)} = (m^{(q-1)})^{k(p-1)}.$$

По малой теореме Ферма ($x^{p-1} \equiv 1 \pmod p$, p – простое):

$$m^{ed} \pmod p \equiv m (1)^{k(q-1)} \pmod p = m \pmod p.$$

Аналогично, заменив p на q , получим:

$$m^{ed} \pmod q \equiv m (1)^{k(p-1)} \pmod q = m \pmod q.$$

Далее, по следствию из китайской теоремы об остатках так как $n=pq$:

$$m^{ed} \pmod n = m \pmod n, \quad \text{ч.т.д.}$$

Таким образом, если криптограмму c возвести в степень K_c :

$$c^{K_c} \pmod n = m,$$

то в результате восстанавливается исходный открытый текст m .

Именно поэтому для вычисления секретного ключа K_c используют соотношение (*).

Процедуры шифрования и расшифрования в криптосистеме RSA

В реальных системах алгоритм RSA реализуется следующим образом. Предположим, что пользователь А хочет передать пользователю В сообщение в зашифрованном виде, используя криптосистему RSA. В таком случае пользователь А выступает в роли отправителя сообщения, а пользователь В - в роли получателя. Криптосистему RSA должен сформировать получатель сообщения, т.е. пользователь В, так как в этом случае не будет необходимости в передаче секретного ключа. Рассмотрим последовательность действий пользователя В и пользователя А.

Формирование криптосистемы (на стороне получателя информации) состоит в выборе параметров алгоритма и вычислении пары ключей:

1. Пользователь В выбирает два больших простых числа p и q . Это *секретные параметры* алгоритма, они хранятся в секрете на стороне получателя.

2. Пользователь В вычисляет значение модуля n криптосистемы как результат умножения первых двух чисел:

$$n = p \cdot q.$$

Это *общедоступный параметр* криптосистемы. Иногда его включают в открытый ключ.

3. Пользователь В, зная секретные параметры алгоритма p и q , вычисляет функцию Эйлера:

$$\varphi(n) = (p-1)(q-1)$$

и выбирает случайным образом простое число e как значение открытого ключа K_o с учетом выполнения условий:

$$K_o < \varphi(n) \text{ и } \text{НОД}(K_o, \varphi(n)) = 1.$$

4. Пользователь В вычисляет значение секретного ключа $K_c = d$ при решении сравнения

$$K_c \equiv K_o^{-1} \pmod{\varphi(n)}.$$

Для вычисления обратного элемента в кольце $Z_{\varphi(n)}$ используют частный режим работы расширенного алгоритма Евклида для определения НОД. Это можно осуществить, так как получатель В знает пару простых чисел (p, q) и может легко найти $\varphi(n)$. Заметим, что K_c и n должны быть взаимно простыми.

(В принципе открытый и закрытый ключи можно поменять местами, главное, чтобы выполнялось соотношение $e \cdot d \equiv 1 \pmod{\varphi(n)}$).

5. Пользователь В пересылает пользователю А пару чисел $\{e, n\}$ по незащищенному каналу.

Шифрование информации (на стороне отправителя). Если пользователь А хочет передать пользователю В сообщение m , он выполняет следующие шаги.

6. Пользователь А разбивает исходный открытый текст m на блоки $m_i, i=1, \dots, N$, каждый из которых может быть представлен в виде числа меньшего n

$$m_i \in \{0, 1, 2, \dots, n-1\}.$$

7. Пользователь А шифрует текст, представленный в виде последовательности чисел m_i с помощью ключа $K_o = e$ по формуле

$$c_i = m_i^e \pmod{n}$$

и отправляет криптограмму c_1, \dots, c_i пользователю В.

Дешифрование информации (на стороне получателя):

8. Пользователь В расшифровывает принятую криптограмму c_1, \dots, c_i используя секретный ключ $K_c = d$, по формуле

$$m_i = c_i^d \pmod{n}.$$

В результате будет получена последовательность чисел, которые представляют собой исходное сообщение m .

Таким образом, когда получатель В, который создает криптосистему, он защищает следующие параметры:

1. секретный ключ K_c ;
2. пару чисел (p, q) ;

Открытый ключ K_o и значение модуля n публикуются и доступны каждому, кто желает послать владельцу ключа сообщение, которое шифруется указанным алгоритмом.

После шифрования сообщение невозможно раскрыть с помощью открытого ключа. Владелец же закрытого ключа без труда может расшифровать принятое сообщение. Противнику же для того, чтобы определить значение секретного ключа K_c нужно суметь разложить число n на множители p и q , чтобы узнать бы "потайной ход" и вычислить значение функции Эйлера как $\varphi(n) = (p-1)(q-1)$.

Пример. Рассмотрим небольшой пример, иллюстрирующий применение алгоритма RSA. Зашифруем сообщение "CAB". Для простоты будем использовать маленькие числа (на практике применяются гораздо большие).

1. Выберем $p=3$ и $q=11$.

2. Определим $n=3 \cdot 11=33$.

3. Найдем $\varphi(n) = (p-1)(q-1) = 2 \cdot 10 = 20$.

4. Выберем в качестве секретного ключа d произвольное число взаимно простое с $\varphi(n)=20$, например, $d=3$.

5. Выберем открытый ключ $e = 7$. В качестве такого числа может быть взято любое число, для которого удовлетворяется соотношение

$$e \cdot d \bmod \varphi(n) = 7 \cdot 3 \bmod 20 = 1.$$

6. Представим шифруемое сообщение как последовательность целых чисел с помощью отображения: A=1, B=2, C=3. Тогда сообщение принимает вид (3,1,2). Зашифруем сообщение с помощью ключа $\{e=7, n=33\}$:

$$\text{ШТ1} = (3^7) \bmod 33 = 2187 \bmod 33 = 9,$$

$$\text{ШТ2} = (1^7) \bmod 33 = 1 \bmod 33 = 1,$$

$$\text{ШТ3} = (2^7) \bmod 33 = 128 \bmod 33 = 29.$$

7. Расшифруем полученное зашифрованное сообщение (9, 1, 29) на основе закрытого ключа $\{d=3, n=33\}$:

$$\text{ИТ1} = (9^3) \bmod 33 = 729 \bmod 33 = 3,$$

$$\text{ИТ2} = (1^3) \bmod 33 = 1 \bmod 33 = 1,$$

$$\text{ИТ3} = (29^3) \bmod 33 = 24389 \bmod 33 = 2.$$

Надежность RSA. RSA многие годы противостоит интенсивному криптоанализу. Доказано, что раскрытие шифра RSA эквивалентно решению задачи факторизации больших (100-200 двоичных разрядов) чисел. Важно, что в этой задаче для любой длины ключа можно дать нижнюю оценку числа операций для раскрытия шифра, а с учетом производительности современных компьютеров оценить и необходимое на это время.

Возможность гарантированно оценить защищенность алгоритма RSA стала одной из причин популярности этой СОК на фоне десятков других схем. Поэтому алгоритм RSA

используется в банковских компьютерных сетях, особенно для работы с удаленными клиентами (обслуживание кредитных карточек). В настоящее время алгоритм RSA активно реализуется как в виде самостоятельных криптографических продуктов, так и в качестве встроенных средств в популярных приложениях, а также используется во многих стандартах.

Криптосистема Эль-Гамала

Данная система является альтернативой RSA и при равном значении ключа обеспечивает ту же криптостойкость. Однако общего мнения по поводу предпочтительности того или иного метода нет.

В отличие от RSA метод Эль-Гамала основан на проблеме дискретного логарифма. Этим он похож на *алгоритм Диффи-Хелмана*. Если возводить число в степень в конечном поле достаточно легко, то восстановить аргумент по значению (то есть найти логарифм в множестве целых чисел) довольно трудно.

Основу системы составляют параметры p и g - числа, первое из которых p — простое, а второе ($g \in Z_p$) — целое. Данные параметры не являются секретными и могут быть общими для группы пользователей. В реальных схемах шифрования необходимо использовать в качестве модуля большое целое простое число, имеющее в двоичном представлении длину 512...1024 бит.

Секретный ключ $x \in Z_{p-1}$ генерируется случайным образом, а открытый ключ y вычисляется по формуле

$$y = g^x \bmod p.$$

Для шифрования сообщения m сначала выбирается случайное число k , взаимно простое с $p-1$, которое называют также сеансовым ключом.

Шифртекстом является пара чисел (a, b) , вычисляемая по формулам:

$$a = g^k \bmod p \text{ и}$$

$$b = y^k m \bmod p.$$

Таким образом, шифртекст в два раза длиннее открытого текста.

Для дешифрования вычисляется

$$m = \frac{b}{a^x} \bmod p.$$

Преобразование обратимо, так как

$$a^x = g^{kx} \bmod p$$

$$\text{и } \frac{b}{a^x} \equiv \frac{y^k m}{a^x} \equiv \frac{g^{xk} m}{a^x} = m \pmod{p}.$$

Число k называют также рандомизатором. Его использование означает, что здесь реализован многозначный шифр замены. При этом для зашифрования различных блоков (чисел) открытого текста необходимо использовать различные значения рандомизатора. При использовании одного и того же значения соответствующие шифртексты (a, b) и (a', b') , полученные для блоков открытых текстов m и m' , связаны соотношением

$$b(b')^{-1} = m(m')^{-1}$$

и текст m' можно вычислить, если известен текст m .

Стойкость криптосистемы Эль Гамала основана на сложности задачи логарифмирования в мультипликативной группе конечного простого поля. Эта криптосистема может быть обобщена для применения в любой конечной циклической группе. В качестве такой группы, помимо рассмотренной Z_p^* , чаще всего используется мультипликативная группа конечного поля $GF(2^m)$ и группа точек на эллиптической кривой над конечным полем (см. приложение П.8).

Алгоритм не запатентован, но попадает под действие патента на метод экспоненциального ключевого обмена Диффи-Хеллмана, рассмотренный ниже. Преобразование шифрования/дешифрования Эль Гамала по сути то же самое, что ключевой обмен по Диффи-Хеллману, за исключением того, что y – это часть ключа, а при шифровании сообщение умножается на y^k . Алгоритм цифровой подписи DSA, разработанный NIST (National Institute of Standard and Technology USA) и являющийся частью стандарта DSS частично опирается на рассмотренный метод.

Комбинированный метод шифрования

Главным достоинством криптосистем с открытым ключом является их потенциально высокая безопасность: нет необходимости ни передавать, ни сообщать кому бы то ни было значения секретных ключей, ни убеждаться в их подлинности. В симметричных криптосистемах существует опасность раскрытия секретного ключа во время передачи.

Однако алгоритмы, лежащие в основе криптосистем с открытым ключом, имеют следующие недостатки:

- генерация секретных и открытых ключей основана на генерации больших простых чисел, а проверка простоты чисел занимает много процессорного времени;
- процедуры шифрования и расшифрования, связанные с возведением в степень многозначного числа, достаточно трудоемки.

Поэтому быстродействие (скорость шифрования и дешифрования) в криптосистемах с открытым ключом обычно в сотни и тысячи раз меньше быстродействия симметричных криптосистем с секретным ключом.

Комбинированный метод шифрования позволяет сочетать преимущества высокой секретности, предоставляемые асимметричными криптоистемами с открытым ключом, с преимуществами высокой скорости работы, присущими симметричным криптосистемам с секретным ключом. При таком подходе криптосистема с открытым ключом применяется для шифрования, передачи и последующего расшифрования только секретного ключа симметричной криптосистемы. А симметричная криптосистема применяется для шифрования и передачи исходного открытого текста. В результате криптосистема с открытым ключом не заменяет симметричную криптосистему с секретным ключом, а лишь дополняет ее, позволяя повысить в целом защищенность передаваемой информации.

Пользователи А и В, использующие комбинированный метод шифрования, имеют каждый по паре асимметричных ключей шифрования

(K_{Ao}, K_{Ac}) и (K_{Bo}, K_{Bc}) .

Если пользователь А хочет передать зашифрованное комбинированным методом сообщение m пользователю В, то порядок его действий будет таков.

1. Создать (например, сгенерировать случайным образом) симметричный ключ, называемый в этом методе сеансовым ключом K_s .
2. Зашифровать сообщение m на сеансовом ключе K_s .
3. Зашифровать сеансовый ключ K_s на открытом ключе K_{Bo} пользователя В и своем секретном ключе K_{Ac} .
4. Передать по открытому каналу связи в адрес пользователя В зашифрованное сообщение вместе с зашифрованным сеансовым ключом.

Действия пользователя В при получении зашифрованного сообщения и зашифрованного сеансового ключа должны быть обратными:

5. Расшифровать на своем секретном ключе K_{Bc} и открытом ключе K_{Ao} пользователя А сеансовый ключ K_s .
6. С помощью полученного сеансового ключа K_s расшифровать и прочитать сообщение m .

При использовании комбинированного метода шифрования можно быть уверенным в том, что только пользователь В сможет правильно расшифровать ключ K_s и прочитать сообщение m .

Выбор длины ключей в комбинированном методе шифрования. Таким образом, при комбинированном методе шифрования применяются криптографические ключи как симметричных, так и асимметричных криптосистем. Очевидно, выбор длин ключей для каждого типа криптосистемы следует осуществлять таким образом, чтобы злоумышленнику было одинаково трудно атаковать любой механизм защиты комбинированной криптосистемы.

В таблице 3.1. приведены распространенные длины ключей симметричных и асимметричных криптосистем, для которых трудность атаки полного перебора примерно равна трудности факторизации соответствующих модулей асимметричных криптосистем.

Таблица 3.1. Длины ключей для симметричных и асимметричных криптосистем при одинаковой их криптостойкости

Симметричные криптосистемы	56	64	80	112	128
Асимметричные криптосистемы	384	512	786	179 ²	230 ⁴

Метод экспоненциального ключевого обмена Диффи-Хеллмана

Одни из авторов идеи криптосистем с открытым ключом, Диффи и Хеллман предложили новую идею - *открытое распределение ключей*.

Они задались вопросом: можно ли организовать такую процедуру взаимодействия абонентов А и В по открытым каналам связи, чтобы решить следующие задачи:

- 1) вначале у А и В нет никакой общей секретной информации, но в конце процедуры такая общая секретная информация (общий ключ) у А и В появляется, т. е. вырабатывается;
- 2) пассивный противник, который перехватывает все передачи информации и знает, что хотят получить А и В, тем не менее не может восстановить выработанный общий ключ А и В.

Метод получил название метода экспоненциального ключевого обмена. Он был первой криптосистемой с открытым ключом, хотя для обмена ключами можно использовать любые криптосистемы с открытым ключом, например, тот же алгоритм RSA.

Криптостойкость данного метода определяется трудоемкостью вычисления дискретного логарифма:

$$f(x) = g^x \bmod p,$$

где p — большое простое число, x — произвольное натуральное число, g — некоторый *примитивный элемент* поля Галуа $GF(p)$. Общеизвестно, что инвертирование функции $g^x \bmod p$, т.е. дискретное логарифмирование, является трудной математической задачей.

Сама процедура или *протокол выработки общего ключа* заключается в следующем. Значения p и g являются общедоступными параметрами протокола. Абоненты А и В независимо друг от друга случайно выбирают по одному большому натуральному числу x_A и x_B . Это их секретные ключи. Далее каждый из них вычисляет открытые ключи:

$$y_A = g^{x_A} \bmod p \text{ и } y_B = g^{x_B} \bmod p.$$

Потом они обмениваются этими элементами по каналу связи. Теперь абонент А, получив y_B и зная свой секретный элемент x_A , вычисляет новый элемент:

$$y_B^{x_A} \bmod p = (g^{x_B})^{x_A} \bmod p.$$

Аналогично поступает абонент В:

$$y_A^{x_B} \bmod p = (g^{x_A})^{x_B} \bmod p.$$

Тем самым у А и В появился общий элемент поля, равный $g^{x_A x_B}$. Этот элемент и объявляется общим ключом абонентов А и В.

Необратимость преобразования в этом случае обеспечивается тем, что достаточно легко вычислить показательную функцию в конечном поле Галуа, состоящем из p элементов (p — простое число). Обратная задача вычисления x из y будет достаточно сложной. Если p выбрано достаточно правильно, то извлечение логарифма потребует вычислений, пропорциональных $L(p) = \exp\{(\ln p \ln \ln p)^{0.5}\}$.

При всей простоте алгоритма Диффи-Хелмана его недостатком по сравнению с системой RSA является отсутствие гарантированной нижней оценки трудоемкости раскрытия ключа.

Алгоритмы практической реализации криптосистем с открытым ключом

Возведение в степень по модулю m

В криптографии используются вычисления по $\bmod m$, так как алфавит любой криптографической системы представляет собой конечное множество целых чисел. Арифметика вычетов к тому же легче реализуется на компьютерах, поскольку она ограничивает диапазон промежуточных значений и результата. Для k -битовых вычетов m промежуточные результаты любого сложения, вычитания или умножения будут не длиннее, чем $2k$ бит. Поэтому в арифметике вычетов мы можем выполнить возведение в степень без огромных промежуточных результатов. Вычисление степени некоторого числа по модулю другого числа,

$$a^d \bmod m,$$

представляет собой просто последовательность умножений и делений, но существуют приемы, ускоряющие это действие. Один из таких приемов стремится минимизировать количество умножений по модулю, другой — оптимизировать отдельные умножения по

модулю. Так как операции дистрибутивны, быстрее выполнить возведение в степень как поток последовательных умножений, каждый раз получая вычеты.

Например, для того, чтобы вычислить $a^8 \bmod m$, не выполняйте семь умножений и одно приведение по модулю; вместо этого выполните три меньших умножения и три меньших приведения по модулю:

$$a^8 \bmod m = ((a^2 \bmod m)^2 \bmod m)^2 \bmod m.$$

$$\text{Точно также, } a^{16} \bmod m = (((a^2 \bmod m)^2 \bmod m)^2 \bmod m)^2 \bmod m.$$

Вычисление $a^d \bmod m$, где d не является степенью 2, не намного труднее. Двоичная запись представляет x в виде суммы степеней 2: число 25 — это бинарное 11001, поэтому $25 = 16 + 8 + 1$. Тогда

$$a^{25} \bmod m = (a^8 a^{16}) \bmod m = (a^*(((a*a^2)^2)^2) \bmod m).$$

С продуманным сохранением промежуточных результатов нам понадобится только шесть умножений:

$$(((((((a^2 \bmod m)*a)^2 \bmod m)^2 \bmod m)^2 \bmod m)^2 \bmod m)^2 \bmod m * a) \bmod m.$$

Такой прием называется методом двоичных квадратов и умножений. Он использует простую и очевидную цепочку сложений, в основе которой лежит двоичное представление числа. Увеличение скорости вычислений при умножении 200-битовых чисел будет очень заметным.

Алгоритм вычисления $a^d \bmod m$.

Пусть натуральные числа a и d не превосходят по величине m .

1. Представим d в двоичной системе счисления:

$$d = d_0 2^r + \dots + d_{r-1} 2 + d_r,$$

где r — число двоичных разрядов, d_i — цифры в двоичном представлении, равные 0 или 1, и $d_0 = 1$.

2. Положим $a_0 = a$ и затем для $i = 1, \dots, r$ вычислим

$$a_i \equiv a^2_{i-1} a^{d_i} \bmod m.$$

Тогда a_r есть искомый вычет $a^d \bmod m$.

Справедливость этого алгоритма вытекает из легко доказываемого индукцией по i сравнения:

$$a_i \equiv a^{d_0 2^i + \dots + d_i} \bmod m.$$

Так как каждое вычисление на шаге 2 требует не более трех умножений по модулю m и этот шаг выполняется $r \leq \log_2 m$ раз, то сложность алгоритма может быть оценена величиной $O(\ln m)$. Говоря о сложности алгоритмов, мы имеем в виду количество арифметических операций.

Второй алгоритм — это классический алгоритм Евклида вычисления наибольшего общего делителя целых чисел. Мы предполагаем заданными два натуральных числа a и b и вычисляем их наибольший общий делитель $\text{НОД}(a, b)$.

Алгоритм Евклида вычисления НОД

Одним из способов вычисления НОД двух чисел является алгоритм Евклида, который описал его в своей книге «Начала» около 300 лет до н.э. Полагают, что он не изобрел его, а сам алгоритм ещё старше лет на 200. Это древнейший нетривиальный алгоритм, который актуален и в наше время.

1. Вычислим r — остаток от деления числа a на b ($a > b$):

$$a = bq + r, \quad 0 \leq r < b.$$

2. Если $r = 0$, то b есть искомое число.

3. Если $r \neq 0$, то заменим пару чисел (a, b) парой (b, r) и перейдем к шагу 1.

Остановка гарантируется, поскольку остатки от делений образуют строго убывающую последовательность.

Оценку сложности этого алгоритма дает следующая теорема.

Теорема. При вычислении наибольшего общего делителя $\text{НОД}(a, b)$ с помощью алгоритма Евклида будет выполнено не более $5p$ операций деления с остатком, где p есть количество цифр в десятичной записи меньшего из чисел a и b .

Вычисление обратных величин в кольце целых чисел

При вычислении ключей в асимметричных криптографических системах необходимо находить обратные элементы в кольце целых чисел Z_m . Элемент $x \in Z_m$ является обратным к $a \in Z_m$, ($x = a^{-1}$), если

$$x \cdot a \equiv 1 \pmod{m} \text{ или } a^{-1} \equiv x \pmod{m}.$$

В общем случае это сравнение может не иметь решений или иметь несколько решений. Оно имеет единственное решение тогда и только тогда, если числа a и m взаимно простые, то есть $\text{НОД}(a, m) = 1$.

Рассмотрим основные способы нахождения обратных величин в Z_m .

1. Метод перебора. Проверить поочередно значения $x \in (1, 2, \dots, m-1)$, пока не выполнится сравнение $x \cdot a \equiv 1 \pmod{m}$.

2. Если известна функция Эйлера (для RSA) $\varphi(m)$, то можно вычислить $a^{-1} \pmod{m} \equiv a^{\varphi(m)-1} \pmod{m}$, используя алгоритм быстрого возведения в степень. (Это следует из утверждения теоремы Эйлера: если a и m взаимно простые, то $a^{\varphi(m)} \equiv 1 \pmod{m}$.)

3. Нахождение обратной величины $a^{-1} \pmod{m}$ с помощью расширенного алгоритма Евклида.

Алгоритм Евклида можно обобщить способом, который имеет большое практическое значение. При этом способе во время вычисления НОД можно попутно вычислить такие целые числа x и y , что

$$a \cdot x + b \cdot y = \text{НОД}(a, b).$$

Этот вариант называется расширенным алгоритмом Евклида. Для вычисления обратной величины используется частный режим работы алгоритма Евклида, при котором

$$b = m \text{ и } \text{НОД}(a, m) = 1.$$

В самом деле сравнение $x \cdot a \equiv 1 \pmod{m}$ означает, что \exists целое k , что верно:

$$a x + m k = 1 \pmod{m}.$$

Эта задача равносильна поиску целых решений уравнения

$$a x + m k = 1.$$

Алгоритм поиска целых решений уравнения $a x + m k = 1$

Немного подправив алгоритм Евклида, можно достаточно быстро решать сравнения

$$a x + m k = 1 \pmod{m}$$

при условии, что $\text{НОД}(a, m) = 1$.

0. Определим матрицу $E = I$ -единичная матрица размером 2×2 .

1. Вычислим r — остаток от деления числа a на m :

$$a = m q + r, \quad 0 \leq r < m.$$

2. Если $r = 0$, то второй столбец матрицы E дает вектор $[x, k]^T$ решений уравнения.

3. Если $r \neq 0$, то заменим матрицу E матрицей

$$E = E \cdot \begin{pmatrix} 1 & 0 \\ 0 & -q \end{pmatrix}.$$

4. Заменим пару чисел (a, m) парой (m, r) и перейдем к шагу 1.

Три приведенных выше алгоритма относятся к разряду так называемых полиномиальных алгоритмов. Это название носят алгоритмы, сложность которых оценивается сверху степенным образом в зависимости от длины записи входящих чисел (см. подробности в [2]). Если наибольшее из чисел, подаваемых на вход алгоритма, не превосходит m , то сложность алгоритмов этого типа оценивается величиной $O(\ln^c m)$, где c — некоторая абсолютная постоянная. Во всех приведенных выше примерах $c=1$.

Полиномиальные алгоритмы в теории чисел — большая редкость. Да и оценки сложности алгоритмов чаще всего опираются на какие-либо недоказанные, но правдоподобные гипотезы, обычно относящиеся к аналитической теории чисел.

Генерация простых чисел

Для алгоритмов с открытыми ключами нужны простые числа. Их нужно множество для любой достаточно большой сети, хватит ли их?

Да, существует приблизительно 10^{151} простых чисел длиной до 512 бит включительно. Для чисел, близких m , вероятность того, что случайно выбранное число окажется простым, равна $1/\ln m$. Поэтому полное число простых чисел, меньших m , равно $m/(\ln m)$. При выборе из 10^{151} простых чисел вероятность совпадения выбора практически равна нулю.

Но если так трудоемко разложение на множители, как может быть простой генерация простых чисел? Дело в том, что ответить "да" или "нет" на вопрос "является ли число простым?" гораздо проще, чем ответить на более сложный вопрос "каковы множители m ?"

Генерация случайных чисел с последующей попыткой разложения их на множители - это неправильный способ поиска простых чисел. Существуют различные вероятностные проверки на простоту чисел, определяющие, является ли число простым, с заданной степенью достоверности. При условии, что эта «степень достоверности» достаточно велика, такие способы проверки достаточно хороши.

Предлагаются [6] следующие тесты проверки чисел на простоту:

Тест Леманна (Lehmann)

Вот последовательность действий при проверке простоты числа p :

(1) Выберите случайно число a , меньшее p .

(2) Вычислите $a^{(p-1)/2} \bmod p$.

(3) Если $a^{(p-1)/2} \not\equiv 1$ или $-1 \bmod p$, то p не является простым.

(4) Если $a^{(p-1)/2} \equiv 1$ или $-1 \bmod p$, то вероятность того, что число p не является простым, не больше 50 процентов.

Число a , которое не показывает, что число p не является простым числом, называется свидетелем. Если p — составное число, вероятность случайного числа a быть свидетелем составной природы числа p не меньше 50 процентов. Повторите эту проверку t раз с t различными значениями a . Если результат вычислений равен 1 или -1, но не всегда равен 1, то p является простым числом с вероятностью ошибки $1/2^t$.

Алгоритм Рабина-Миллера (Rabin-Miller)

Повсеместно используемым является простой алгоритм, разработанный М. Рабином, частично основанным на идеях Миллера.

Выберите для проверки случайное число p . Вычислите b — число делений $p-1$ на 2 (т.е. 2^b — это наибольшая степень числа 2, на которую делится $p-1$). Затем вычислите m , такое, что $p = 1 + 2^b \cdot m$.

- 1) Выберите случайное число a , меньшее p .
- 2) Установите $j = 0$ и $z = a \cdot m \bmod p$.
- 3) Если $z = 1$ или если $z = p-1$, то p проходит проверку и может быть простым числом.
- 4) Если $j > 0$ и $Z = 1$, то p не является простым числом.
- 5) Установите $j = j + 1$. Если $j < b$ и $z < p - 1$, установите $z = z^2 \bmod p$ и вернитесь на этап (4). Если $z = p - 1$, то p проходит проверку и может быть простым числом.
- 6) Если $j = b$ и $z \neq p - 1$, то p не является простым числом.

В этом тесте вероятность прохождения проверки составным числом убывает быстрее, чем в предыдущих. Гарантируется, что три четверти возможных значений a окажутся свидетелями. Это означает, что составное число проскользнет через t проверок с вероятностью, не большей $1/4^t$, где t - число итераций. На самом деле и эти оценки слишком пессимистичны. Для большинства случайных чисел около 99.9 процентов возможных значений являются свидетелями [4].

Существуют более точные оценки [4]. Для n -битового кандидата в простые числа (где $n > 100$) вероятность ошибки в одном тесте меньше, чем $4n2^{\sqrt{k}/2}$. И для 256-битового n вероятность ошибки в шести тестах меньше, чем $1/2^{51}$.

Практические соображения

В реальных приложениях генерация простых чисел происходит быстро. В [6] описан следующий алгоритм:

- 1) Сгенерируйте случайное n -битовое число p .
- 2) Установите старший и младший биты равными 1. (Старший бит гарантирует требуемую длину простого числа, а младший бит обеспечивает его нечетность.)
- 3) Убедитесь, что p не делится на небольшие простые числа: 3, 5, 7, 11 и т.д. Во многих реализациях проверяется делимость p на все простые числа, меньшие 256. Наиболее эффективной является проверка на делимость для всех простых чисел, меньших 2000.
- 4) Выполните тест Рабина-Миллера для некоторого случайного a . Если p проходит тест, сгенерируйте другое случайное a и повторите проверку. Выбирайте небольшие значения a для ускорения вычислений. Выполните пять тестов. (Одного может показаться достаточным, но выполните пять.) Если p не проходит одной из проверок, сгенерируйте другое p и попробуйте снова.

Иначе, можно не генерировать p случайным образом каждый раз, но последовательно перебирать числа, начиная со случайно выбранного до тех пор, пока не будет найдено простое число.

Этап 3 не является обязательным, но это хорошая идея. Проверка, что случайное нечетное p не делится на 3, 5 и 7 отсекает 54% нечетных чисел еще до этапа 4. Проверка делимости на все простые числа, меньшие 100, убирает 76% нечетных чисел, проверка делимости на все простые числа, меньшие 256, убирает 80% нечетных чисел. В общем случае, доля нечетных кандидатов, которые не делятся ни на одно простое число, меньшее m , равна $1.12/1n m$. Чем больше проверяемое m , тем больше предварительных вычислений нужно выполнить до теста Рабина-Миллера.

Сильные простые числа.

Если m – произведение двух простых чисел p и q , то может понадобиться использовать в качестве и q *сильные простые числа*.

Такие простые числа обладают рядом свойств, которые усложняют разложение произведения m определенными методами разложения на множители. Среди таких свойств были предложены [4]:

1. Наибольший общий делитель $p-1$ и $q-1$ должен быть небольшим.
2. И $p-1$, и $q-1$ должны иметь среди своих множителей большие простые числа, соответственно p' и q' .
3. И $p'-1$, и $q'-1$ должны иметь среди своих множителей большие простые числа.
4. И $p+1$, и $q+1$ должны иметь среди своих множителей большие простые числа.
5. И $(p-1)/2$, и $(q-1)/2$ должны быть простыми (обратите внимание, при выполнении этого условия выполняются и два первых).

Насколько существенно применение именно сильных простых чисел остается предметом продолжающихся споров. Эти свойства были разработаны, чтобы затруднить выполнение ряда старых алгоритмов разложения на множители. Однако самые быстрые алгоритмы одинаково быстры при разложении на множители любых чисел, как удовлетворяющих приведенным условиям, так и нет.

Некоторые авторы [4] выступают против специальной генерации сильных простых чисел, утверждая, что длина простых чисел гораздо важнее их структуры. Более того, сама структура уменьшает случайность числа и может снизить устойчивость системы.

Но все может измениться. Могут быть созданы новые методы разложения на множители, которые лучше работают с числами, обладающими определенными свойствами. В этом случае снова могут потребоваться сильные простые числа.

8.2. Компьютерный практикум для шифров с открытым ключом

8.2.1. Изучение отечественных стандартов функции хэширования и цифровой подписи.

Цель работы

Целью данной лабораторной работы является изучение процессов вычисления функции хэширования по ГОСТ Р 34.11-94, формирования и проверки ЭЦП на эллиптических кривых по ГОСТ Р 34.10-2001.

Краткие теоретические сведения

Электронная Цифровая Подпись

При обмене электронными документами по сети связи существенно снижаются затраты на обработку и хранение документов, убыстряется их поиск. Но при этом возникает проблема аутентификации автора документа и самого документа, т.е. установления подлинности автора и отсутствия изменений в полученном документе. В обычной (бумажной) информатике эти проблемы решаются за счет того, что информация в документе и рукописная подпись автора жестко связаны с физическим носителем (бумагой). В электронных документах на машинных носителях такой связи нет.

Целью аутентификации электронных документов является их защита от возможных видов злоумышленных действий, к которым относятся:

- *активный перехват* – нарушитель, подключившийся к сети, перехватывает документы (файлы) и изменяет их;
- *маскарад* – абонент *C* посылает документ абоненту *B* от имени абонента *A*;
- *рenegатство* – абонент *A* заявляет, что не посылал сообщения абоненту *B*, хотя на самом деле послал;
- *подмена* – абонент *B* изменяет или формирует новый документ и заявляет, что получил его от абонента *A*;
- *повтор* – абонент *C* повторяет ранее переданный документ, который абонент *A* посылал абоненту *B*.

Эти виды злоумышленных действий могут нанести существенный ущерб банковским и коммерческим структурам, государственным предприятиям и организациям, частным лицам, применяющим в своей деятельности компьютерные информационные технологии.

При обработке документов в электронной форме совершенно непригодны традиционные способы установления подлинности по рукописной подписи и оттиску печати на бумажном

документе. Принципиально новым решением является электронная цифровая подпись (ЭЦП).

Электронная цифровая подпись предназначена для аутентификации лица, подписавшего электронное сообщение. Функционально она аналогична обычной рукописной подписи.

Электронная цифровая подпись обладает следующими свойствами:

- осуществляет контроль целостности передаваемого подписанного сообщения;
- предоставляет возможность доказательно подтвердить авторство лица, подписавшего сообщение;
- не дает самому этому лицу возможности отказаться от обязательств, связанных с подписанным текстом;
- обеспечивает защиту от возможностей подделки сообщения.

Цифровая подпись представляет собой относительно небольшое количество дополнительной цифровой информации, передаваемой вместе с подписываемым текстом.

Система ЭЦП включает две процедуры:

- процедуру формирования подписи;
- процедуру проверки подписи.

В процедуре постановки подписи используется секретный ключ отправителя сообщения, в процедуре проверки подписи – открытый ключ отправителя.

При формировании ЭЦП отправитель прежде всего вычисляет хэш-функцию $h(M)$ подписываемого текста M . Вычисленное значение хэш-функции $h(M)$ представляет собой один короткий блок информации t , характеризующий весь текст M в целом. Затем число t шифруется секретным ключом отправителя. Получаемая при этом пара чисел представляет собой ЭЦП для данного текста M .

При проверке ЭЦП получатель сообщения снова вычисляет хэш-функцию $t = h(M)$ принятого по каналу текста M , после чего при помощи открытого ключа отправителя проверяет, соответствует ли полученная подпись вычисленному значению t хэш-функции.

Принципиальным моментом в системе ЭЦП является невозможность подделки ЭЦП пользователя без знания его секретного ключа подписывания.

В качестве подписываемого документа может быть использован любой файл.

Функция хэширования

Функция хэширования (хэш-функция, или дайджест-функция) представляет собой отображение, на вход которого подается сообщение переменной длины M , а выходом является строка фиксированной длины $h(M)$. Иначе говоря, хэш-функция $h(\cdot)$ принимает в качестве аргумента сообщение (документ) M произвольной длины и возвращает хэш-значение $h(M)$ фиксированной длины.

Хэш-значение $h(M)$ – это сжатое двоичное представление (дайджест) основного сообщения M произвольной длины. Функция хэширования позволяет сжать подписываемый документ M до нескольких десятков или сотен бит (например, 128 или 256 бит), тогда как M может быть размером в мегабайт или более. Следует отметить, что значение хэш-функции $h(M)$ зависит от документа M сложным образом и не позволяет восстановить сам документ M .

Функция хэширования может использоваться для обнаружения модификации сообщения, то есть служить в качестве криптографической контрольной суммы (также называемой кодом обнаружения изменений или кодом аутентификации сообщения). В этом качестве функция хэширования применяется при формировании и проверке электронной цифровой подписи.

Модулярная арифметика

Пусть a и n – натуральные числа. «Разделить число a на число n с остатком» – это значит найти целые числа q и r , удовлетворяющие условию

$$a = q \cdot n + r, 0 \leq r < n. \quad (32)$$

При этом число q называют неполным частным, а r – остатком от деления числа a на число n .

Если остаток r равен нулю, то говорят, что число n делит число a , или, по-другому, n является делителем числа a .

Целые числа a и b называются *сравнимыми по модулю n* , если их остатки при делении на n совпадают. Обычно для обозначения этого факта используется запись

$$a \equiv b \pmod{n}. \quad (33)$$

Отсюда, в частности, следует, что число n делит разность чисел a и b . Для обозначения остатка обычно используют бесскобочную запись

$$b = a \bmod n. \quad (34)$$

Операцию нахождения числа $b = a \bmod n$ называют *приведением числа a по модулю n* .

Конечное множество целых чисел a_0, \dots, a_{n-1} , таких, что для любого целого числа b найдется $k \in \{0, \dots, n-1\}$ со свойством $a_k \equiv b \pmod{n}$, называется *полной системой вычетов по модулю n* . Обычно используется полная система вычетов $\{0, \dots, n-1\}$.

Эллиптические кривые

Эллиптические кривые применяются в криптографии с 1985 года. Интерес к ним обусловлен тем, что на их основе обеспечиваются те же криптографические свойства, которыми обладают числовые или полиномиальные криптосистемы, но при существенно меньшем размере ключа.

Пусть задано простое число $p > 3$. Тогда эллиптической кривой E , определённой над конечным простым полем F_p , называется кривая, которая задается уравнением вида

$$y^2 \equiv x^3 + ax + b \pmod{p}, \quad (35)$$

где $a, b \in F_p$, и $4a^3 + 27b^2$ не сравнимо с нулем по модулю p .

Эллиптическая кривая E состоит из всех точек (x, y) , $x, y \in F_p$, удовлетворяющих тождеству (1.4), и бесконечно удаленной точки O , которая называется *нулевой точкой*.

Точки эллиптической кривой будем обозначать $Q = (x, y)$ или просто Q . Две точки эллиптической кривой равны, если равны их соответствующие x - и y -координаты.

Инвариантом эллиптической кривой называется величина $J(E)$, удовлетворяющая тождеству

$$J(E) \equiv 1728 \cdot \frac{4a^3}{4a^3 + 27b^2} \pmod{p} \quad (36)$$

Коэффициенты a, b эллиптической кривой E , по известному инварианту $J(E)$, определяются следующим образом

$$\begin{cases} a \equiv 3k \pmod{p}, \\ b \equiv 2k \pmod{p}, \end{cases} \quad \text{где } k \equiv \frac{J(E)}{1728 - J(E)} \pmod{p}, J(E) \neq 0 \text{ или } 1728. \quad (37)$$

Таким образом эллиптическая кривая E может быть однозначно задана либо своим инвариантом $J(E)$, либо коэффициентами $a, b \in F_p$.

Функция хэширования ГОСТ Р 34.11-94

Российский стандарт ГОСТ Р 34.11-94 определяет алгоритм и процедуру вычисления хэш-функции для любых последовательностей двоичных символов, применяемых в криптографических методах обработки и защиты информации. Этот стандарт базируется на блочном алгоритме шифрования ГОСТ 28147-89, хотя в принципе можно было бы использовать и другой блочный алгоритм шифрования с 64-битовым блоком и 256-битовым ключом.

Введем обозначения, которые будут использоваться в данном пункте:

$\{01\}^*$ – множество двоичных строк произвольной конечной длины;

$\{01\}^n$ – множество двоичных строк длиной n бит;

$\{0\}^n$ – двоичная строка из n нулей;

$|A|$ – длина слова A в битах;

$A \oplus B$ – побитное сложение слов A и B по $\text{mod } 2$, или попросту *XOR*;

$A \otimes B$ – сложение слов A и B по $\text{mod } 2^k$, где $k = |A| = |B|$;

$A \boxplus B$ – сложение слов A и B по $\text{mod } 2^{256}$;

\leftarrow оператор присвоения;

$\|$ – конкатенация.

Сообщения с произвольной длины можно сжать используя хэш-функцию с фиксированным размером входа при помощи двух методов:

- последовательного (итерационного);
- параллельного.

Создатели ГОСТ Р 34.11-94 пошли по первому пути и использовали метод последовательного хэширования использующий хэш-функцию с фиксированным размером входа $h: \{01\}^{2n} \rightarrow \{01\}^n$ (см. рисунок 1.1), то есть функцию сжатия с коэффициентом 2.

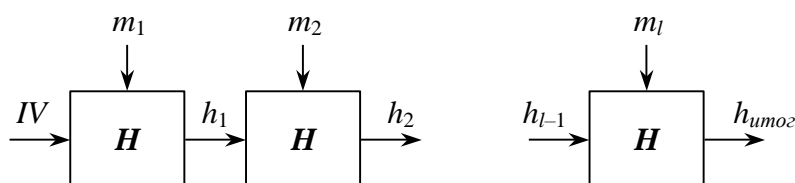


Рис. 8.1. Метод последовательного хэширования

Если необходимо хэшировать сообщение $m = (m_1, m_2, \dots, m_l)$, то хэширование выполняется следующим образом:

$$h_0 \leftarrow IV,$$

$$h_i \leftarrow H(m_i, h_{i-1}), \text{ для } i = 1, 2, \dots, l,$$

$$h_{\text{умог}} \leftarrow h_l.$$

Здесь H – функция сжатия, а h_i – переменная сцепления, IV – начальный вектор.

Если последний блок меньше чем n бит, то он набивается одним из существующих методов до достижения длины кратной n . В отличие от стандартных предпосылок, что сообщение разбито на блоки и произведена набивка последнего блока (форматирование входа априори), если необходимо, до начала хэширования, то в ГОСТ Р 34.11-94 процедура хэширования ожидает конца сообщения (форматирование входного сообщения постериори). Набивка производится следующим образом: последний блок сдвигается вправо, а затем набивается нулями до достижения длины в 256 бит (рисунок 8.3).

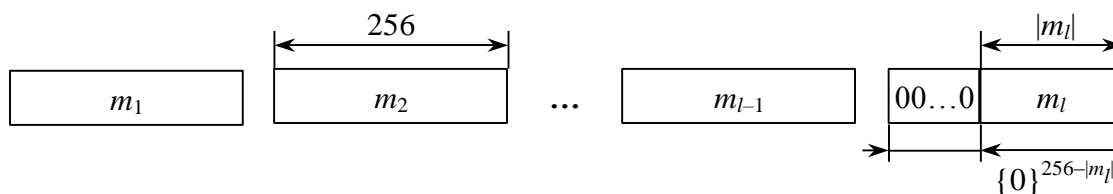


Рис. 8.2. Набивка сообщения

Указывать в передаваемом сообщении сколько было добавлено нулей к последнему блоку не требуется, так как длина сообщения участвует в хэшировании.

Параллельно рассчитываются контрольная сумма, представляющая собой сумму всех блоков сообщения (последний суммируется уже набитым) по правилу $A \otimes B$, и битовая длина хэшируемого сообщения, приводимая по $\text{mod } 2^{256}$, которые в финальной функции сжатия используются для вычисления итогового хэша (рисунок 1.3).

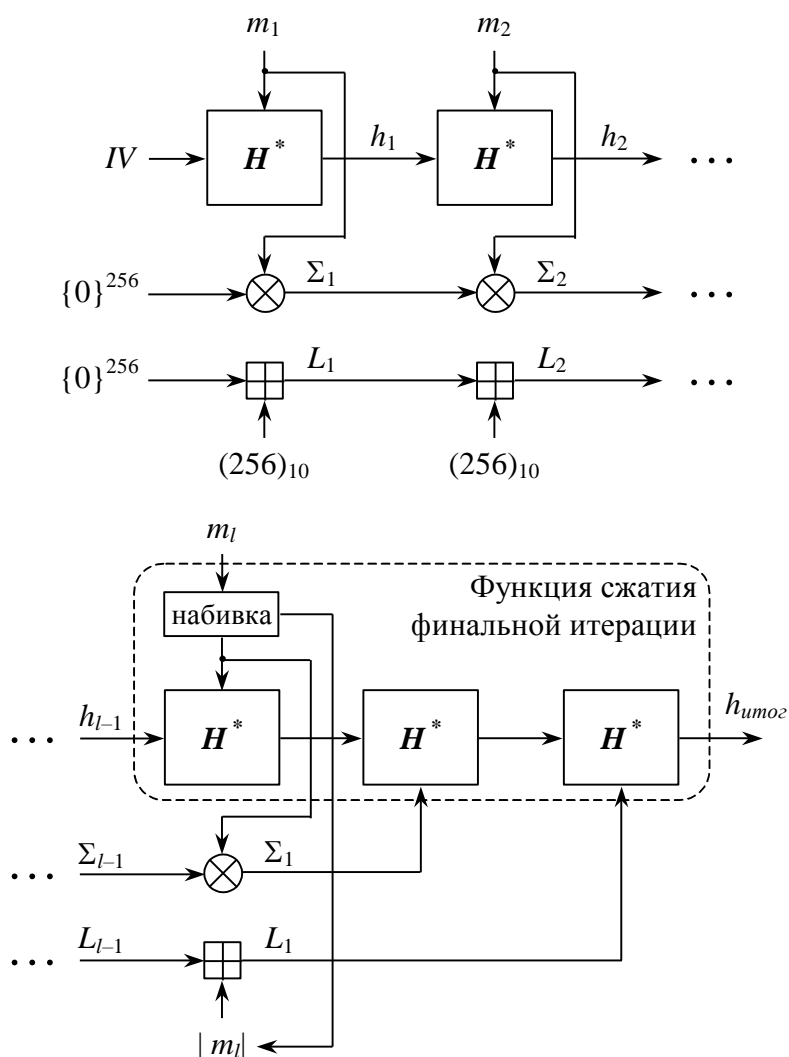


Рис. 8.3. Общая схема функции хэширования по ГОСТ Р 34.11-94

Согласно ГОСТ Р 34.11-94 IV – произвольное фиксированное слово длиной 256 бит ($IV \in \{01\}^{256}$). В таком случае, если он априорно не известен верифицирующему целостность сообщения, то он должен передаваться вместе с сообщением с гарантией целостности. При небольших сообщениях можно усложнить задачу противнику, если IV выбирается из небольшого множества допустимых величин (но при этом увеличивается вероятность угадывания хэш величины противником). Также он может задаваться в рамках организации, домена как константа (обычно как в тестовом примере).

Функция сжатия внутренних итераций

Функция сжатия внутренних итераций (по ГОСТ “шаговая функция хэширования”) H^* отображает два слова длиной 256 бит в одно слово длиной 256 бит:

$$H^*: \{01\}^{256} \times \{01\}^{256} \rightarrow \{01\}^{256},$$

и состоит из трех процедур (рисунок 3.5):

- перемешивающего преобразования;
- шифрующего преобразования;
- генерирования ключей.

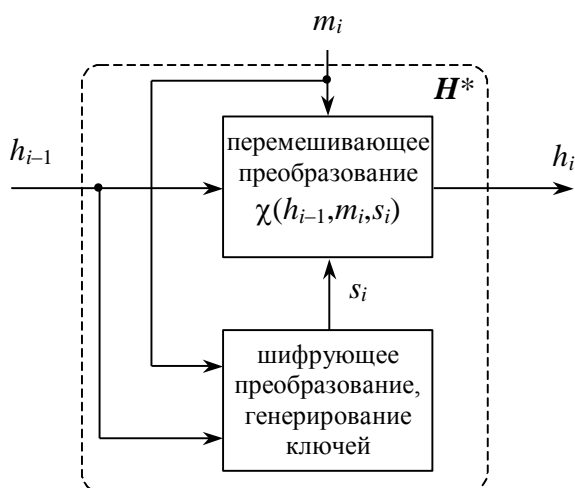


Рис. 8.4. Структура функции сжатия в ГОСТ Р 34.11-94

Генерирования ключей

Данная подпрограмма использует следующие функции и константы:

- Константы

$$C_2 = C_4 = \{0\}^{256} \text{ и } C_3 = 1^8 0^8 1^{16} 0^{24} 1^{16} 0^8 (0^8 1^8)^2 1^8 0^8 (0^8 1^8)^4 (1^8 0^8)^4,$$

где степень обозначает количество повторений 0 или 1.

- Функции

- $P: \{01\}^{256} \rightarrow \{01\}^{256}$.

Пусть $X = \xi_{32} \parallel \xi_{31} \parallel \dots \parallel \xi_1$,

тогда $P(X) = \xi_{\varphi(32)} \parallel \xi_{\varphi(31)} \parallel \dots \parallel \xi_{\varphi(1)}$, где $\varphi(i+1+4(k-1)) = 8i+k$, $i = 0 \dots 3$, $k = 1 \dots 8$.

- $A: \{01\}^{256} \rightarrow \{01\}^{256}$.

Пусть $X = x_4 \parallel x_3 \parallel x_2 \parallel x_1$,

тогда $A(X) = (x_1 \oplus x_2) \parallel x_4 \parallel x_3 \parallel x_2$.

При вычислении ключей реализуется следующий алгоритм (рисунок 1.5):

1. $j \leftarrow 1, V \leftarrow m_i, U \leftarrow h_{i-1}$;
2. $W \leftarrow U \oplus V, K_1 \leftarrow P(W)$;
3. $j \leftarrow j + 1$;
4. Проверить условие $j = 5$. При положительном исходе перейти к шагу 7, при отрицательном – к шагу 5;
5. $U \leftarrow A(U) \oplus C_j, V \leftarrow A(A(V)), W \leftarrow U \oplus V, K_j \leftarrow P(W)$;
6. Перейти к шагу 3;
7. Выход (K_1, K_2, K_3, K_4) .

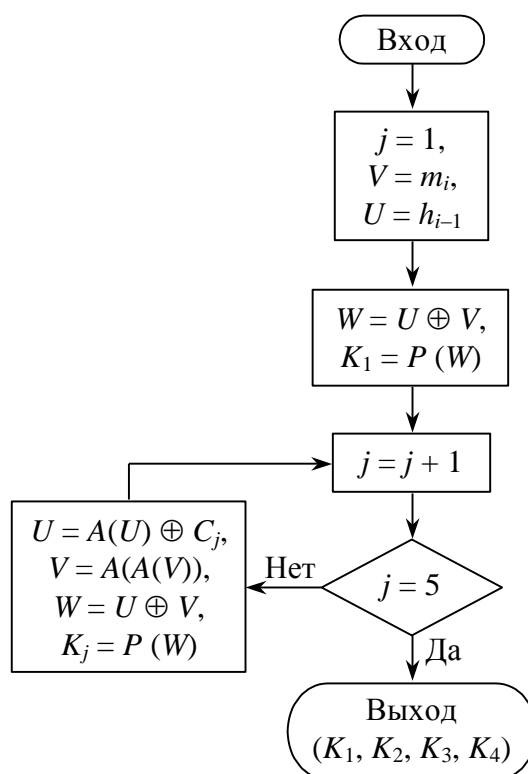


Рис. 8.5. Алгоритм генерирования ключей

Шифрующее преобразование

Основным функциональным предназначением является получение s_i из h_{i-1} .

Пусть $h_{i-1} = h_{i-1}^4 \parallel h_{i-1}^3 \parallel h_{i-1}^2 \parallel h_{i-1}^1$, где $h_{i-1}^j \in \{01\}^{64}$, $j = 1 \dots 4$, а

$s_i = s_i^4 \parallel s_i^3 \parallel s_i^2 \parallel s_i^1$, где $s_i^j \in \{01\}^{64}$, $j = 1 \dots 4$.

Тогда

$$s_i^j \leftarrow E_{K_j}(h_{i-1}^j),$$

где $j = 1 \dots 4$, E_{K_j} – шифрование по ГОСТ 28147-89 в режиме простой замены.

Схематически это изображено на рисунке 1.6.

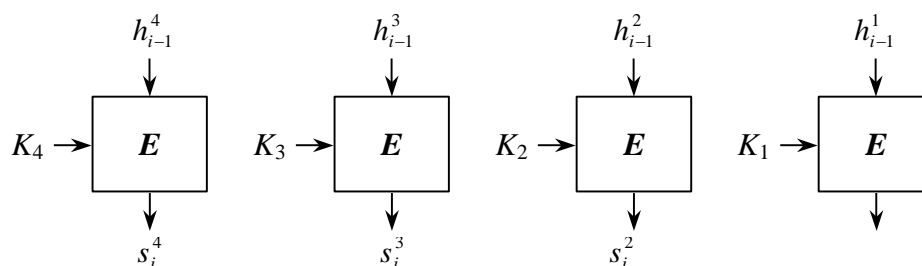


Рис. 8.6. Шифрующее преобразование

ГОСТ 28147-89 предусматривает три следующих режима шифрования данных:

- простая замена,
- гаммирование,
- гаммирование с обратной связью,
- и один дополнительный режим выработки имитовставки.

В любом из этих режимов данные обрабатываются блоками по 64 бита, на которые разбивается массив, подвергаемый криптографическому преобразованию, именно поэтому ГОСТ относится к блочным шифрам.

Если внимательно изучить оригинал ГОСТа 28147-89, можно заметить, что в нем содержится описание алгоритмов нескольких уровней. На самом верхнем находятся практические алгоритмы, предназначенные для шифрования массивов данных и выработки для них имитовставки. Все они опираются на три алгоритма низшего уровня, называемые в тексте ГОСТа *циклами*. Эти фундаментальные алгоритмы упоминаются в данной работе как *базовые циклы*, чтобы отличать их от всех прочих циклов. Они имеют следующие названия и обозначения:

- C_{32-3} – цикл зашифрования (32-3);
- C_{32-P} – цикл расшифрования (32-P);

В свою очередь, каждый из базовых циклов представляет собой многократное повторение одной единственной процедуры, называемой для определенности далее в настоящей работе *основным шагом криптопреобразования*.

В соответствии с принципом Кирхгофа, которому удовлетворяют все современные известные широкой общественности шифры, секретность зашифрованного сообщения обеспечивается секретностью ключевой информации. В ГОСТе ключевая информация

состоит из двух структур данных. Помимо собственно *ключа*, необходимого для всех шифров, она содержит еще и *таблицу замен*:

1. **Ключ** является массивом из восьми 32-битных элементов кода: $K = \{K_i\}_{0 \leq i \leq 7}$. В ГОСТе элементы ключа используются как 32-разрядные целые числа без знака: $0 \leq K_i < 2^{32}$. Таким образом, размер ключа составляет $32 \cdot 8 = 256$ бит или 32 байта.

2. **Таблица замен** является матрицей 8×16 , содержащей 4-битовые элементы, которые можно представить в виде целых чисел от 0 до 15: $H = \{H_{i,j}\}_{\substack{0 \leq i \leq 7 \\ 0 \leq j \leq 15}}$, $0 \leq H_{i,j} \leq 15$. Строки *таблицы замен* называются *узлами замен*, они должны содержать различные значения, то есть каждый узел замен должен содержать 16 различных чисел от 0 до 15 в произвольном порядке. Таким образом, общий объем таблицы замен равен: $8 \text{ узлов} \times 16 \text{ элементов/узел} \times 4 \text{ бита/элемент} = 512 \text{ бит}$ или 64 байта.

Основной шаг криптопреобразования по своей сути является оператором, определяющим преобразование 64-битового блока данных. Дополнительным параметром этого оператора является 32-битовый блок, в качестве которого используется какой-либо элемент ключа. Схема алгоритма основного шага приведена на рисунке 1.7.

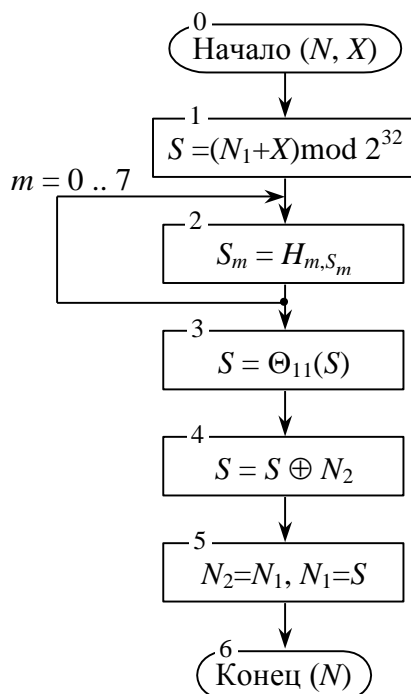


Рис. 8.7. Схема основного шага криптопреобразования алгоритма ГОСТ 28147-89

Ниже даны пояснения к алгоритму основного шага:

0. Определяет исходные данные для основного шага криптопреобразования:

- N – преобразуемый 64-битовый блок данных, в ходе выполнения шага его младшая (N_1) и старшая (N_2) части обрабатываются как отдельные 32-битовые целые числа без знака. Таким образом, можно записать $N = (N_1, N_2)$.
- X – 32-битовый элемент ключа.

1. Сложение с ключом. Младшая половина преобразуемого блока складывается по модулю 2^{32} с используемым на шаге элементом ключа, результат передается на следующий шаг.

2. Поблочная замена. 32-битовое значение, полученное на предыдущем шаге, интерпретируется как массив из восьми 4-битовых блоков кода:

$$S = (S_7, S_6, S_5, S_4, S_3, S_2, S_1, S_0).$$

Далее значение каждого из восьми блоков заменяется на новое, которое выбирается по таблице замен следующим образом: значение блока S_i заменяется на S_i -тый по порядку элемент (нумерация с нуля) i -того узла замен (т.е. i -той строки таблицы замен, нумерация также с нуля). Другими словами, в качестве замены для значения блока выбирается элемент из таблицы замен с номером строки, равным номеру заменяемого блока, и номером столбца, равным значению заменяемого блока как 4-битового целого неотрицательного числа. Теперь становится понятным размер таблицы замен: число строк в ней равно числу 4-битных элементов в 32-битном блоке данных, то есть восьми, а число столбцов равно числу различных значений 4-битного блока данных, равному как известно 2^4 , шестнадцати.

3. Циклический сдвиг на 11 бит влево. Результат предыдущего шага сдвигается циклически на 11 бит в сторону старших разрядов и передается на следующий шаг. На схеме алгоритма символом Θ_{11} обозначена функция циклического сдвига своего аргумента на 11 бит в сторону старших разрядов.

4. Побитовое сложение: значение, полученное на шаге 3, побитно складывается по модулю 2 со старшей половиной преобразуемого блока.

5. Сдвиг по цепочке: младшая часть преобразуемого блока сдвигается на место старшей, а на ее место помещается результат выполнения предыдущего шага.

6. Полученное значение преобразуемого блока возвращается как результат выполнения алгоритма основного шага криптопреобразования.

Базовые циклы заключаются в многократном выполнении *основного шага* с использованием разных элементов ключа и отличаются друг от друга только числом

повторения шага и порядком использования ключевых элементов. Ниже приведен этот порядок для различных циклов.

1. Цикл зашифрования 32-З:

$K_0, K_1, K_2, K_3, K_4, K_5, K_6, K_7, K_0, K_1, K_2, K_3, K_4, K_5, K_6, K_7, K_0, K_1, K_2, K_3, K_4, K_5, K_6, K_7, K_7, K_6, K_5, K_4, K_3, K_2, K_1, K_0.$

2. Цикл расшифрования 32-Р:

$K_0, K_1, K_2, K_3, K_4, K_5, K_6, K_7, K_7, K_6, K_5, K_4, K_3, K_2, K_1, K_0, K_7, K_6, K_5, K_4, K_3, K_2, K_1, K_0, K_7, K_6, K_5, K_4, K_3, K_2, K_1, K_0.$

Каждый из циклов имеет собственное буквенно-цифровое обозначение, где первый элемент обозначения, задает число повторений основного шага в цикле, а второй элемент обозначения, буква, задает порядок зашифрования («З») или расшифрования («Р») в использовании ключевых элементов.

Цикл расшифрования должен быть обратным циклу зашифрования, то есть последовательное применение этих двух циклов к произвольному блоку должно дать в итоге исходный блок, что отражается следующим соотношением:

$$C_{32-P}(C_{32-Z}(T))=T,$$

где T – произвольный 64-битный блок данных,

$C_X(T)$ – результат выполнения цикла X над блоком данных T .

Для выполнения этого условия для алгоритмов, подобных ГОСТу, необходимо и достаточно, чтобы порядок использования ключевых элементов соответствующими циклами был взаимно обратным. Из двух взаимно обратных циклов любой может быть использован для зашифрования, тогда второй должен быть использован для расшифрования данных, однако стандарт ГОСТ28147-89 закрепляет роли за циклами и не предоставляет пользователю права выбора в этом вопросе.

Схемы базовых циклов приведены на рисунках 1.8а, 1.8б. Каждый из них принимает в качестве аргумента и возвращает в качестве результата 64-битный блок данных, обозначенный на схемах N .

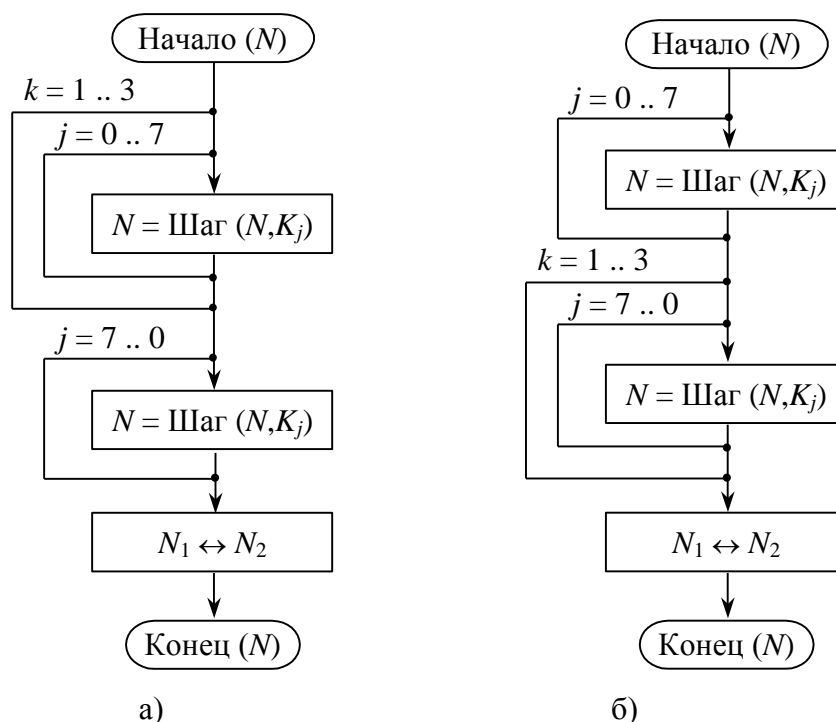


Рис. 8.8. Схемы цикла зашифрования 32-3 (а) и расшифрования 32-Р (б)

Символ $\text{Шаг}(N, X)$ обозначает выполнение основного шага криптопреобразования для блока N с использованием ключевого элемента X .

В конце базовых циклов шифрования старшая и младшая часть блока результата меняются местами, это необходимо для их взаимной обратимости.

Как уже упоминалось выше, нас интересует шифрование в режиме простой замены. Зашифрование в режиме простой замены заключается в применении цикла 32-3 к блокам открытых данных, расшифрование – цикла 32-Р к блокам зашифрованных данных. Это наиболее простой из режимов, 64-битовые блоки данных обрабатываются в нем независимо друг от друга. Схемы алгоритмов зашифрования и расшифрования в режиме простой замены приведены на рисунках 3.10а и 3.10б соответственно.

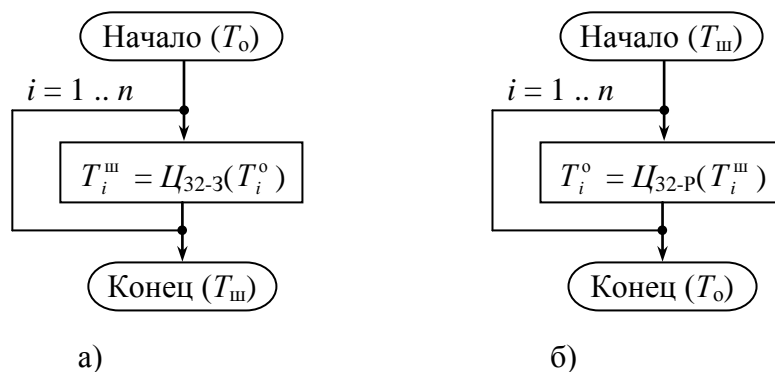


Рис. 8.9. Алгоритмы зашифрования (а) и расшифрования (б)

данных в режиме простой замены

На рисунке 3.10 используются следующие обозначения:

- $T_o, T_{\text{ш}}$ – массивы соответственно открытых и зашифрованных данных;
- $T_i^o, T_i^{\text{ш}}$ – i -тые по порядку 64-битные блоки соответственно открытых и зашифрованных данных: $T_o = (T_n^o, \dots, T_2^o, T_1^o), T_{\text{ш}} = (T_n^{\text{ш}}, \dots, T_2^{\text{ш}}, T_1^{\text{ш}}), 1 \leq i \leq n$;
- n – число 64-битных блоков в массиве данных.

Размер массива открытых или зашифрованных данных, подвергающийся соответственно зашифрованию или расшифрованию, должен быть кратен 64 битам: $|T_o|=|T_{\text{ш}}|=64 \cdot n$, после выполнения операции размер полученного массива данных не изменяется.

Перемешивающее преобразование

Перемешивающее преобразование имеет вид

$$h_i = \chi(m_i, h_{i-1}, s_i) = \psi^{61}(h_{i-1} \oplus \psi(m_i \oplus \psi^{12}(s_i))),$$

где ψ^j – j -я степень преобразования ψ .

Схематически данное преобразование представлено на рисунке 1.10.

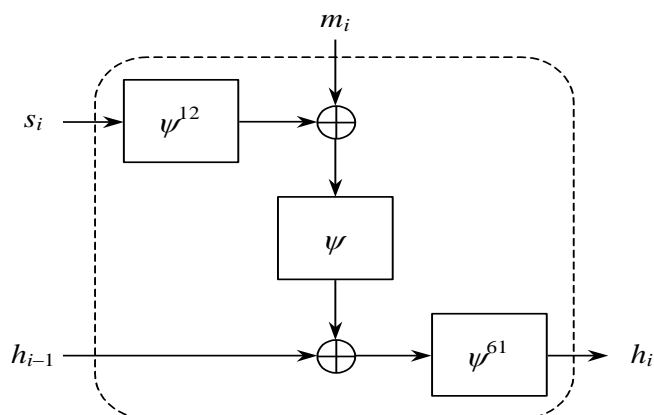


Рис. 8.10. Перемешивающее преобразование ГОСТ Р 34.11-94

Заметим, что s_i выводится из h_{i-1} (см. рисунок 1.4).

Функция $\psi: \{01\}^{256} \rightarrow \{01\}^{256}$ преобразует слово $\eta_{16} \parallel \eta_{15} \parallel \dots \parallel \eta_1 \in \{01\}^{16}, i = 1 \dots 16$ в слово $\eta_1 \oplus \eta_2 \oplus \eta_3 \oplus \eta_4 \oplus \eta_{13} \oplus \eta_{16} \parallel \eta_{16} \parallel \eta_{15} \parallel \dots \parallel \eta_2$ (рисунок 1.11).

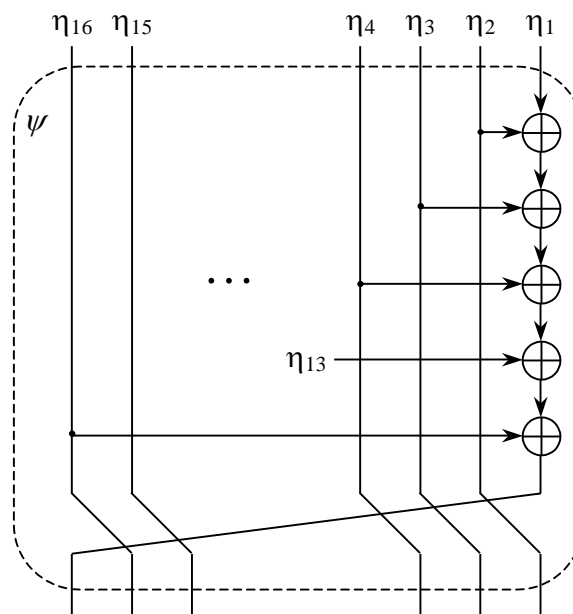


Рис. 8.11. Функция ψ перемешивающего преобразования

Функция сжатия финальной итерации

Функция сжатия финальной итерации производит итоговую хэш величину, зависящую от результата хэширования последовательным методом, контрольной суммы по mod 2 и длины сообщения в битах приведенной по mod 2^{256} :

$$m'_l \times h_{l-1} \times \Sigma_l \times L_l \rightarrow h_{итог},$$

где $m'_l, h_{l-1}, \Sigma_l, L_l \in \{01\}^{256}$, m'_l – последний набитый нулями блок (если необходимо, см. рисунок 1.2).

Сначала вычисляется битовая длина последнего (ненабитого) блока сообщения $|m_l| \leq 256$ бит. Если $|m_l| < 256$, то производится его набивка справа нулями до достижения длины 256 бит и получается новый блок $m'_l \leftarrow \{0\}^{256-|m_l|} \parallel m_l$. Вычисляются итоговая контрольная сумма $\Sigma_l \leftarrow \Sigma_{l-1} \otimes m'_l$ и длина всего сообщения $L_l \leftarrow L_{l-1} + |m_l| \pmod{2^{256}}$. Параллельно выполняется последняя итерация $h_l = H^*(m'_l, h_{l-1})$. Затем вычисляются значения $h_{l+1} \leftarrow H^*(\Sigma_l, h_l)$ и $h_{итог} \leftarrow H^*(L_l, h_{l+1})$, давая в результате итоговую хэш величину $h_{итог}$.

Теперь дадим формальное описание алгоритма:

Вход: двоичное сообщение M , блоки замен для шифрования в режиме простой замены по ГОСТ 28147-89, начальный вектор $IV \in \{01\}^{256}$.

Выход: хэш величина $h_{итог}$ для сообщения M .

1. *Инициализация алгоритма:* $h \leftarrow IV, \Sigma \leftarrow \{0\}^{256}, L \leftarrow \{0\}^{256}$.
2. *Функция сжатия внутренних итераций:* Пока $|M| > 256$ выполняем следующее:

2.1. $h \leftarrow H^*(m_s, h)$ (итерация метода последовательного хэширования);

- 2.2. $L \leftarrow L + 256 \pmod{2^{256}}$ (итерация вычисления длины сообщения);
- 2.3. $\Sigma \leftarrow \Sigma \otimes m_s$ (итерация вычисления контрольной суммы).
3. Иначе (функция сжатия финальной итерации)
- 3.1. $L \leftarrow L + |m| \pmod{2^{256}}$ (вычисление полной длины сообщения);
- 3.2. $m' \leftarrow \{0\}^{256-|m|} \parallel m$ (набивка последнего блока);
- 3.3. $\Sigma \leftarrow \Sigma \otimes m'$ (вычисление контрольной суммы сообщения);
- 3.4. $h \leftarrow H^*(m', h)$;
- 3.5. $h \leftarrow H^*(\Sigma, h)$;
- 3.6. $h_{\text{итог}} \leftarrow H^*(L, h)$.
4. Выход ($h_{\text{итог}}$).

Российский стандарт ЭЦП ГОСТ Р 34.10-2001

Новый отечественный стандарт цифровой подписи обозначается как ГОСТ Р 34.10-2001.

В нем используются следующие параметры:

- простое число p – модуль эллиптической кривой, удовлетворяющее неравенству $p > 2^{255}$. Верхняя граница данного числа должна определяться при конкретной реализации схемы цифровой подписи;
- эллиптическая кривая E , задаваемая своим инвариантом $J(E)$ или коэффициентами $a, b \in F_p$;
- целое число m – порядок группы точек эллиптической кривой E ;
- простое число q – порядок циклической подгруппы группы точек эллиптической кривой E , для которого выполнены следующие условия:

$$\begin{cases} m = nq, & n \in \mathbb{Z}, \quad n \geq 1, \\ 2^{254} < q < 2^{256}; \end{cases} \quad (38)$$

- точка $P \neq O$ эллиптической кривой E , с координатами (x_p, y_p) , удовлетворяющая равенству $qP = O$.
- хэш-функция, отображающая сообщения, представленные в виде двоичных векторов произвольной конечной длины, в двоичные вектора длины 256 бит. Хэш-функция определена в ГОСТ Р 34.11, ее алгоритм будет рассмотрен ниже.

Каждый пользователь схемы цифровой подписи должен обладать личными ключами:

- ключом подписи – целым числом d , удовлетворяющим неравенству $0 < d < q$;
- ключом проверки – точкой эллиптической кривой Q с координатами (x_q, y_q) , удовлетворяющей равенству $dP = Q$.

На приведенные выше параметры схемы цифровой подписи накладываются следующие требования:

- должно быть выполнено условие $p^t \neq 1 \pmod{q}$, для всех целых $t = 1, 2, \dots, B$, где B удовлетворяет неравенству $B \geq 31$;
- должно быть выполнено неравенство $m \neq p$;
- инвариант кривой должен удовлетворять условию $J(E) \neq 0$ или 1728.

Для получения цифровой подписи под сообщением M необходимо выполнить следующие действия (рис. 1.12а):

1. вычислить хэш-значение сообщения M : $\bar{h} = h(M)$;
2. вычислить целое число α , двоичным представлением которого является вектор \bar{h} , и определить $e \equiv \alpha \pmod{q}$. Если $e = 0$, то определить $e = 1$;
3. сгенерировать случайное (псевдослучайное) целое число k , удовлетворяющее неравенству $0 < k < q$;
4. вычислить точку эллиптической кривой $C = kP$ и определить $r \equiv x_c \pmod{q}$, где x_c - x -координата точки C . Если $r = 0$, то вернуться к шагу 3;
5. вычислить значение $s \equiv (rd + ke) \pmod{q}$. Если $s = 0$, то вернуться к шагу 3;

Подпись для сообщения M составит пара чисел (r, s) .

Для проверки цифровой подписи ζ под полученным сообщением M необходимо выполнить следующие действия (рис. 1.12б):

1. Если выполнены неравенства $0 < r < q$, $0 < s < q$, то перейти к следующему шагу. В противном случае подпись неверна;
2. вычислить хэш-значение полученного сообщения M : $\bar{h} = h(M)$;
3. вычислить целое число α , двоичным представлением которого является вектор \bar{h} и определить $e \equiv \alpha \pmod{q}$. Если $e = 0$, то определить $e = 1$;
4. вычислить значение $v = e^{-1} \pmod{q}$;
5. вычислить значения $z_1 \equiv sv \pmod{q}$, $z_2 \equiv -rv \pmod{q}$;
6. вычислить точку эллиптической кривой $C = z_1P + z_2Q$ и определить $R \equiv x_c \pmod{q}$, где x_c - x -координата точки C ;
7. если выполнено равенство $R = r$, то подпись принимается, в противном случае, подпись неверна.

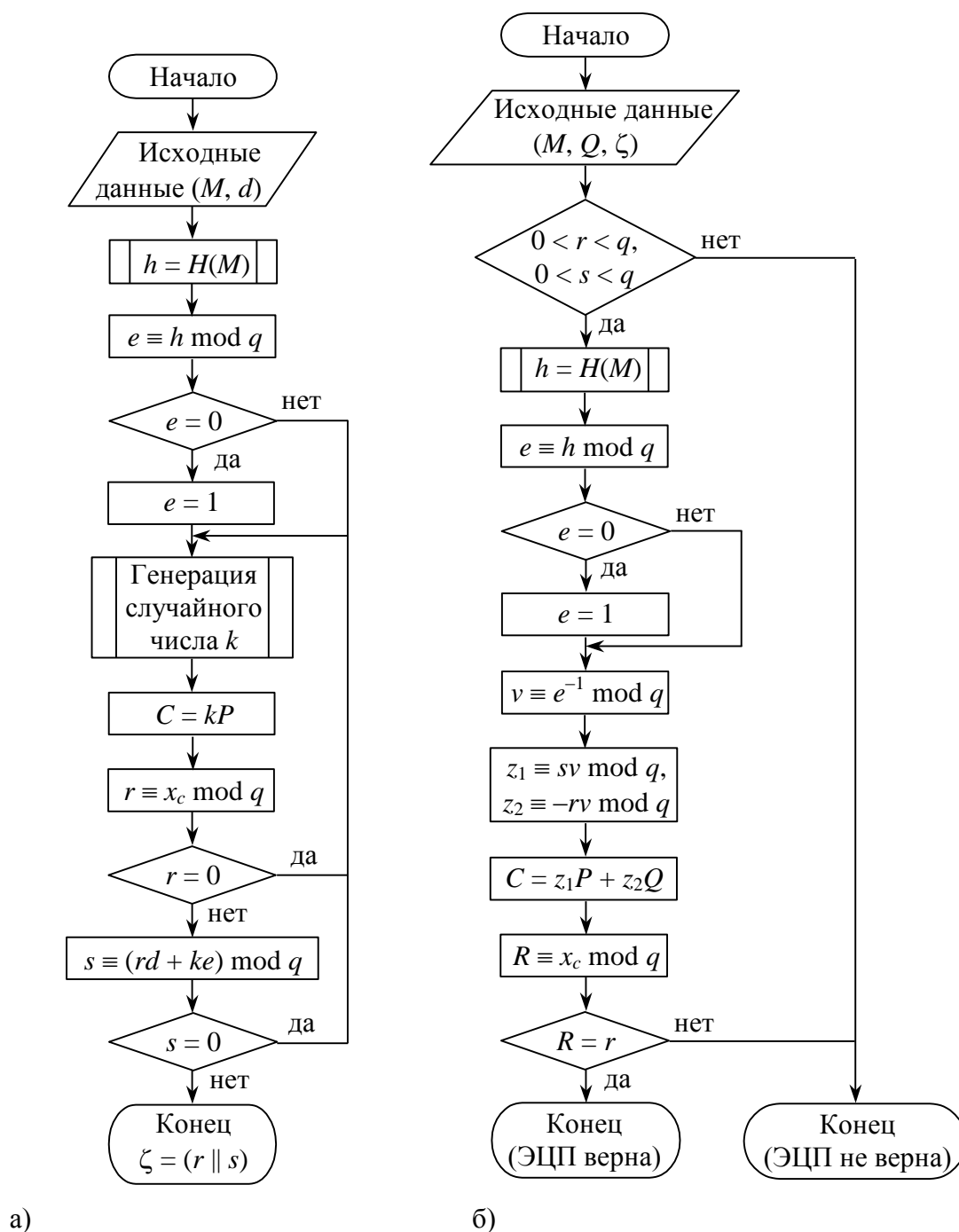


Рис. 8.12. Процесс формирования цифровой подписи (а), процесс проверки цифровой подписи (б).

Интерфейс учебного программного комплекса DigitSign

На рисунке 8.13 приведен интерфейс главного окна учебного программного комплекса.

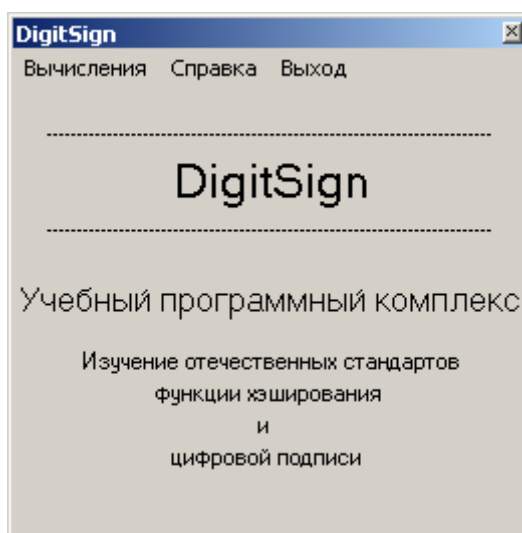


Рис. 8.13. Интерфейс главного окна учебного программного комплекса

Пункт меню “Вычисления” содержит следующие опции:

- Вычисление хэш-функции.

При выборе данной опции открывается окно “Вычисление хэш-функции” (рисунок 8.15), которое позволяет определить хэш-значение произвольного файла при заданных пользователем начальном хэш-векторе и таблице замен. Полученный результат может быть сохранен в файл. Также пользователь получает информацию о времени хэширования сообщения.

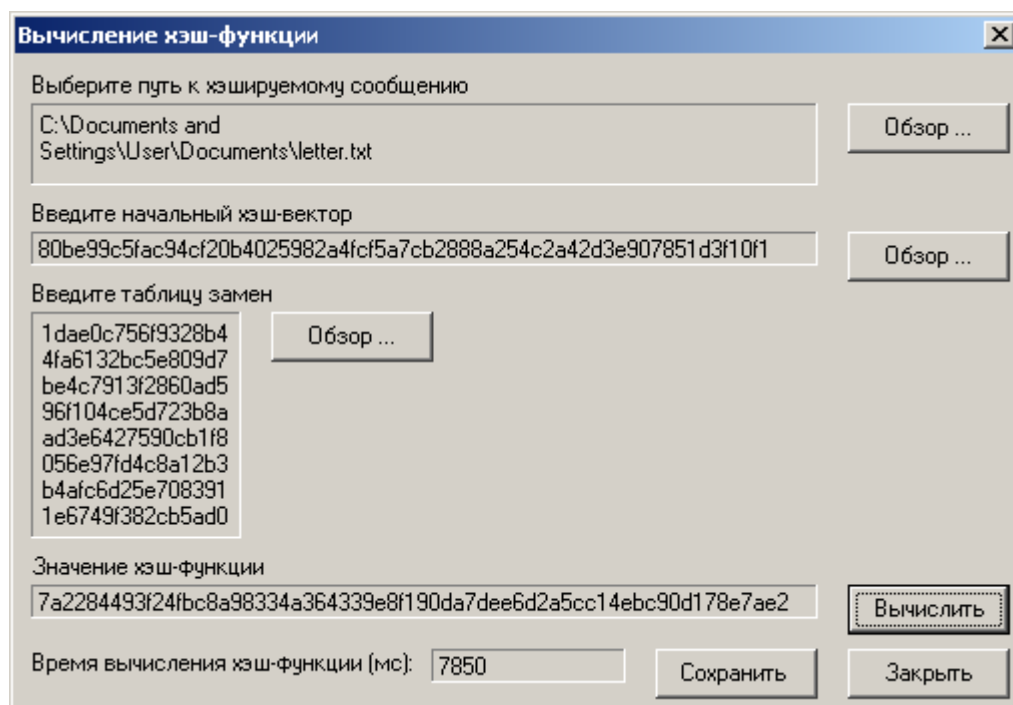


Рис. 8.14. Окно “Вычисление хэш-функции”

- Генератор псевдослучайных чисел.

При выборе данной опции открывается окно “Генератор псевдослучайных чисел” (рисунок 3.16).

Порядок выполнения работы

Во время выполнения лабораторной работы делайте скриншоты. После вставить в отчет.

1. Ознакомьтесь с теорией по функциям хэширования, эллиптическим кривым и цифровым подписям.
2. Ознакомьтесь с ГОСТ Р 34.11-94 и ГОСТ Р 34.10-2001.
3. Откройте папку DigitSign и запустите программу DigitSign.exe.
4. Изучение функции хэширования по ГОСТ Р 34.11-94.

4.1 Открываем пункт меню «Вычисления» и выбираем «Вычисление хэш-функции».

4.2 Задать начальные данные для вычисления хэш-функции (файлы находятся в папке «Хэширование»):

- путь к файлу, который подвергается сжатию (название text0.txt);
- начальный хэш-вектор (HeshVHex.txt);
- таблицу замен (TabZamen.txt).

4.3 Вычислить хэш-функцию.

4.4 Повторить пункты 4.2 – 4.3 для файлов с названием text1, text2, text3, text4, text5. Сохранить полученные значения хэш-функции в папку «Временное», назвать hash1.txt, hash2.txt, hash3.txt, hash4.txt, hash5.txt соответственно. Определить время хэширования. Полученные результаты занести в первые две строки таблицы (l – размер текстового файла, t_h – время хэширования, для text0 делать данный пункт не обязательно). **ВНИМАНИЕ**, хэширование может производиться до 3х минут, лучше в этот момент не чего не делать на компьютере, иначе возможно получение недостоверных значений.

l , Мбайт					
t_h , мс					
$t_{\text{ЭЦП}}$, мс					
t_{Σ} , мс					

5. Изучение цифровой подписи ГОСТ Р 34.10-2001.

5.1 Вычислить секретный ключ.

5.1.1 Открываем пункт меню «Вычисления» и выбираем «Генератор псевдослучайных чисел».

5.1.2 Ввести параметры датчика (файлы находятся в папке «ГПСЧ»):

- начальное значение X_0 (X0.txt);

- множитель a (a.txt);
- приращение c (c.txt);
- модуль m (m.txt).

5.1.3 Вычислить псевдослучайное число, оно же является секретным ключом.

5.1.4 Сохранить полученное значение в папку «Временное», назвав «KeySkr.txt».

5.2 Вычислить открытый ключ.

5.2.1. Открываем пункт меню «Вычисления» и выбираем «Вычисление открытого ключа».

5.2.2. Задать параметры, необходимые для вычисления (файлы находятся в папке «Открытый ключ»).

- Модуль эллиптической кривой p (modP.txt);
- Коэффициент эллиптической кривой a (kofA.txt);
- Координаты точки P (координата x – Px.txt, координата y – Py.txt);
- Секретный ключ d (ранее полученный в пункте 5.1.4).

5.2.3. Вычислить открытый ключ.

5.2.4. Сохранить полученное значение в папку «Временное», назвав «Qx.txt» для координаты X и «Qy.txt» для координаты Y .

5.3 Вычислить цифровую подпись.

5.3.1. Выполнить предварительное вычисление случайного числа k , необходимого для формирования цифровой подписи. Для этого следует выполнить п. 5.1. Сохранить результат в папку «Временное», назвав «psK.txt».

5.3.2. Открываем пункт меню «Вычислить» и выбираем «Формирование цифровой подписи».

5.3.3. Задать параметры, необходимые для вычисления (файлы находятся в папке «Открытый ключ»):

- Модуль эллиптической кривой p (modP.txt);
- Коэффициент эллиптической кривой a (kofA.txt);
- Порядок циклической подгруппы эллиптической кривой q (q.txt);

- Координаты точки P (координата X – Px.txt, координата Y – Py.txt);
- Секретный ключ d (ранее полученный в пункте 5.1.4);
- Хэш сообщения (ранее полученное в пункте 4.4, с названием hash1.txt);
- Псевдо случайное число k (ранее полученное в пункте 5.4.1).

5.3.4. Сформировать цифровую подпись и занести в таблицу время создания ЭЦП (строка $t_{\text{ЭЦП}}$).

5.3.5. Сохранить значение ЭЦП в папку «Временное», назвав «podpis.txt»

5.3.6. Повторить пункты 5.4.3-5.4.4 для хэш-значений hash2.txt; hash3.txt; hash4.txt; hash5.txt. Обратите внимание на то, что пункт 5.4.5 выполнять повторно не нужно.

5.3.7. Рассчитать суммарное время вычисления ЭЦП и занести результаты расчетов в таблицу ($t_{\Sigma} = t_h + t_{\text{ЭЦП}}$).

5.4 Передайте другому студенту, который сидит за другим компьютером, открытый ключ (Qx.txt; Qy.txt), файл (text1.txt) и соответствующую ему цифровую подпись (podpis.txt) при помощи дискеты или сети и получите от него такой же набор. Полученные файлы переименовать:

- «Qx.txt» в «QxPol.txt»;
- «Qy.txt» в «QyPol.txt»;
- «text1.txt» в «text1Pol.txt»;
- «podpis.txt» в «podpisPol.txt»;
- И поместить в папку «Временное».

5.5 Проверить цифровую подпись.

5.5.1 Сделать пункт 4, для файла text1Pol.txt, результат сохранить в папку «Временное», назвав «hashPol.txt»;

5.5.2 Открываем пункт меню «Вычислить», и выбираем «Проверка цифровой подписи».

5.5.3 Задать параметры, необходимые для вычисления (файлы находятся в папке «Открытый ключ»):

- Модуль эллиптической кривой p (modP.txt);
- Коэффициент эллиптической кривой a (kofA.txt);

- Порядок циклической подгруппы эллиптической кривой q (q.txt);
- Координаты точки P (координата X – «Px.txt», Y – «Py.txt»);
- Открытый ключ Q (полученный в пункте 5.4, координата X – «QxPol.txt»; координата Y – «QyPol.txt»);
- Хэш сообщения (полученный в пункте 5.4.1, «hashPol.txt»);
- ЭЦП (полученное в пункте 5.4, «podpisPol.txt»).

5.5.4 Осуществить проверку.

ВНИМАНИЕ! Данная программа неверно вычисляет открытый ключ Q , координату y . Если «Подпись не верна»: открываете файл QyPol.txt через блокнот и добавляете к **последней** цифре единицу (программа все время выдает последнюю цифру равную 0) и снова проверяете, если опять неверна, то еще добавляете единицу до тех пор пока подпись не будет верна или пока не достигнете f и подпись по прежнему не верна, тогда подпись действительно не верна (программа использует 16-ти ричную систему счисления, {0; 1; 2; 3; 4; 5; 6; 7; 8; 9; a; b; c; d; e; f}).

5.6 Подделка сообщения.

5.6.1. Откройте файл «text1Pol.txt» (полученный в пункте 5.4) с помощью блокнота и добавьте, удалите или измените любой символ в тексте, после сохраните изменение в этот же файл.

5.6.2. Повторите пункт 5.5.

6. По полученным результатам построить два графика: на первом графике зависимость времени вычисления хэш-значения от размера файла $t_h = f(l)$, на втором графике зависимость $t_{\text{ЭЦП}} = f(l)$.
7. По проделанной работе сделать выводы.
8. Оформить отчет по проделанной работе.
9. Очистить папку «Временное».

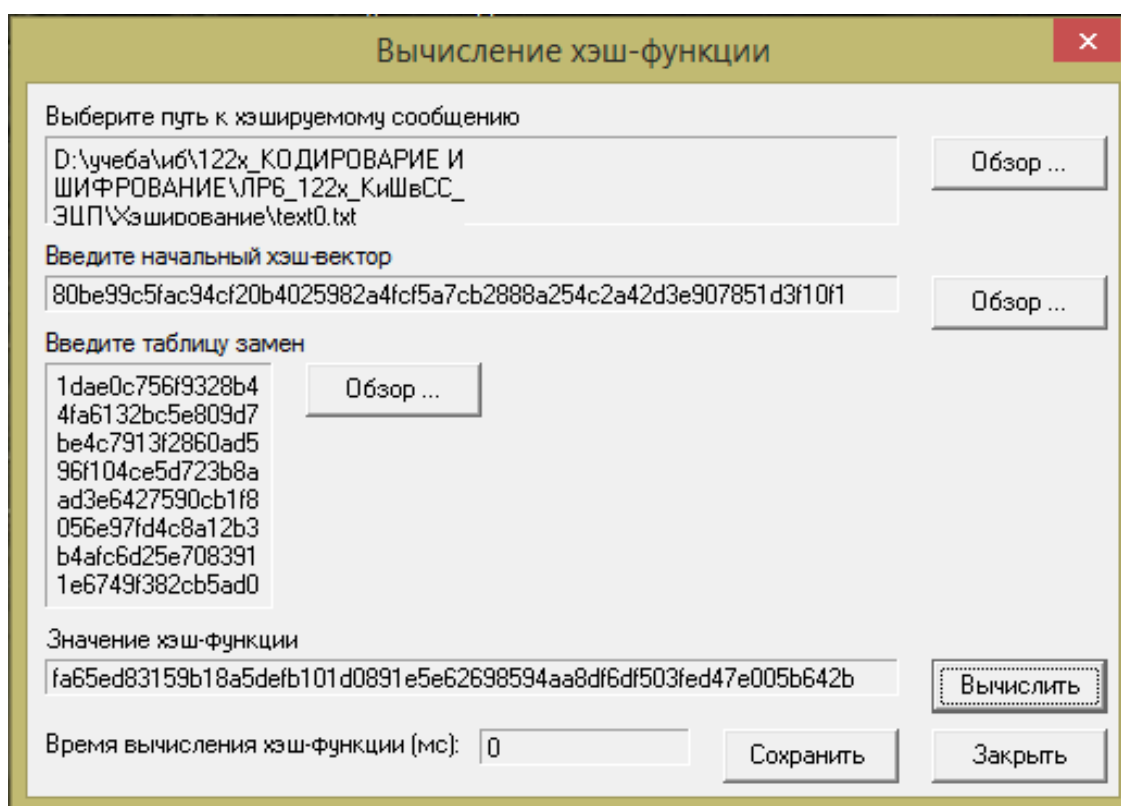
Контрольные вопросы

1. Что такое асимметричная криптографическая система? В чем ее преимущества перед симметричной системой? В чем недостатки?
2. Что такое хэш-функция? Какими свойствами она должна обладать и в чем они заключаются?
3. Опишите процесс формирования хэш-функции ГОСТ Р 34.11-94.

4. Что такое “функция сжатия внутренних итераций”? В чем она заключается?
5. Что такое цифровая подпись? В чем ее отличие от рукописной подписи?
6. Опишите процесс формирования цифровой подписи ГОСТ Р 34.10-2001.
7. Опишите процесс проверки цифровой подписи ГОСТ Р 34.10-2001.
8. От каких злоумышленных действий позволяет защититься ЭЦП?

Выполнение работы

10. Изучение функции хэширования по ГОСТ Р 34.11-94.
 1. Открываем пункт меню «Вычисления» и выбираем «Вычисление хэш-функции».
 2. Задать начальные данные для вычисления хэш-функции (файлы находятся в папке «Хэширование»):
 3. путь к файлу, который подвергается сжатию (название text0.txt);
 4. начальный хэш-вектор (HeshVHex.txt);
 5. таблицу замен (TabZamen.txt).
 6. Вычислить хэш-функцию.



7. Повторить пункты 4.2 – 4.3 для файлов с названием text1, text2, text3, text4, text5. Сохранить полученные значения хэш-функции в папку «Временное», назвать hash1.txt, hash2.txt, hash3.txt, hash4.txt, hash5.txt соответственно.

5.3.5. Сохранить значение ЭЦП в папку «Временное», назвав «podpis.txt»

5.3.6. Повторить пункты 5.4.3-5.4.4 для хэш-значений hash2.txt; hash3.txt; hash4.txt; hash5.txt. Обратите внимание на то, что пункт 5.4.5 выполнять повторно не нужно.

5.3.7. Рассчитать суммарное время вычисления ЭЦП и занести результаты расчетов в таблицу ($t_{\Sigma} = t_h + t_{\text{ЭЦП}}$).

5.6 Передайте другому студенту, который сидит за другим компьютером, открытый ключ (Qx.txt; Qy.txt), файл (text1.txt) и соответствующую ему цифровую подпись (podpis.txt) при помощи дискеты или сети и получите от него такой же набор. Полученные файлы переименовать:

- «Qx.txt» в «QxPol.txt»;
- «Qy.txt» в «QyPol.txt»;
- «text1.txt» в «text1Pol.txt»;
- «podpis.txt» в «podpisPol.txt»;
- И поместить в папку «Временное».

5.7 Проверить цифровую подпись.

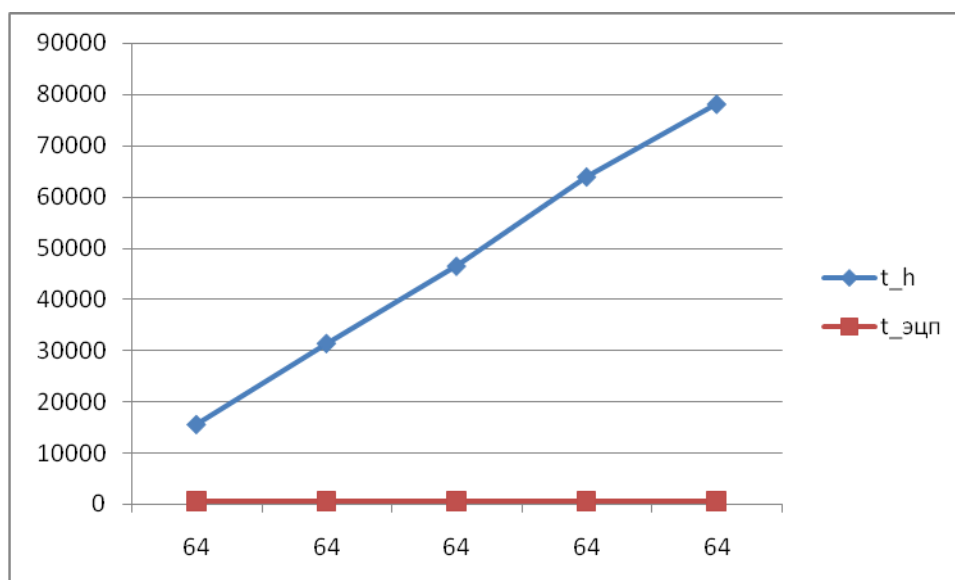
5.5.5 Сделать пункт 4, для файла text1Pol.txt, результат сохранить в папку «Временное», назвав «hashPol.txt»;

5.5.6 Открываем пункт меню «Вычислить», и выбираем «Проверка цифровой подписи».

5.5.7 Задать параметры, необходимые для вычисления (файлы находятся в папке «Открытый ключ»):

- Модуль эллиптической кривой p (modP.txt);
- Коэффициент эллиптической кривой a (kofA.txt);
- Порядок циклической подгруппы эллиптической кривой q (q.txt);
- Координаты точки P (координата X – «Px.txt», Y – «Py.txt»);
- Открытый ключ Q (полученный в пункте 5.4, координата X – «QxPol.txt»; координата Y – «QyPol.txt»);
- Хэш сообщения (полученный в пункте 5.4.1, «hashPol.txt»);
- ЭЦП (полученное в пункте 5.4, «podpisPol.txt»).

5.5.8 Осуществить проверку.



Вывод: В результате выполнения данной лабораторной работы мы изучили процессы вычисления функции хэширования по ГОСТ Р 34.11-94, формирования и проверки ЭЦП на эллиптических кривых по ГОСТ Р 34.10-2001.

8.3.2. PGP кодирование и шифрование с открытым ключом

PGP (англ. Pretty Good Privacy) — компьютерная программа, также библиотека функций, позволяющая выполнять операции шифрования и цифровой подписи сообщений, файлов и другой информации, представленной в электронном виде, в том числе прозрачное шифрование данных на запоминающих устройствах, например, на жёстком диске.

Первоначально разработана Филиппом Циммерманном в 1991 году.

Общие сведения

PGP имеет множество реализаций, совместимых между собой и рядом других программ (GnuPG, FileCrypt и др.) благодаря стандарту OpenPGP (RFC 4880), но имеющих разный набор функциональных возможностей. Существуют реализации PGP для всех наиболее распространённых операционных систем. Кроме свободно распространяемых реализаций есть еще и коммерческие.

Совместимость

Так как PGP развивается, некоторые системы позволяют создавать зашифрованные сообщения с использованием новых возможностей, которые отсутствуют в старых системах. Отправитель и получатель должны знать возможности друг друга или, по крайней мере, согласовать настройки PGP.

Защищённость

В 1996 году криптограф Брюс Шнайер охарактеризовал раннюю версию PGP как «ближайшую к криптосистемам военного уровня». На данный момент не известно ни одного способа взлома данных, зашифрованных PGP, при помощи полного перебора или уязвимости криптоалгоритма. Ранние версии PGP обладали теоретическими уязвимостями, поэтому рекомендуется пользоваться современными версиями.

Криптографическая стойкость PGP основана на предположении, что используемые алгоритмы устойчивы к криптоанализу на современном оборудовании. Например, в PGP первых версий для шифрования ключей сессии использовался алгоритм RSA, основанный на односторонней функции (факторизация). В PGP версии 2 дополнительно можно использовать алгоритм IDEA. В последующем были добавлены дополнительные алгоритмы шифрования. Ни у одного используемого алгоритма нет известных уязвимостей.

В 2010 году группе учёных из Швейцарии, Японии, Франции, Нидерландов, Германии и США удалось декодировать данные, зашифрованные по алгоритму RSA при помощи ключа длиной 768 бит. Нахождение простых сомножителей осуществлялось общим методом решета числового поля. На первый шаг (выбор пары полиномов степени 6 и 1) было потрачено около полугода вычислений на 80 процессорах, что составило около 3 % времени, потраченного на главный этап алгоритма (просеивание), который выполнялся на сотнях компьютеров в течение почти двух лет. Если интерполировать это время на работу одного процессора AMD Opteron 2.2ГГц с 2Гб памяти, то получилось бы порядка 1500 лет. Обработка данных после просеивания для следующего ресурсоёмкого шага (линейной алгебры) потребовала несколько недель на малом количестве процессоров. Заключительный шаг после нахождения нетривиальных решений ОСЛУ занял не более 12 часов.

Решение ОСЛУ проводилось с помощью метода Видемана на нескольких отдельных кластерах и длилось чуть менее 4 месяцев. При этом размер разреженной матрицы составил $192\,796\,550 \times 192\,795\,550$ при наличии $27\,795\,115\,920$ ненулевых элементов. Для хранения матрицы на жёстком диске понадобилось около 105 гигабайт. В то же время понадобилось около 5 терабайт сжатых данных для построения данной матрицы.

В итоге группе удалось вычислить 232-цифровой ключ, открывающий доступ к зашифрованным данным.

Исследователи уверены, что с использованием их метода факторизации взломать 1024-битный RSA-ключ будет возможно в течение следующего десятилетия.

По словам исследователей, после их работы в качестве надежной системы шифрования можно рассматривать только RSA-ключи длиной 1024 бита и более. Причём от шифрования ключом длиной в 1024 бит стоит отказаться в ближайшие три-четыре года.

Зная разложение модуля на произведение двух простых чисел, противник может легко найти секретную экспоненту и тем самым взломать RSA. Однако на сегодняшний день самый быстрый алгоритм факторизации — решето обобщённого числового поля (General Number Field Sieve), скорость которого для k -битного целого числа составляет $\exp((c + o(1))k^{\frac{1}{3}} \log^{\frac{2}{3}} k)$ для некоторого $c < 2$, не позволяет разложить большое целое за приемлемое время.

Механизм работы PGP

Шифрование PGP осуществляется последовательно хешированием, сжатием данных, шифрованием с симметричным ключом, и, наконец, шифрованием с открытым ключом, причём каждый этап может осуществляться одним из нескольких поддерживаемых алгоритмов. Симметричное шифрование производится с использованием одного из семи симметричных алгоритмов (AES, CAST5, 3DES, IDEA, Twofish, Blowfish, Camellia) на сеансовом ключе. Сеансовый ключ генерируется с использованием криптографически стойкого генератора псевдослучайных чисел. Сеансовый ключ зашифровывается открытым ключом получателя с использованием алгоритмов RSA или Elgamal (в зависимости от типа ключа получателя). Каждый открытый ключ соответствует имени пользователя или адресу электронной почты. Первая версия системы называлась Сеть Доверия и противопоставлялась системе X.509, использовавшей иерархический подход, основанной на удостоверяющих центрах, добавленный в PGP позже. Современные версии PGP включают оба способа.

Ключи

Пользователь PGP создаёт ключевую пару: открытый и закрытый ключ. При генерации ключей задаются их владелец (имя и адрес электронной почты), тип ключа, длина ключа и срок его действия. Открытый ключ используется для шифрования и проверки цифровой подписи. Закрытый ключ - для декодирования и создания цифровой подписи.

PGP поддерживает три типа ключей RSA v4, RSA legacy (v3) и Diffie-Hellman/DSS (Elgamal в терминологии GnuPG).

Для ключей RSA legacy длина ключа может составлять от 1024 до 2048 бит, а для Diffie-Hellman/DSS и RSA — от 1024 до 4096. Ключи RSA legacy содержат одну ключевую пару, а ключи Diffie-Hellman/DSS и RSA могут содержать один главный ключ и дополнительные ключи для шифрования. При этом ключ электронной подписи в ключах Diffie-Hellman/DSS всегда имеет размер 1024. Срок действия для каждого из типов ключей может быть определён как неограниченный или до конкретной даты. Для защиты ключевого контейнера используется секретная фраза.

Цифровая подпись

PGP поддерживает аутентификацию и проверку целостности посредством цифровой подписи. По умолчанию она используется совместно с шифрованием, но также может быть применена и к открытому тексту. Отправитель использует PGP для создания подписи алгоритмом RSA или DSA. При этом сначала создаётся хеш открытого текста (также известный как дайджест), затем — цифровая подпись хеша при помощи закрытого ключа отправителя. Для формирования хеша могут использоваться алгоритмы MD5, SHA-1, RIPEMD-160, SHA-256, SHA-384, SHA-512. В новых версиях PGP поддержка MD5 осуществляется для сохранения совместимости с ранними версиями. Для подписи используются алгоритмы RSA или DSA (в зависимости от типа ключа).

Сжатие данных

В целях уменьшения объёма сообщений и файлов и, возможно, для затруднения криптоанализа PGP производит сжатие данных перед шифрованием. Сжатие производится по одному из алгоритмов ZIP, ZLIB, BZIP2. Для сжатых, коротких и слабосжимаемых файлов сжатие не выполняется.

Сеть доверия

Как при шифровании сообщений, так и при проверке цифровой подписи, необходимо, чтобы принятый получателем открытый ключ действительно принадлежал отправителю. При простом скачивании открытого ключа он может быть подменён. С первых версий PGP поддерживает сертификаты открытых ключей, с помощью которых подмены (или случайные ошибки передачи) легко распознаются. Однако недостаточно просто создать сертификат, защищённый от модификации, так как при этом гарантируется лишь целостность сертификата после его создания. Пользователи также должны каким-нибудь способом проверить, что открытый ключ в сертификате действительно принадлежит отправителю. С первых версий продукты PGP включают в себя внутреннюю схему проверки сертификатов, названную сетью доверия. Заданная пара «имя пользователя — открытый ключ» может быть подписана третьим лицом, удостоверяющим соответствие ключа и владельца. В таких подписях может быть несколько вложенных уровней доверия. Хотя многие программы читают и пишут эту информацию, очень немногие учитывают этот уровень сертификата, принимая решение о принятии или отклонении сертификата.

Протокол сети доверия был впервые описан Циммерманном в 1992 году в руководстве PGP версии 2.0: «С течением времени вы будете накапливать ключи других людей, которых вы можете назвать доверенными рекомендателями. Кто-нибудь ещё может выбрать своих доверительных рекомендателей. И все будут постепенно накапливать и распространять со своими ключами набор заверенных подписей других людей, ожидая, что любой получатель

доверяет по крайней мере одной или двум подписям. Это позволяет создать децентрализованную устойчивую к сбоям сеть всех открытых ключей.»

Механизм сети доверия обладает преимуществами над централизованной инфраструктурой управления открытыми ключами, например, используемой в S/MIME, но не получил повсеместного применения. Пользователи хотели проверять корректность сертификатов вручную или не проверять вовсе.

Сертификаты

В последних спецификациях OpenPGP доверенные подписи могут использоваться для поддержки создания центров сертификации. Доверенность сертификата означает, что ключ действительно принадлежит указанному владельцу и может использоваться для подписи сертификатов одним уровнем ниже. Сертификат уровня 0 означает обычную подпись. Уровень 1 означает, что при помощи подписанного ключа можно создавать сертификаты уровня 0. При помощи сертификата уровня 2 можно создавать сертификаты уровня 1. Уровень 2 практически идентичен степени доверия, с которой полагаются пользователи на списки доверенных сертификатов, встроенные в браузеры.

Все версии PGP включают в себя способ отмены сертификата. Это необходимо, если требуется сохранять безопасность связи при потере или компрометации закрытого ключа. Отмена сертификата похожа на списки отзыва сертификатов в централизованной инфраструктуре открытых ключей. Современные версии PGP также поддерживают сроки истечения сертификатов.

Проблема корректного определения принадлежности открытого ключа владельцу характерна для всех криптографических систем с асимметричным шифрованием. У неё не существует достаточно хороших решений. Оригинальная схема PGP позволяет решить пользователю, использовать ли схему проверки сертификатов, в то время как большинство других инфраструктур открытых ключей требуют проверки каждого сертификата.

История

В 1991 году Филипп Циммерман создал первую версию PGP. Первая версия включала в себя симметричный алгоритм шифрования BassOmatic, созданный самим Циммерманом. Циммерман участвовал в движении против ядерной энергии и создал PGP для защищённого использования BBS и хранения файлов и сообщений. Для некоммерческого использования не требовалось лицензии, со всеми копиями распространялся весь исходный код. PGP распространилась в Usenet, а затем и в Интернете.

Уголовное расследование

Вскоре после выпуска PGP стала использоваться за пределами США, и в 1993 году правительство США начало расследование против Циммермана по подозрению в

нарушении экспортного законодательства, которое регулирует распространение криптографических систем с длиной ключа более 40 бит. В PGP использовались ключи длиной 128 бит и более.

Циммерман остроумно обошёл ограничения законодательства США. Он опубликовал исходный код в книге, изданной MIT Press. Код можно было сосканировать, распознать и скомпилировать. Экспорт книг не может быть запрещён, так как защищён первой поправкой к Конституции США.

OpenPGP

PGP Inc. была обеспокоена по поводу патентов. В компании был создан внутренний стандарт Unencumbered PGP («необременённый PGP»), не использующий алгоритмы, имеющие проблемы с лицензиями. Так как PGP широко использовалась во всём мире, многие хотели создавать собственное ПО, совместимое с PGP 5. В 1997 году PGP Inc. предложила IETF стандарт, названный OpenPGP. В IETF были созданы стандарты RFC 2440 (1998 год) и RFC 4880 (2007 год).

В 1999 году силами Фонда свободного программного обеспечения была создана свободная реализация OpenPGP под названием GNU Privacy Guard (GnuPG).

Поглощение Network Associates

В декабре 1997 года PGP Inc. была поглощена Network Associates Inc (ныне McAfee). NAI продолжила экспорт посредством печати исходных текстов. В составе NAI команда PGP разработала средства для шифрования дисков, брандмауэр, средства для обнаружения вторжений и IPsec VPN. После легализации экспорта криптографического ПО в 2000 году NAI прекратила публикацию исходных текстов, несмотря на возражения команды PGP.

В 2001 году Циммерман покинул NAI, NAI объявила о продаже PGP и остановке разработки PGP. В 2002 году NAI прекратила поддержку всех продуктов PGP PGP E-Business Server (исходной консольной версии PGP).

Современное состояние

В 2002 году несколько бывших разработчиков PGP основали PGP Corporation и выкупили PGP (кроме консольной версии). В 2003 году PGP Corporation разработала новый серверный продукт, PGP Universal.

В 2010-м году Symantec Corp. выкупил PGP за 300 млн долларов.

Криптографические приложения PGP Corporation

PGP изначально разрабатывалась для шифрования электронной почты на стороне клиента, но с 2002 года включает также шифрование жёстких дисков переносных компьютеров, файлов и директорий, сессий программ мгновенного обмена сообщениями, пакетной передачи файлов, защиту файлов и директорий в сетевых хранилищах, а в

современных версиях — ещё и шифрование HTTP-запросов и ответов на стороне сервера (mod openpgp) и клиента (Enigform).

Клиентские программы объединены в семейство PGP Desktop (включает в себя PGP Desktop EMail, PGP Whole Disk Encryption и PGP NetShare).

PGP Universal Server позволяет из командной строки централизованно администрировать клиенты на основе PGP Desktop.

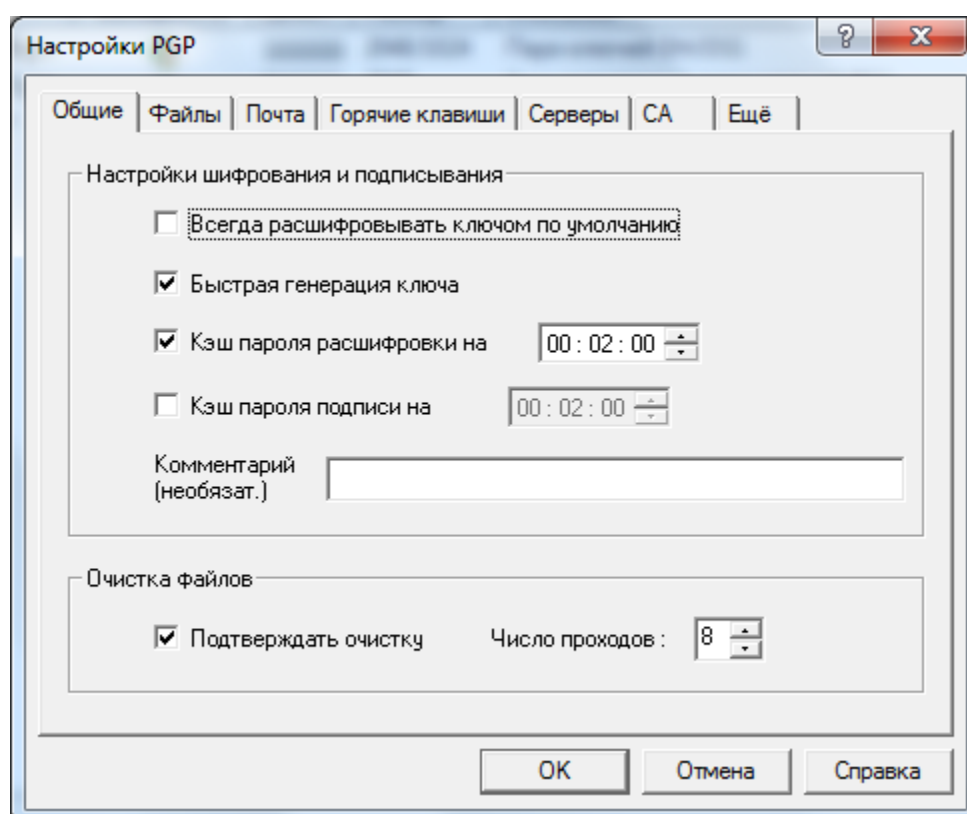
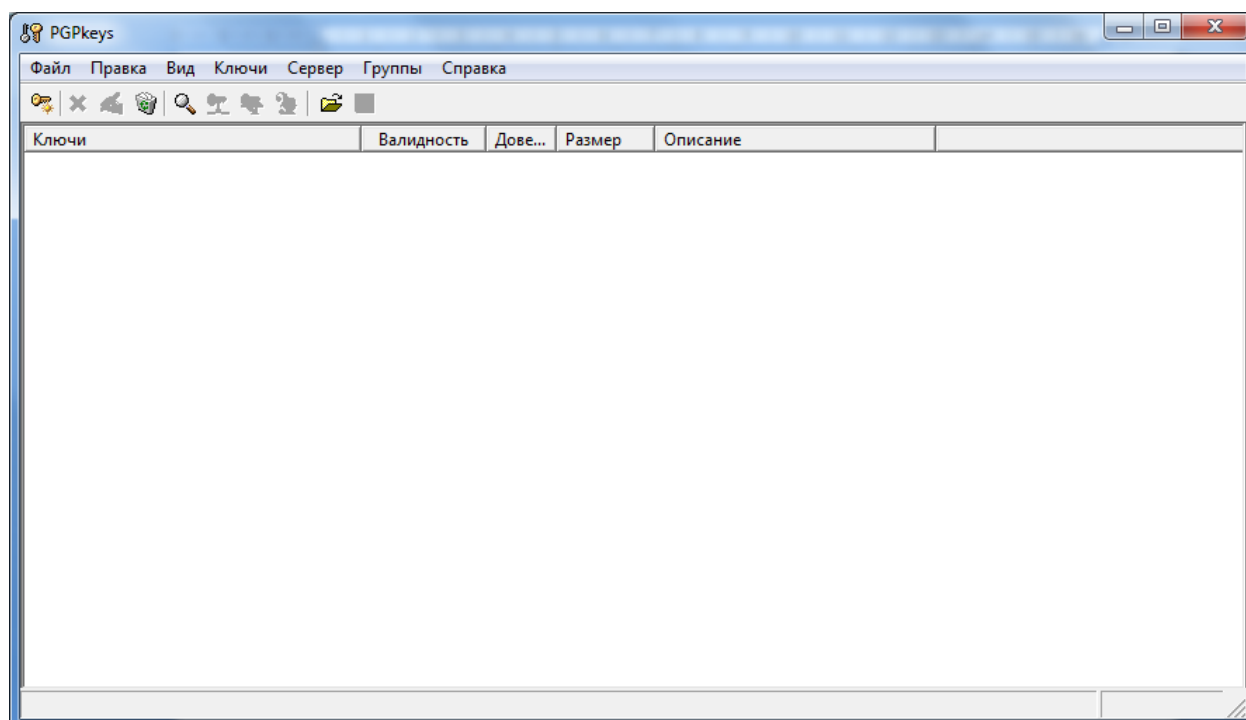
В 2010 году права на приложение были приобретены компанией Symantec за 300 млн. долларов.

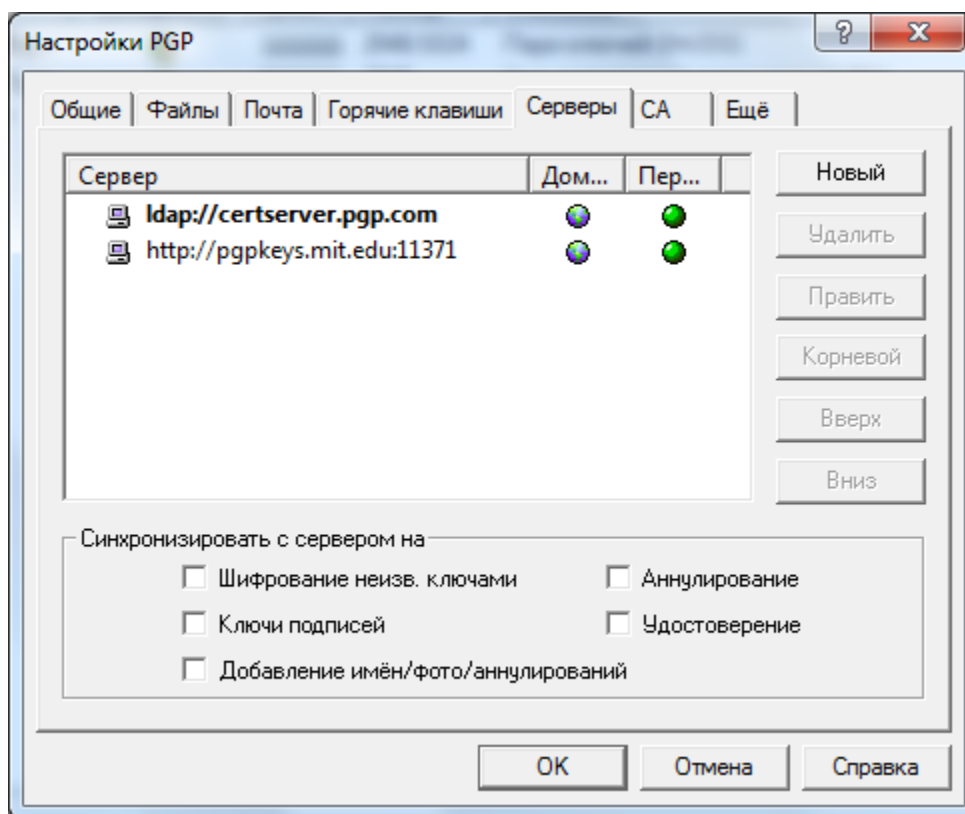
Правовые аспекты использования в России

На сегодняшний день прямых законодательных запретов на использование PGP в России нет. Законодательно ограничивается использование криптографии только в государственных и муниципальных учреждениях. ФСБ предписывает всем государственным структурам использовать только сертифицированные средства криптографии. Физические лица и компании сами устанавливают, какая информация является для них коммерческой тайной, методы хранения и передачи такой информации. Закон «Об информации, информационных технологиях и защите информации» также указывает, что способ защиты информации, представляющей тайну, для негосударственных структур определяется оператором. Информационный ресурс Helpdesk24 в статье «Правомерность использования криптографических средств защиты информации» приводит выдержки из федеральных законов, поясняющие данный вопрос. Также авторы проекта «openPGP в России» утверждают, что не существует законов, запрещающих использование PGP. Указ от 3 апреля 1995 г. N 334 «О мерах по соблюдению законности в области разработки, производства, реализации и эксплуатации шифровальных средств» отменен. Электронная подпись, генерируемая с помощью PGP и её несертифицированных аналогов, имеет юридическую силу в Российской Федерации, т.к. согласно пункту 3 статьи 5 63-ФЗ "Об электронной подписи" попадает под определение усиленной неквалифицированной электронной подписи. Согласно пункту 2 статьи 6 этого ФЗ для признания такой ЭП необходимо соглашение между участниками электронного взаимодействия.

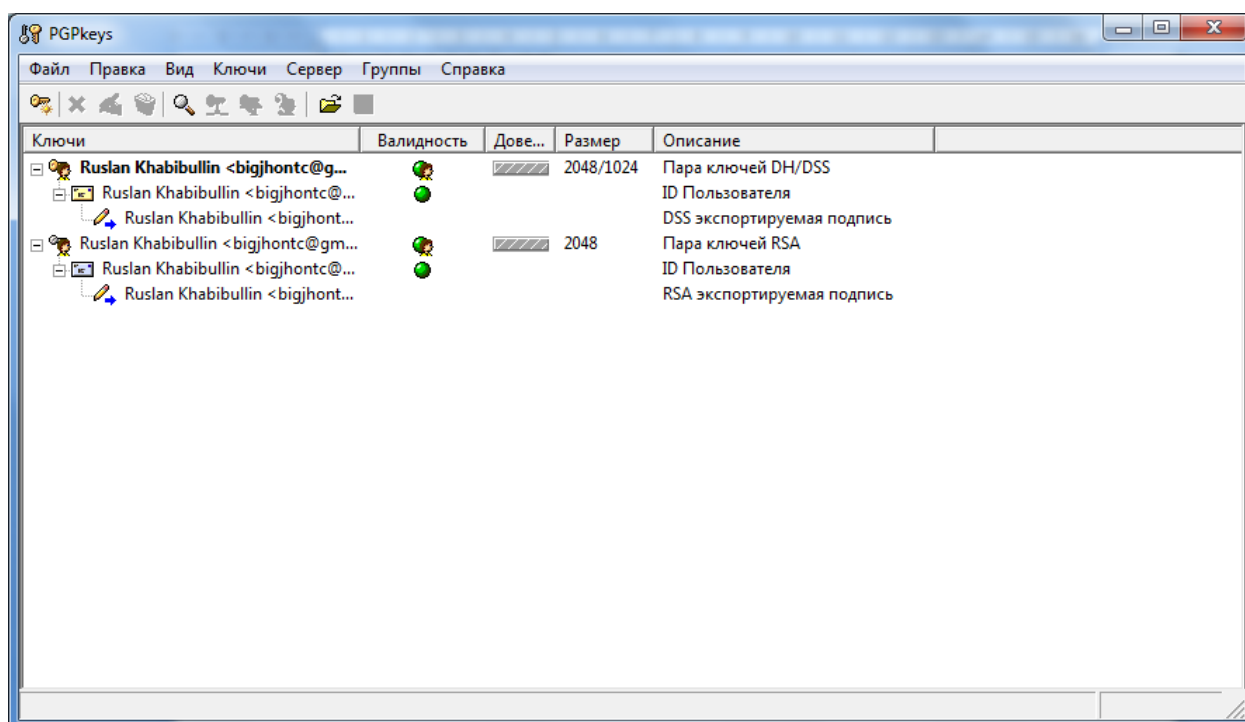
2.ХОД РАБОТЫ

1. Изучите вкладки окна “Настройки PGP”.

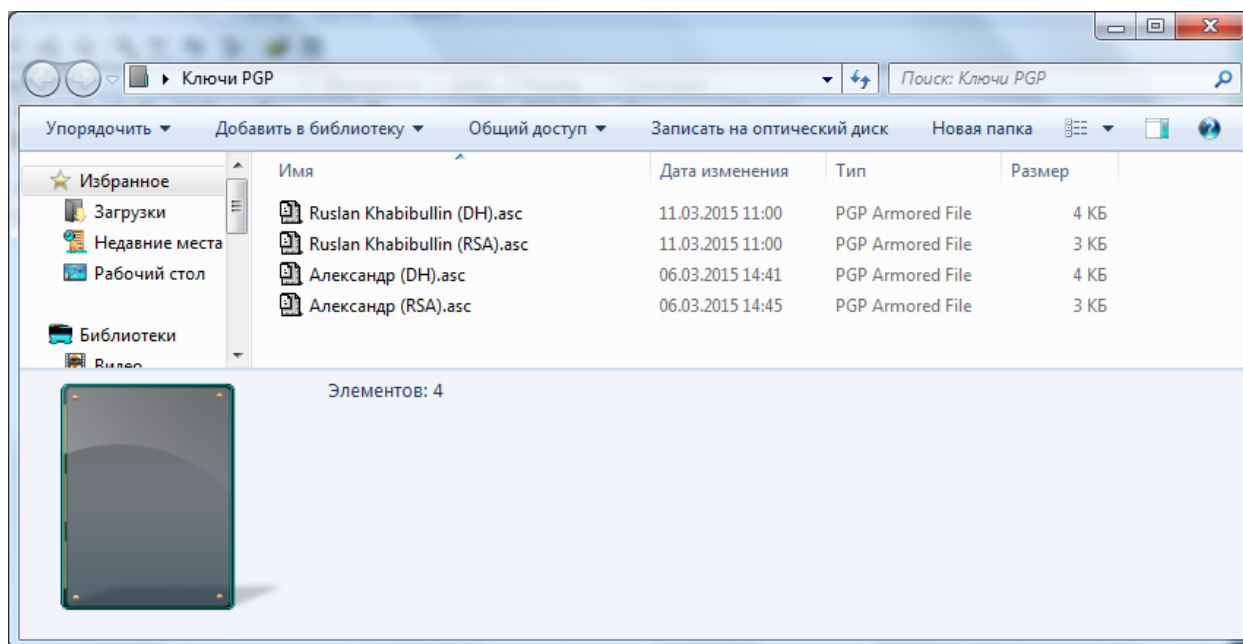




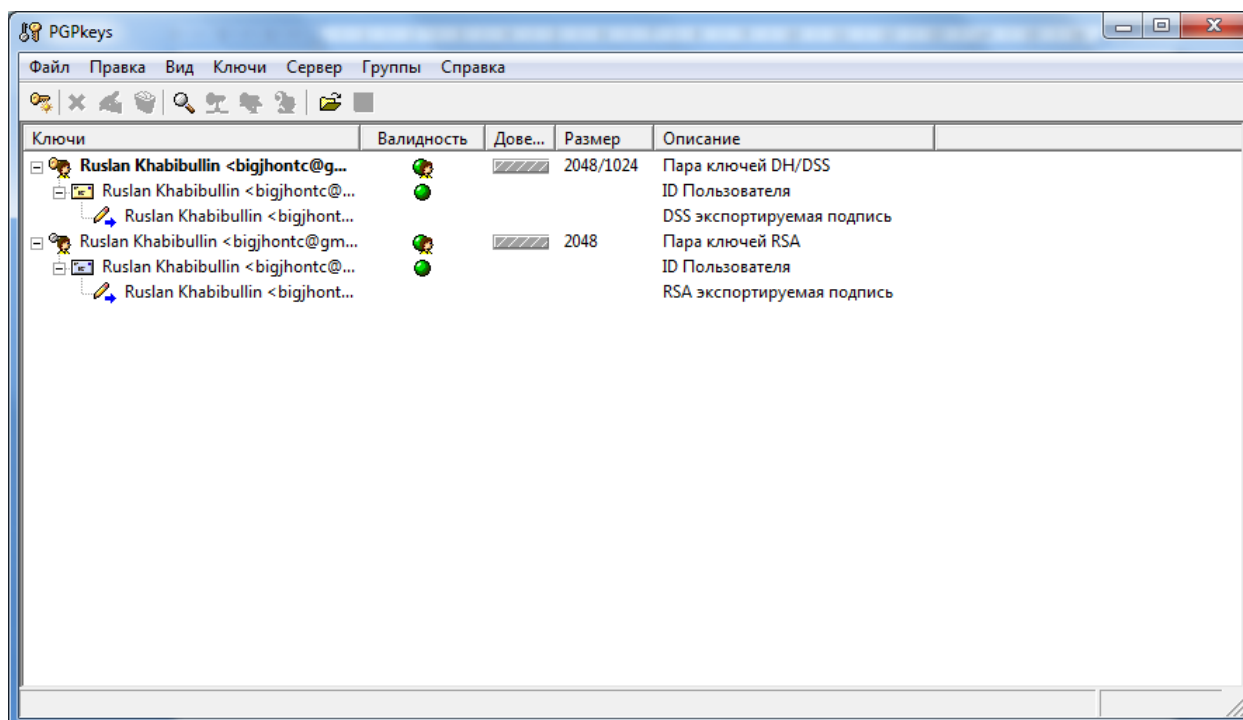
2. Создайте с помощью менеджера PGPkeys ключи шифрования двух типов: RSA и DH/DSS.



3. Сохраните полученные ключи (открытые и секретные) в отдельный файл.

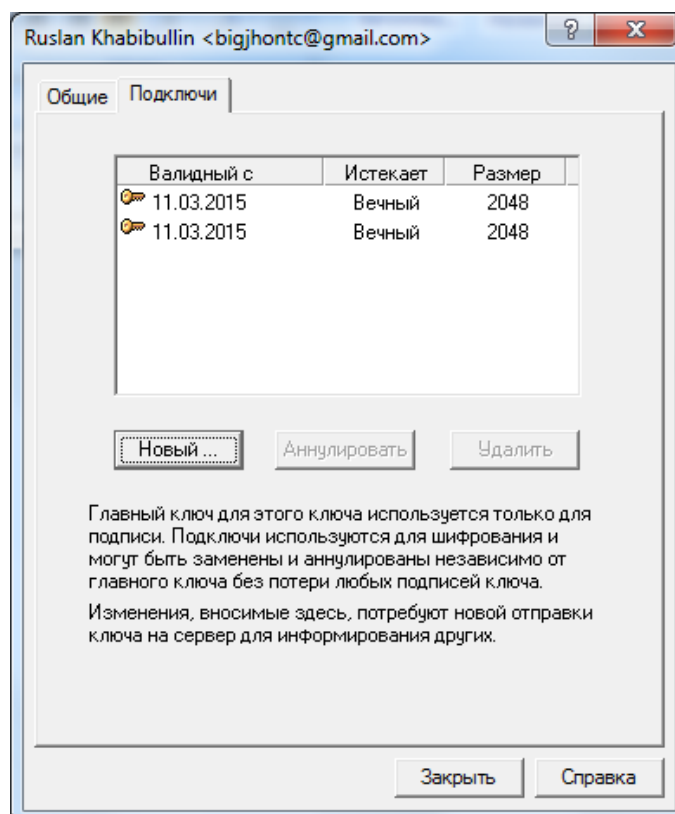


4. Назначьте ключ DH/DSS используемым по умолчанию.

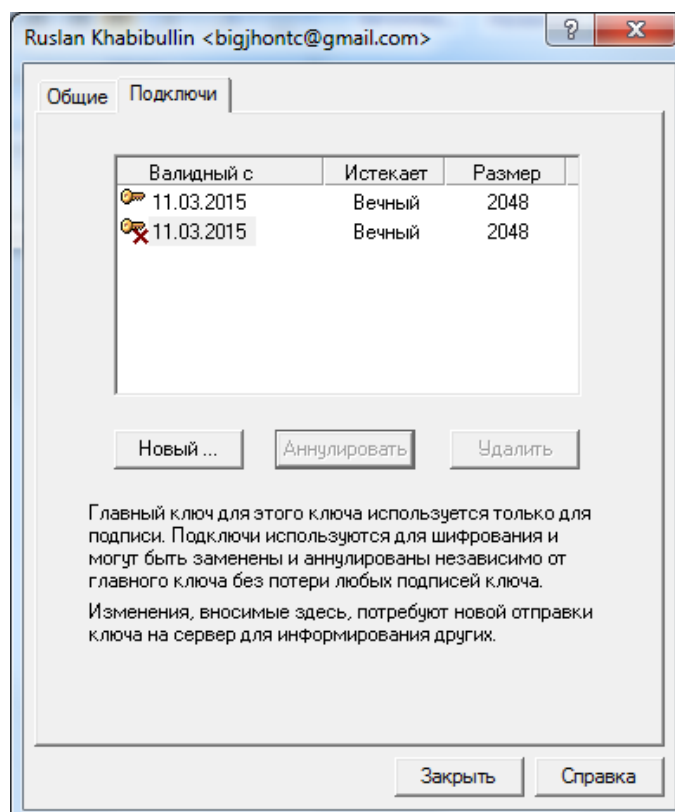


5. Изучите свойства ключевой пары DH/DSS.

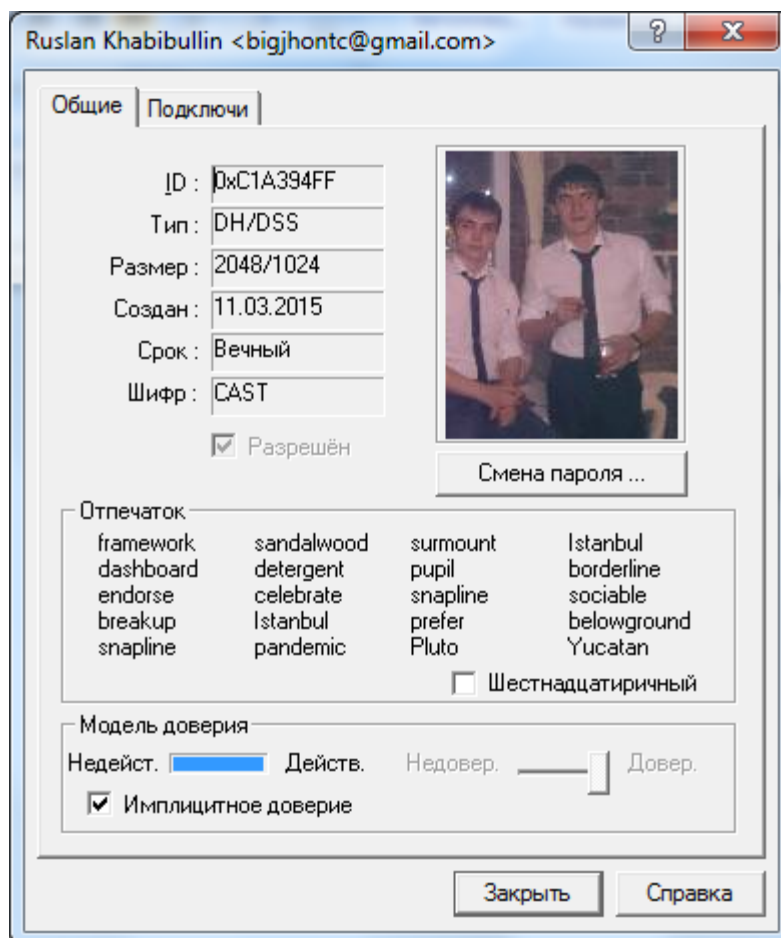
Создание подключа



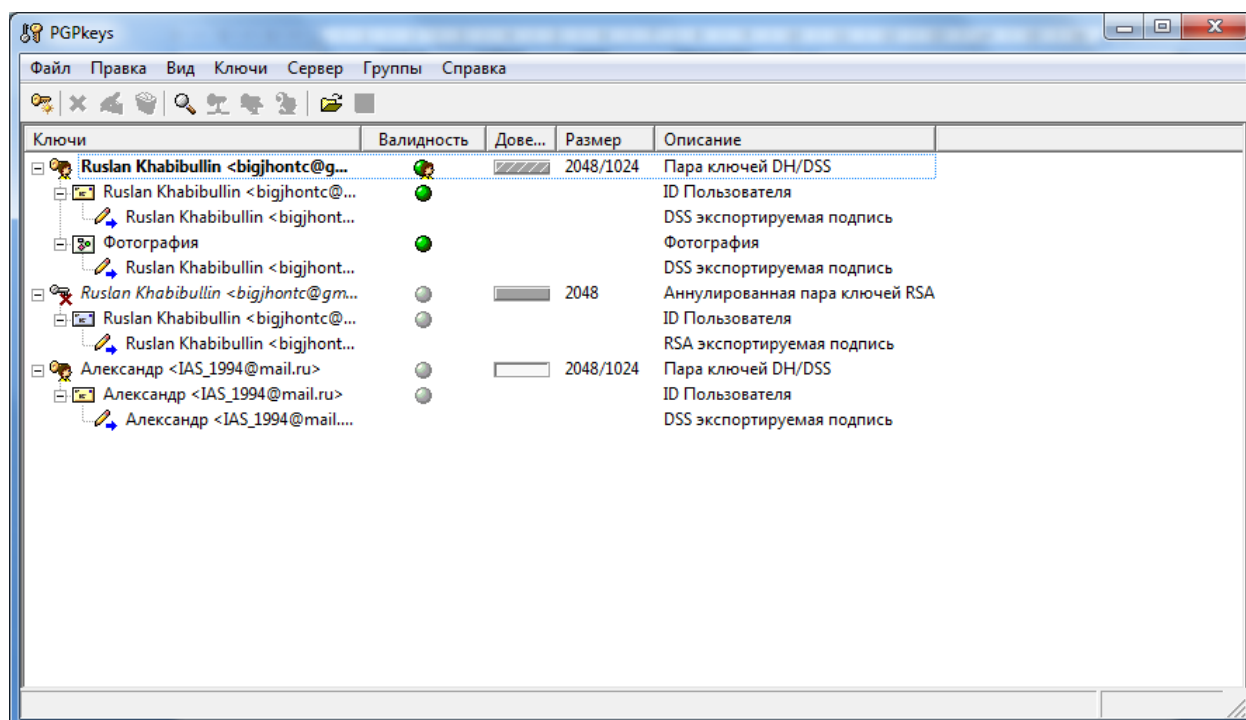
Аннулирование подключа



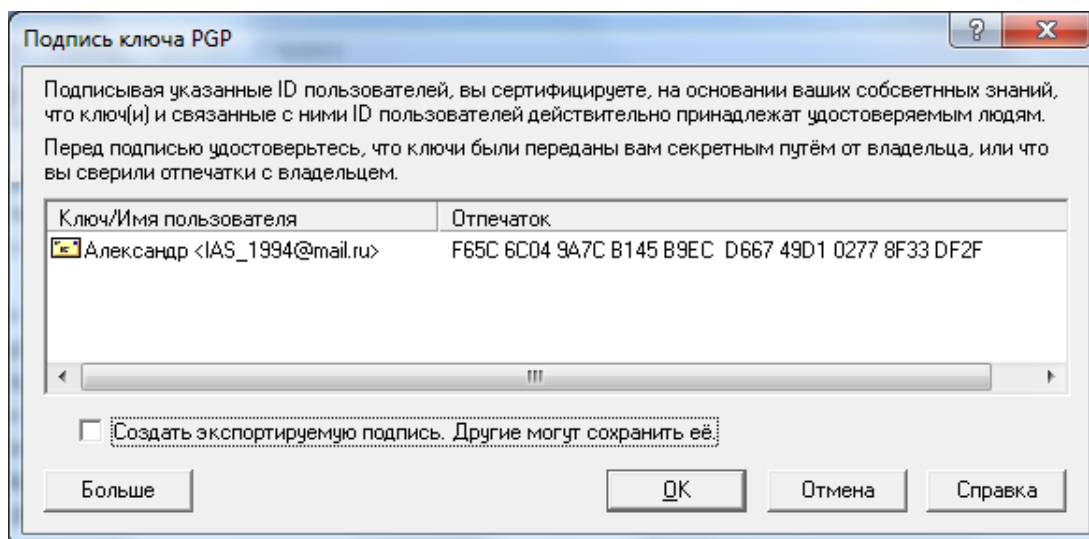
Добавление фотографии



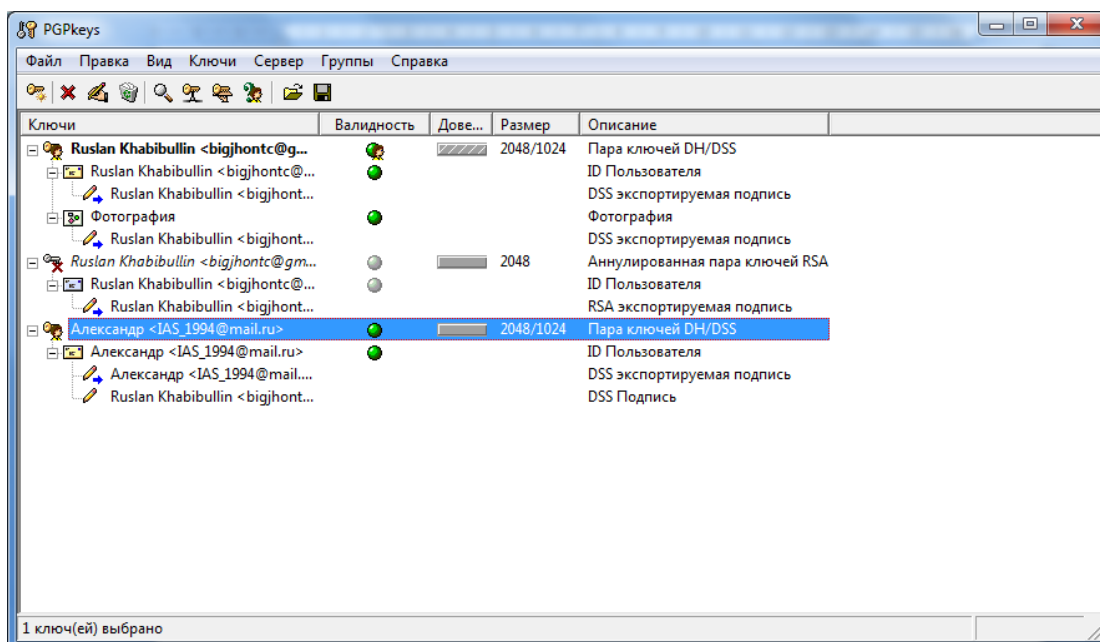
6. Обменяйтесь с другим студентом открытыми ключами DH/DSS.



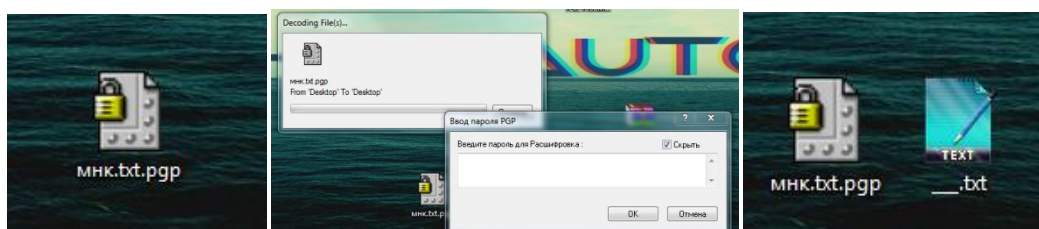
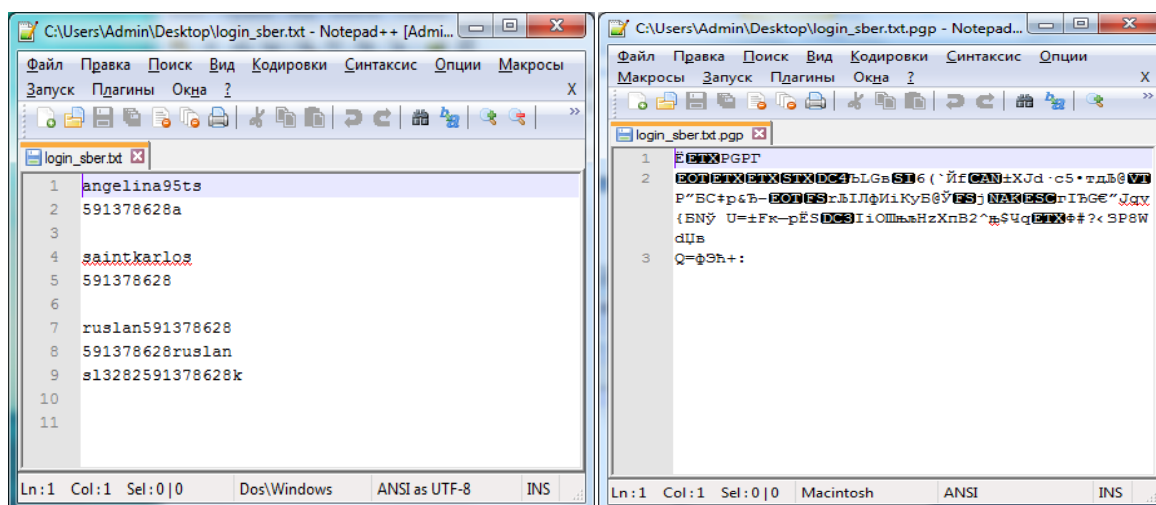
7. Установите подлинность полученного ключа с помощью его отпечатка.



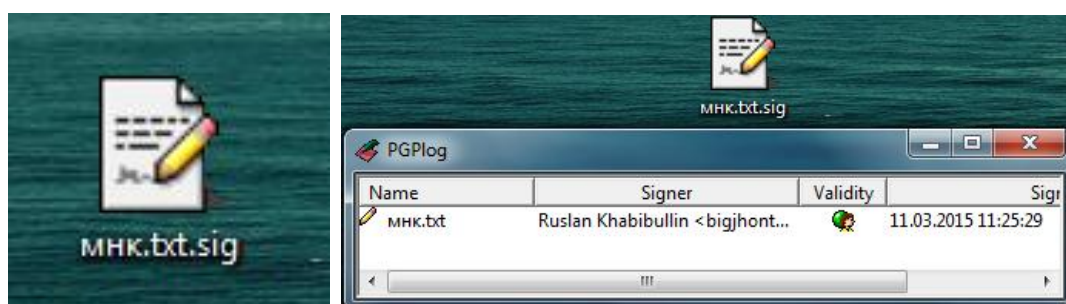
8. Установите степень доверия к владельцу полученного ключа на максимальный уровень.



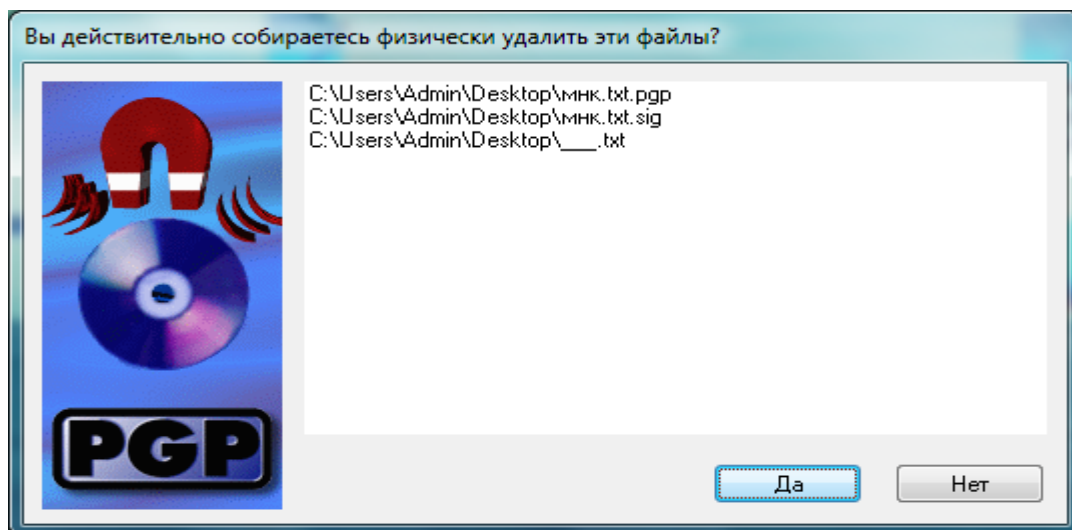
9. Зашифруйте произвольное сообщение/файл и обменяйтесь полученным результатом с другим студентом. Расшифруйте полученное сообщение.



10. Подпишите произвольное сообщение/файл и обменяйтесь полученным результатом с другим студентом. Проверьте достоверность источников полученного сообщения.



11. Уничтожьте все ненужные файлы, используя утилиту PGP Wipe.



Контрольные вопросы.

1. **Криптографическая система с открытым ключом (или асимметричное шифрование, асимметричный шифр)** — система шифрования и/или электронной подписи (ЭП), при которой открытый ключ передаётся по открытому (то есть незащищённому, доступному для наблюдения) каналу и используется для проверки ЭП и для шифрования сообщения. Для генерации ЭП и для расшифровки сообщения используется закрытый ключ. Криптографические системы с открытым ключом в настоящее время широко применяются в различных сетевых протоколах, в частности, в протоколах TLS и его предшественнике SSL (лежащих в основе HTTPS), в SSH. Также используется в PGP, S/MIME.

2. **Электронная подпись (ЭП), Электронная цифровая подпись (ЭЦП)** — реквизит электронного документа, полученный в результате криптографического преобразования информации с использованием закрытого ключа подписи и позволяющий установить отсутствие искажения информации в электронном документе с момента формирования подписи и проверить принадлежность подписи владельцу сертификата ключа подписи.

3. **Аннулирование** – после аннулирования открытого ключа PGP синхронизирует его с сервером, дабы в дальнейшем ваши корреспонденты не могли его применять. **Деактивация** – временное отключение неиспользуемого ключа или ключевой пары.

4. **Отпечаток ключа** — серия знаков, которой сопровождается каждый ключ. Сам по себе он не секретный, но уникальный. Если отпечаток того ключа, который вы получили (например) по электронной почте, и отпечаток, которым поделился ваш друг в Skype, совпадут, значит, у вас в руках правильный, настоящий ключ.

5. ИмPLICITное доверие – полное доверие зарезервировано для ключевых пар, расположенных на локальном ключе. Если одна часть ключевой пары находится на вашей связке, PGP предполагает, что вы владелец пары ключей и что Вы без проблем сможете доверять себе.

9. КРИПТОГРАФИЧЕСКИЕ ПРОТОКОЛЫ В СЕТЯХ ПЕРЕДАЧИ ДАННЫХ

9.1. Теория шифров с открытым ключом криптографические протоколы в сетях передачи данных

Безопасность сети передачи данных на транспортном уровне SSL и TLS [15]

Безопасность транспортного уровня обеспечивает услуги безопасности "из конца в конец" для приложений, которые используют протоколы транспортного уровня, такие как TCP. Основные идеи предназначены для того, чтобы обеспечить услуги безопасности на сети Интернет. Например, когда в сети имеются интерактивно работающие онлайн(online)-магазины, то желательны следующие услуги безопасности:

1. Клиент должен убедиться, что сервер принадлежит фактическому продавцу, а не самозванцу. Клиент не хочет сообщать самозванцу номер кредитной карточки (установление подлинности объекта).

2. Клиент и продавец должны быть убеждены, что содержание сообщения не изменено в течение передачи (целостность сообщения).

3. Клиент и продавец должны быть убеждены, что самозванец не перехватит чувствительную информацию, такую как номер кредитной карточки (конфиденциальность).

Сегодня применяются в основном два протокола обеспечения безопасности на транспортном уровне: *Протокол "Уровень безопасных розеток"* (SSL - *Secure Socket Layer*) и *Протокол Безопасности Транспортного уровня* (TLS - *Transport Layer Security*). Мы сначала обсудим SSL, затем TLS, а потом их сравним и покажем их отличия друг от друга.

Одна из целей этих протоколов состоит в том, чтобы обеспечить сервер и клиента услугами установления подлинности, конфиденциальности и целостности данных. Прикладной уровень программ клиент-сервер (client-server), таких как Язык передачи гипертекста (HTTP), который использует услуги TCP, может инкапсулировать свои данные в пакеты SSL. Если сервер и клиент согласованы с функционирующими программами SSL (или TLS), то клиент может использовать URL `https:// ...` вместо `http:// ...`, для того чтобы разрешить сообщениям HTTP инкапсулироваться в пакеты

SSL (или TLS). Например, номера кредитной карточки могут быть безопасно переданы через Интернет для онлайн-покупателей.

Протокол SSL

SSL (англ. secure sockets layer — уровень защищённых сокетов) — криптографический протокол, который подразумевает более безопасную связь. Он использует асимметричную криптографию для аутентификации ключей обмена, симметричное шифрование для сохранения конфиденциальности, коды аутентификации сообщений для целостности сообщений. Протокол широко использовался для обмена мгновенными сообщениями и передачи голоса через IP (англ. Voice over IP — VoIP), в таких приложениях, как электронная почта, Интернет-факс и др. В настоящее время известно, что протокол не является безопасным. SSL должен быть исключен из работы в пользу TLS (см. CVE-2014-3566).

SSL изначально разработан компанией Netscape Communications для добавления протокола HTTPS в свой веб-браузер Netscape Navigator. Впоследствии, на основании протокола SSL 3.0 был разработан и принят стандарт RFC, получивший имя TLS.

Протокол SSL обеспечивает защищенный обмен данных за счет двух следующих элементов:

- Аутентификация
- Шифрование

SSL использует асимметричную криптографию для аутентификации ключей обмена, симметричный шифр для сохранения конфиденциальности, коды аутентификации сообщений для целостности сообщений.

Протокол SSL предоставляет "безопасный канал", который имеет три основных свойства:

1. Канал является частным. Шифрование используется для всех сообщений после простого диалога, который служит для определения секретного ключа.
2. Канал аутентифицирован. Серверная сторона диалога всегда аутентифицируется, а клиентская делает это опционально.
3. Канал надежен. Транспортировка сообщений включает в себя проверку целостности.

Преимуществом SSL является то, что он независим от прикладного протокола. Протоколы приложений (HTTP, FTP, TELNET и т.д.) могут работать поверх протокола SSL совершенно прозрачно, т.е. SSL может согласовывать алгоритм шифрования и ключ сессии, а также аутентифицировать сервер до того, как приложение примет или передаст первый байт сообщения.

2.1 Принцип работы

SSL использует среду с несколькими слоями, что обеспечивает безопасность обмена информацией. Конфиденциальность общения присутствует за счет того, что безопасное соединение открыто только целевым пользователям.

2.1.1 Многослойная среда

Протокол SSL размещается между двумя протоколами: протоколом, который использует программа-клиент (HTTP, FTP, LDAP, TELNET etc) и транспортным протоколом TCP/IP. SSL защищает данные выступая в роли фильтра для обеих сторон и передает их далее на транспортный уровень. Работу протокола можно разделить на два уровня:

1. Слой протокола подтверждения подключения (Handshake Protocol Layer)
2. Слой протокола записи

Первый слой, в свою очередь, состоит из трех подпротоколов:

1. Протокол подтверждения подключения (Handshake Protocol)
2. Протокол изменения параметров шифра (Cipher Spec Protocol)
3. Предупредительный протокол (Alert Protocol)

Протокол подтверждения подключения используется для согласования данных сессии между клиентом и сервером. К данным сессии относятся:

- Идентификационный номер сессии
- Сертификаты обеих сторон
- Параметры алгоритма шифрования
- Алгоритм сжатия информации
- "Общий секрет" применен для создания ключей; открытый ключ

Протокол подтверждения подключения производит цепочку обмена данными, что в свою очередь начинает аутентификацию сторон и согласовывает шифрование, хэширование и сжатие. Следующий этап - аутентификация участников, которая осуществляется также протоколом подтверждения подключения.

Протокол изменения параметров шифра используется для изменения данных ключа (keyingmaterial) - информации, которая используется для создания ключей шифрования. Протокол состоит всего из одного сообщения, в котором сервер говорит, что отправитель хочет изменить набор ключей.

Предупредительный протокол содержит сообщение, которое показывает сторонам изменение статуса или сообщает о возможной ошибке. Обычно предупреждение отсылается

тогда, когда подключение закрыто и получено неправильное сообщение, сообщение невозможно расшифровать или пользователь отменяет операцию.

Цифровые сертификаты

Протокол SSL использует сертификаты для проверки соединения. Сертификаты расположены на безопасном сервере и используются для шифрования данных и идентификации Web-сайта.

Способы получения SSL-сертификата:

1. Использовать сертификат, выданный CA
2. Использовать самоподписанный сертификат
3. Использовать "пустой" сертификат

Самоподписанный сертификат - сертификат, созданный самим пользователем - в этом случае издатель сертификата совпадает с владельцем сертификата. “Пустой” сертификат - сертификат, содержащий фиктивную информацию, используемую в качестве временной для настройки SSL и проверки его функциональности в данной среде.

Механизмы образования ключа для текущей сессии в SSL/TLS

Для обмена подлинными и конфиденциальными сообщениями клиенту и серверу нужны шесть криптографических объектов секретности (четыре ключа и два вектора инициализации). Однако чтобы создать их, между этими двумя сторонами должен быть установлен один предварительный главный секретный код (pre-master secret). SSL определяет шесть методов обмена ключами, чтобы установить этот предварительный объект секретности: NULL, RSA, анонимный Диффи-Хеллман (Diffie-Hellman), кратковременный Диффи-Хеллман, фиксированный Диффи-Хеллман и Fortezza

RSA

В этом методе предварительный главный секретный код - 48-байтовое случайное число, созданное клиентом, зашифрованное открытым ключом RSA-сервера и передаваемое серверу. Сервер должен передать свой сертификат шифрования/дешифрования RSA.

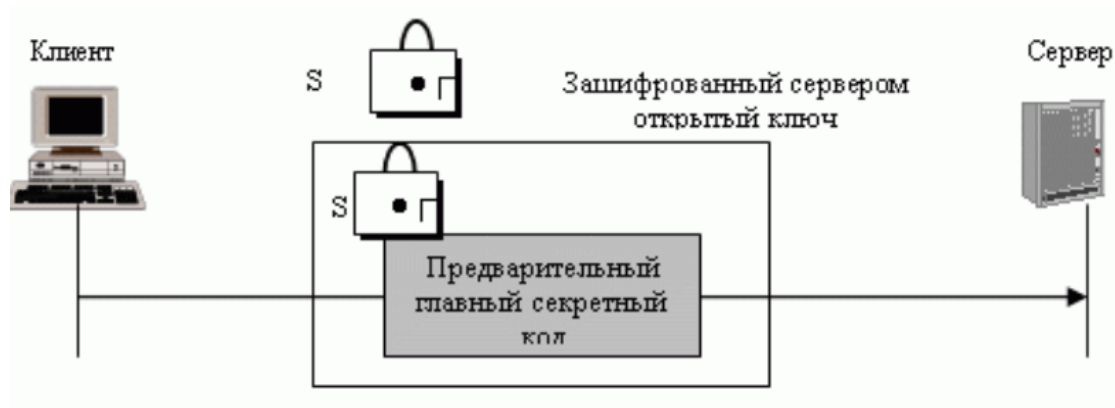


Рис. 9.1 – RSA – смена ключа; открытый ключ сервера

Анонимный протокол Диффи-Хеллмана

Это самый простой и наиболее ненадежный метод. Предварительный главный секретный код устанавливают между клиентом и сервером, используя протокол Диффи-Хеллмана. При этом передают половину ключа в исходном тексте - это называется анонимным протоколом Диффи-Хеллмана, потому что ни одна сторона не известна другой. Самый серьезный недостаток этого метода - возможность атаки "посредника".

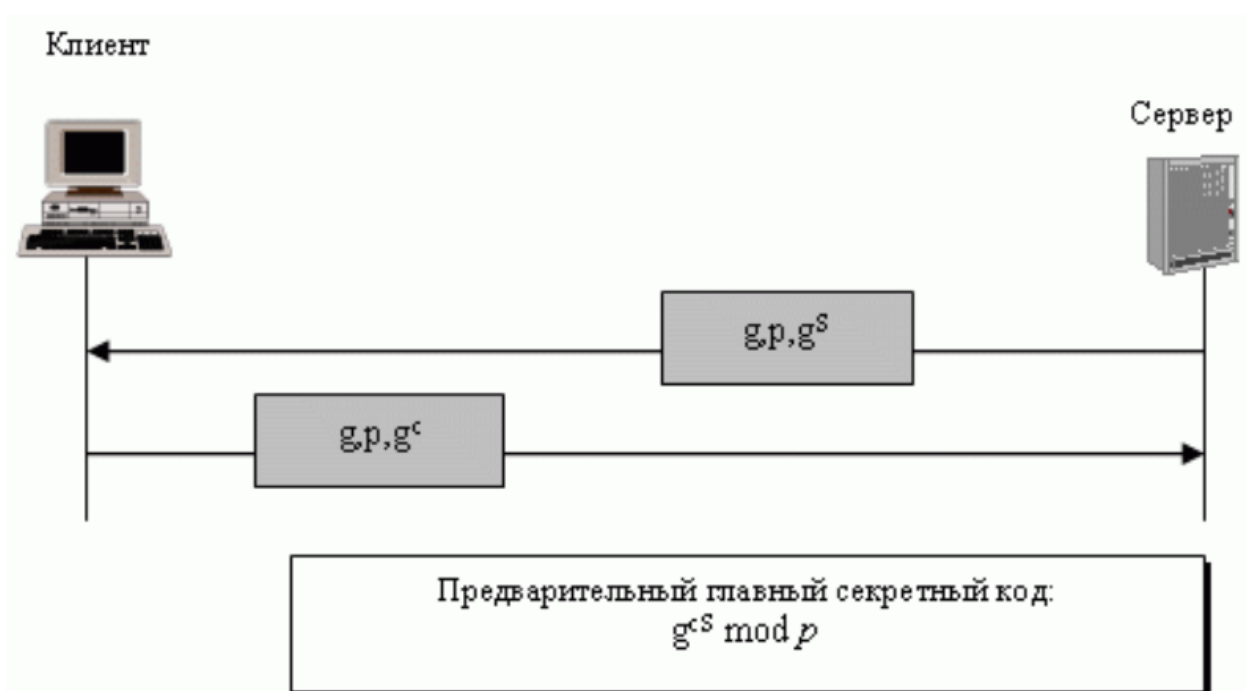


Рис. 9.2. Анонимный протокол Диффи-Хеллмана смены ключей

Кратковременный метод Диффи-Хеллмана

Чтобы сорвать атаку "посредника", может быть использована кратковременная смена ключей методом Диффи-Хеллмана. Каждая сторона передает ключ Диффи-Хеллмана, подписанный своим секретным ключом. На приемной стороне должны проверить подпись,

используя открытый ключ передатчика. Обмен открытыми ключами для проверки использует либо RSA-, либо DSS-сертификат цифровой подписи.

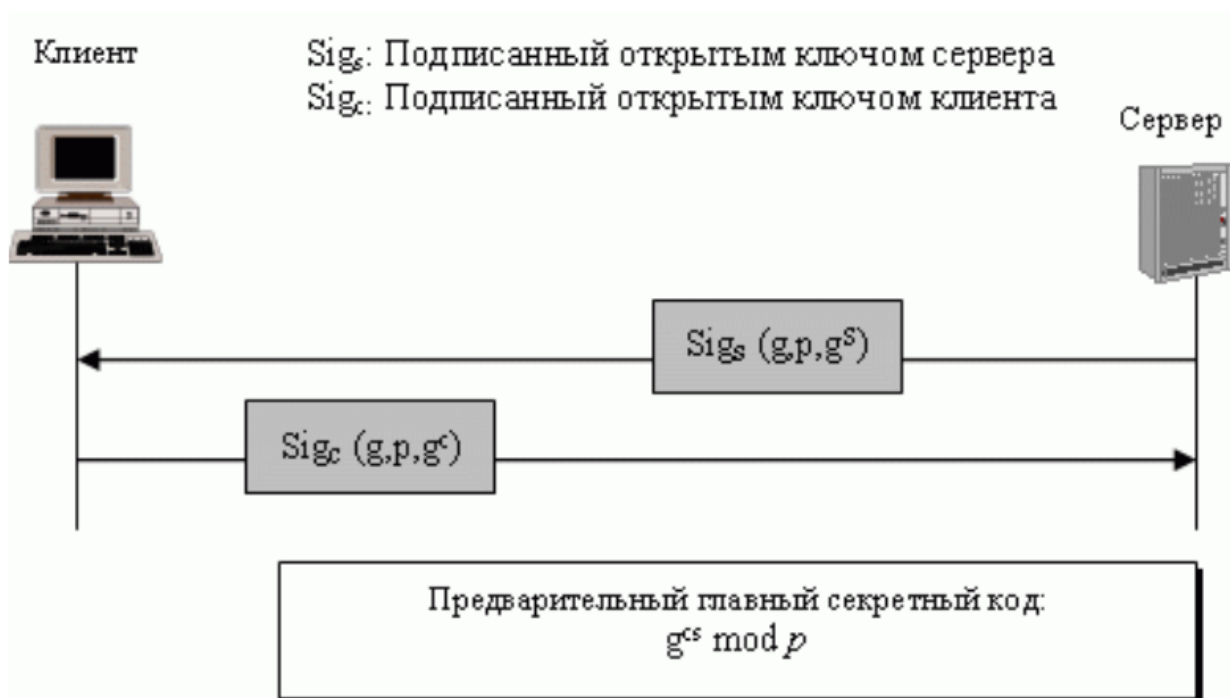


Рис. 9.3. Кратковременный протокол Диффи-Хеллмана смены ключей

Фиксированный метод Диффи-Хеллмана

Другое решение - фиксированный метод Диффи-Хеллмана. Все объекты в группе могут подготовить фиксированные параметры (g и p). Затем каждый объект может создать фиксированную половину ключа (gx). Для дополнительной безопасности каждая отдельная половина ключа Диффи-Хеллмана вставляется в сертификат, проверенный центром сертификации (CA). Другими словами, две стороны отдельно не обмениваются полуключами; CA передает полуключи в специальном сертификате RSA или DSS. Когда клиент должен вычислить *предварительный главный секретный код*, он использует свой собственный фиксированный полуключ и полуключ сервера, полученный в сертификате. Сервер делает то же самое, но в обратном порядке. Обратите внимание, что в этом методе не передаются сообщения смены ключей, а происходит только обмен сертификатами.

Алгоритмы шифрования/дешифрования

Есть несколько возможностей выбора алгоритма шифрования/дешифрования. Мы можем разделить алгоритмы на 6 групп, как это показано на рис. 2.4. Все протоколы блока используют 8-байтовый вектор инициализации (IV), кроме Fortezza, который применяет 20 байтов IV.

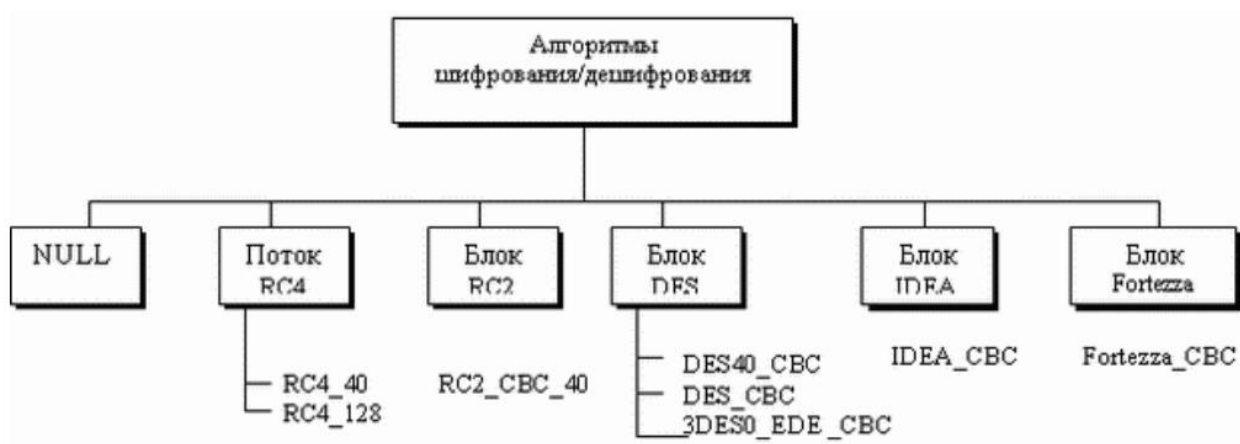


Рис. 9.4. Алгоритмы шифрования/дешифрования

NULL

NULL - категория, которая просто определяет отсутствие алгоритма шифрации/дешифрации.

Поток RC

В режиме потока RC определены два потока алгоритма: RC4-40 (ключ на 40 битов) и RC4-128 (ключ на 128 битов).

Блок RC

В режиме блока RC определен один алгоритм: RC2_CBC_40 (ключ на 40 битов). CBC (Cipher Block Chaining) - сцепление шифрованных блоков.

DES

Все алгоритмы DES определены в режиме блока. DES40_CBC использует ключ на 40 битов. Стандартные DES определены как DES_CBC. 3DES_EDE_CBC используют ключ на 168 битов.

IDEA

В режиме блока IDEA определен один алгоритм - IDEA_CBC, с ключом на 128 битов.

Fortezza

В режиме блока Fortezza определен один алгоритм - FORTEZZA_CBC, с ключом на 96 бит.

2.1.5 Алгоритмы хэширования

SSL использует алгоритмы хэширования, чтобы обеспечить целостность сообщения (установление подлинности сообщения). Имеются хэш-функции, показанные на рис. 2.5.



Рис. 9.5. Алгоритмы хэширования

Null (Пустой указатель)

Две стороны могут отказаться использовать алгоритм хэширования. В этом случае сообщение не заверено.

MD5

Две стороны могут выбрать MD5 как алгоритм хэширования. В этом случае используется алгоритм хэширования MD5 - 128-битовый.

SHA-1

Две стороны могут выбрать SHA как алгоритм хэширования. В этом случае используется алгоритм хэширования SHA-1 на 160 битов.

Алгоритмы сжатия

Как мы уже говорили, сжатие является дополнительной услугой в SSLv3. Для SSLv3 не определен алгоритм сжатия. Поэтому заданным по умолчанию методом сжатия служит NULL. Однако система может использовать любой алгоритм сжатия по выбору сторон.

Генерирование криптографических параметров

Чтобы обеспечить целостность и конфиденциальность сообщения, в SSL необходимо иметь: шесть криптографических объектов секретности, четыре ключа и два инициализирующих вектора (IV). Клиенту нужно: один ключ для передачи сообщения установления подлинности (HMAC - HASH-BASED MESSAGE AUTHENTICATION CODE), один ключ для шифрования и один IV для шифрования блока. Сервер нуждается в том же самом. SSL требует, чтобы ключи для одного направления отличались от ключей для другого направления. Если будет атака в одном направлении, она не затронет другое направление. Для генерации параметров используют следующую процедуру:

1. Клиент и сервер обмениваются двумя случайными числами, одно из которых создано клиентом, а другое - сервером.
2. Клиент и сервер обмениваются одним предварительным главным секретным кодом.

3. Создается 48-байтовый главный секретный код (master secret) из предварительного главного секретного кода (pre-master secret), с применением хэш-функций (SHA-1 и MD5), как это показано на рис. 4.6.

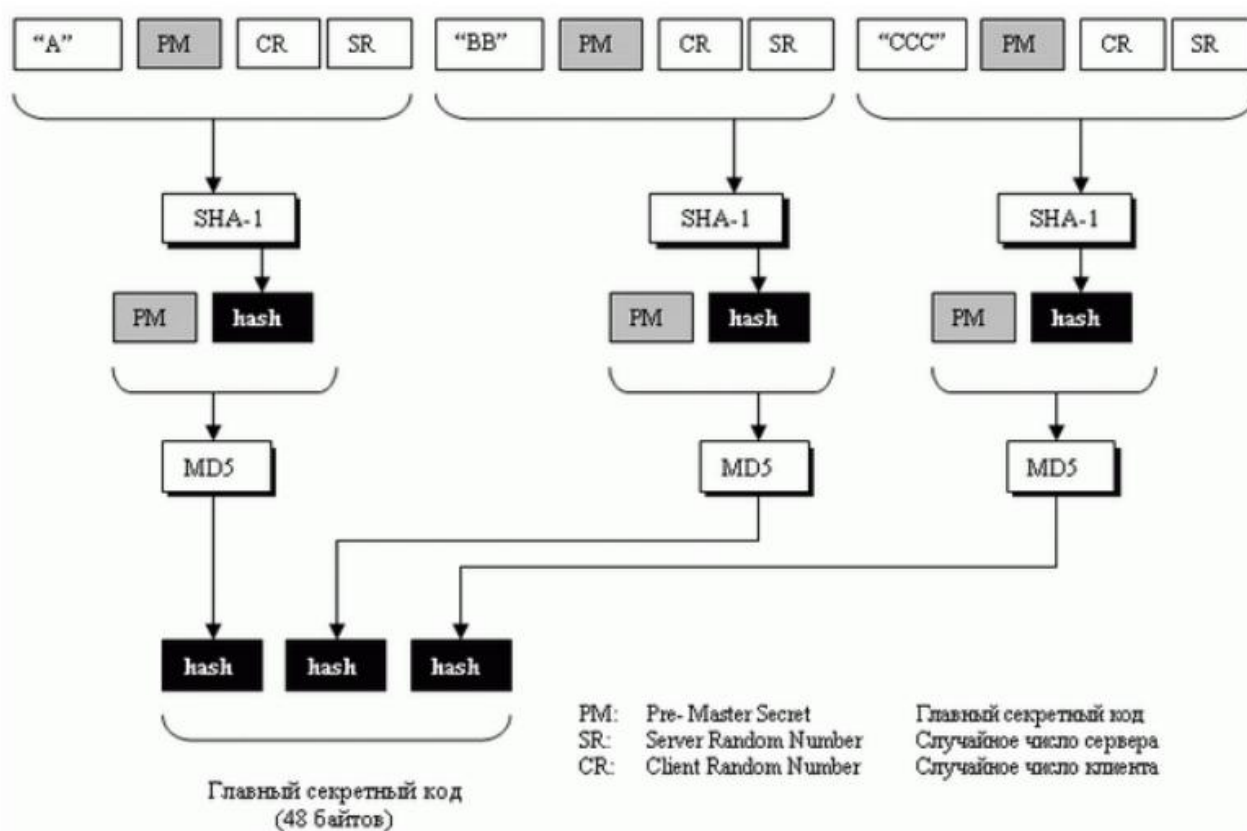


Рис. 9.6. Вычисление главного секретного кода из предварительного главного секретного кода

4. Главный секретный код используется для того, чтобы создать материал для ключей (key material), который имеет переменную длину. Для этого применяют то же самое множество хэш-функций, что и в предыдущем случае, и подставляют спереди различные константы, как это показано на рис. 4.7. Алгоритм повторяется, пока не получится материал для ключа адекватного размера.

Длина блока материала для ключей зависит от выбранного набора шифра и размера ключей, необходимых для этого набора.

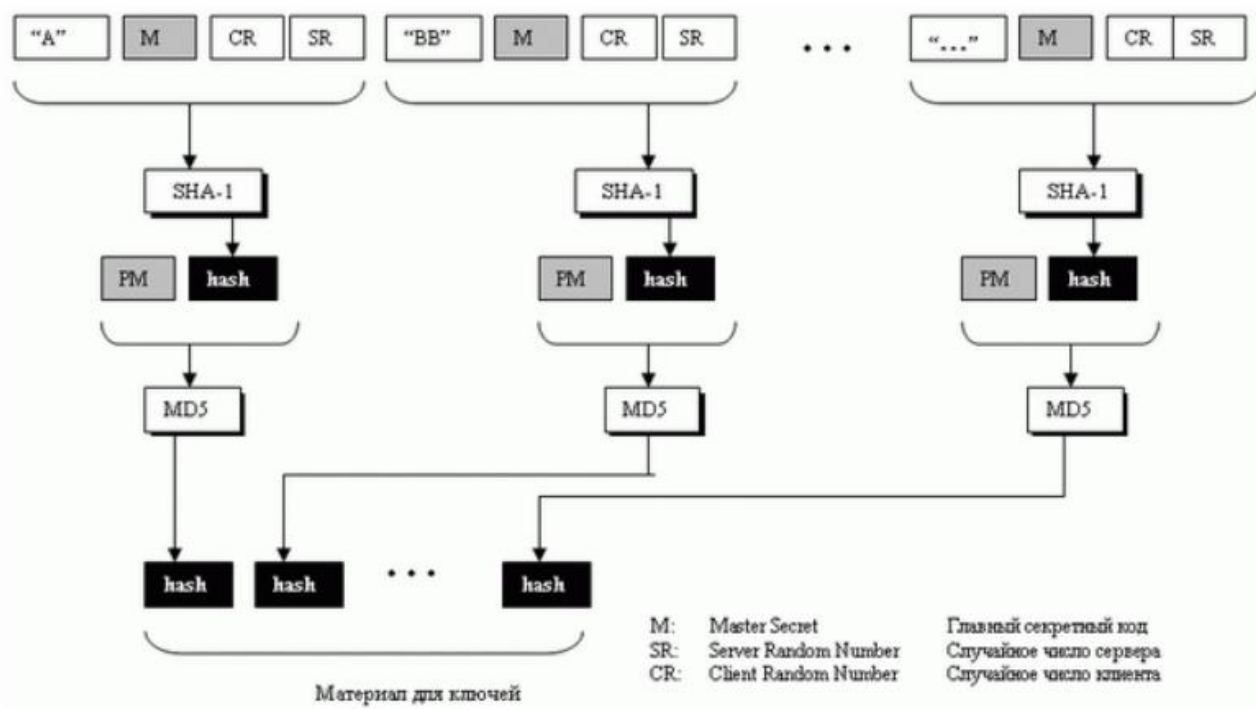


Рис. 9.7. Вычисление материала для ключей из главного секретного кода

5. Из материала для ключей извлекаются шесть различных ключей, как показано на рис. 2.8.

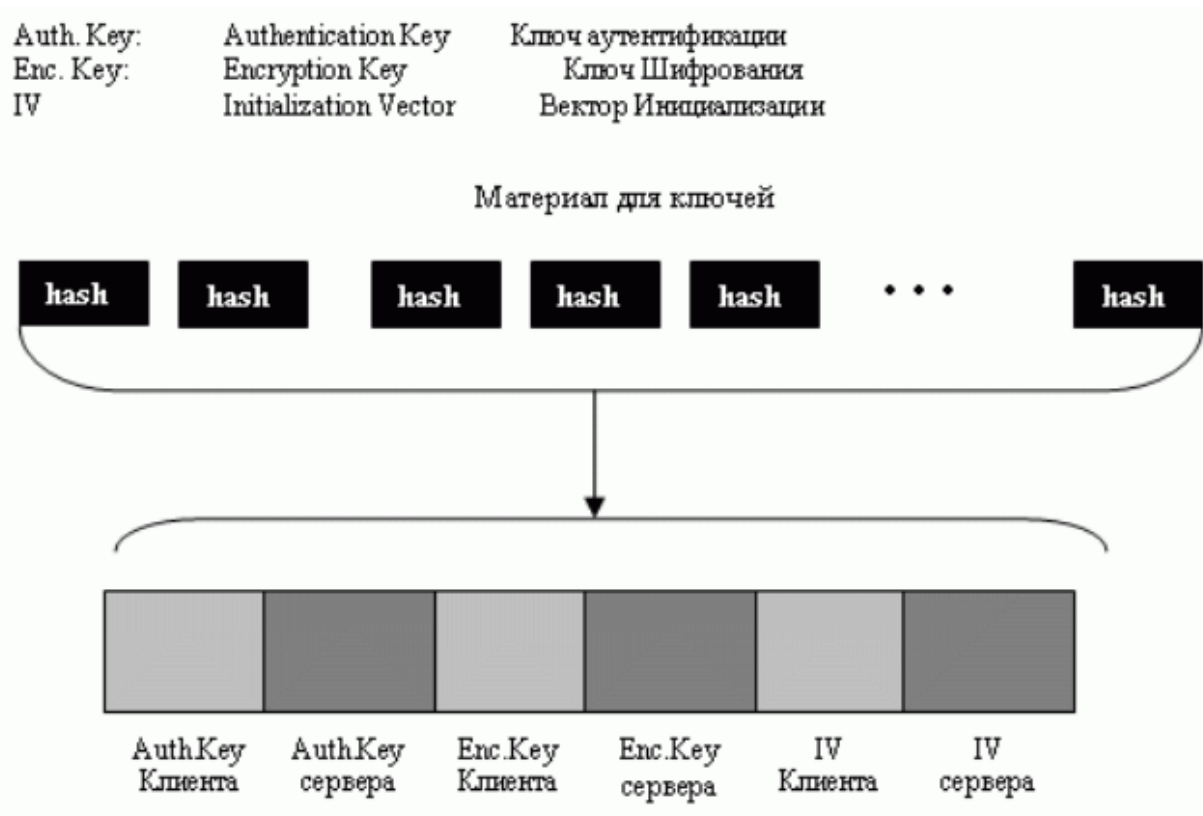


Рис. 9.8. Извлечение криптографических секретных кодов из материала

2.1.8 Сеансы и соединение

SSL отличает соединение от сеанса. Давайте рассмотрим эти два термина. Сеанс – связь между клиентом и сервером. После того как сеанс установлен, эти две стороны имеют общую информацию, такую как идентификатор сеанса, сертификат, подтверждающий подлинность каждого из них (в случае необходимости), метод сжатия(если необходимо), набор шифров и главный секретный код. Эта информация используется для того, чтобы создать ключи для сообщения, содержащего шифр установления подлинности.

Для двух объектов, чтобы начать обмен данными, установление сеанса необходимо, но не достаточно; они должны создать между собой соединение. Эти два объекта обмениваются двумя случайными числами и создают, используя главный секретный код, ключи и параметры, необходимые для того, чтобы обмениваться сообщениями, включая установление подлинности и секретность.

Сеанс может состоять из многих соединений. Соединение между двумя сторонами может быть закончено и восстановлено в пределах одного и того же сеанса. Когда соединение закончено, эти две стороны могут также закончить сеанс, но это необязательно. Сеанс может быть приостановлен и продолжен снова.

Чтобы создавать новый сеанс, эти две стороны должны пройти процесс переговоров. Чтобы возобновлять старый сеанс и создавать только новое соединение, эти две стороны могут пропустить часть переговоров, что уменьшает время вхождения в связь. Не надо создавать главный секретный код, когда сеанс продолжается.

Разделение сеанса от соединения предотвращает высокую стоимость создания главного секретного кода. Если мы разрешаем приостановления и продолжения сеанса, процесс вычисления главного секретного кода может быть устранен. рис. 2.9 иллюстрирует идею сеанса и соединения в этом сеансе.



Рис. 9.9. Сеанс и соединение

В сеансе одна сторона играет роль клиента и другая - роль сервера. При соединении обе стороны имеют равные роли, они равны по уровню.

Четыре протокола

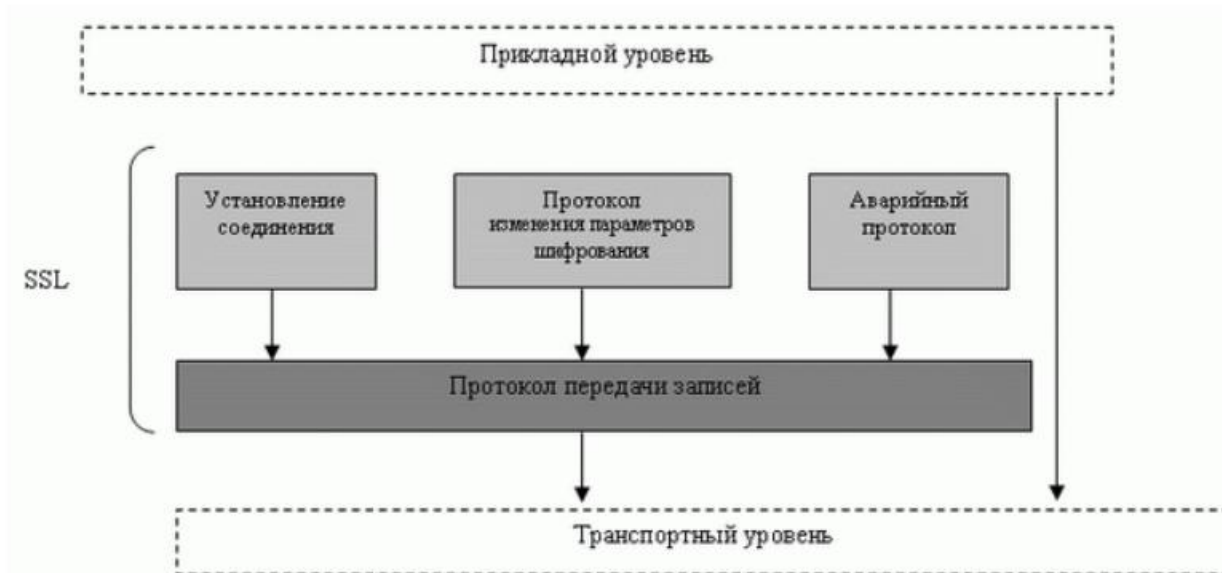


Рис. 9.10. Четыре протокола SSL

Протокол передачи записей - переносящий информацию. Он переносит на транспортный уровень сообщения от трех других протоколов, а также данные, поступающие от прикладного уровня. Сообщения из протокола записей - это полезная нагрузка для транспортного уровня, обычно TCP. Протокол установления соединения обеспечивает параметры безопасности для Протокола записей. Он устанавливает набор шифров и задает ключи и параметры безопасности.

Он также подтверждает, если необходимо, подлинность сервера клиенту и подлинность клиента серверу. Протокол изменения параметров шифрования используется, чтобы передавать сигналы для подготовки к криптографической безопасности. Аварийный протокол нужен, чтобы известить о ситуациях, отклоняющихся от нормы.

Протокол TLS

Безопасность транспортного уровня (TLS - Transport Layer Security) - протокол IETF, стандартная версия протокола SSL. Эти два протокола очень похожи, но имеют небольшие отличия. Вместо того чтобы описывать TLS полностью, в этой секции мы только отметим отличия между протоколами TLS и SSL.

Генерация криптографической секретности

Генерация криптографической секретности в TLS более сложная, чем в SSL. TLS сначала определяет две функции: функцию расширения данных и псевдослучайную функцию.

Функция расширения данных

Функция расширения данных использует заранее заданный код аутентификации на основе хэширования (HMAC-HASH-BASED MESSAGE AUTHENTICATION CODE), или MD5, или SHA-1 для того, чтобы расширить информацию засекречивания. Эту функцию можно рассматривать как функцию, содержащую множество секций, где каждая секция

создает одно значение хэширования. Расширенная секретность – последовательное соединение значений хэширования. Каждая секция использует два HMAC, информацию засекречивания и начальное число. Функция расширения данных - это формирование цепочки в виде многих секций. Однако чтобы сделать следующую секцию зависимой от предыдущей, второе начальное число – фактически выход первого HMAC предыдущей секции, как это показано на рис. 9.11.

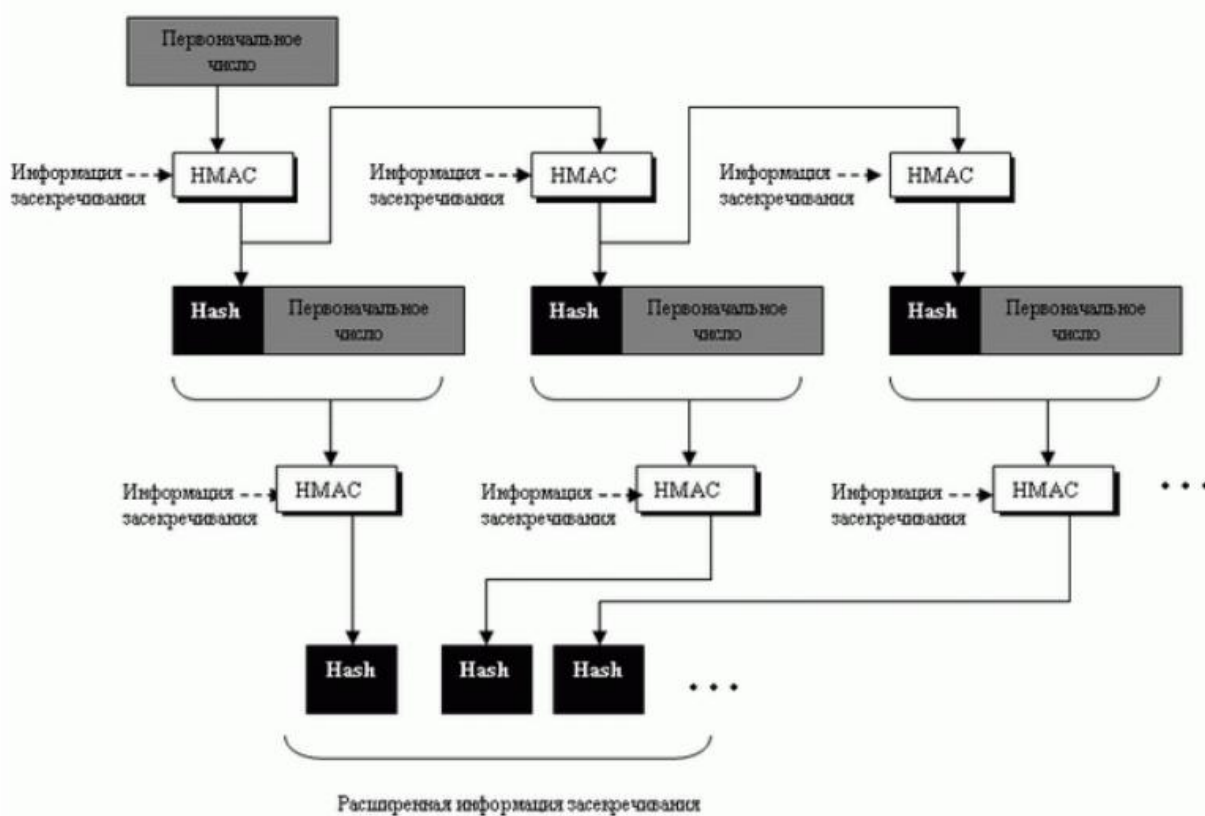


Рис. 9.11. Функция расширения данных

Псевдослучайная функция

TLS определяет псевдослучайную функцию (PRF - PseudoRandom Function), чтобы получить комбинацию двух функций расширения данных: одна из них использует MD5 и другая - SHA-1. На PRF поступает три части информации: секретный код, метка и начальное число.

Метка и начальное число связаны и служат начальным числом для каждой функции расширения данных. Информация засекречивания разделена на две части; каждая часть используется как информация засекречивания для каждой функции расширения данных. Выходы двух функций расширения данных складывают по модулю два, чтобы создать конечную расширенную информацию засекречивания. Обратите внимание, что поскольку хэш создается MD5 и SHA-1, он имеет различные размеры, поэтому должны быть созданы дополнительные секции функций на базе MD5, чтобы сделать два вывода с одинаковым размером. рис. 4.12 показывает идею применения PRF.

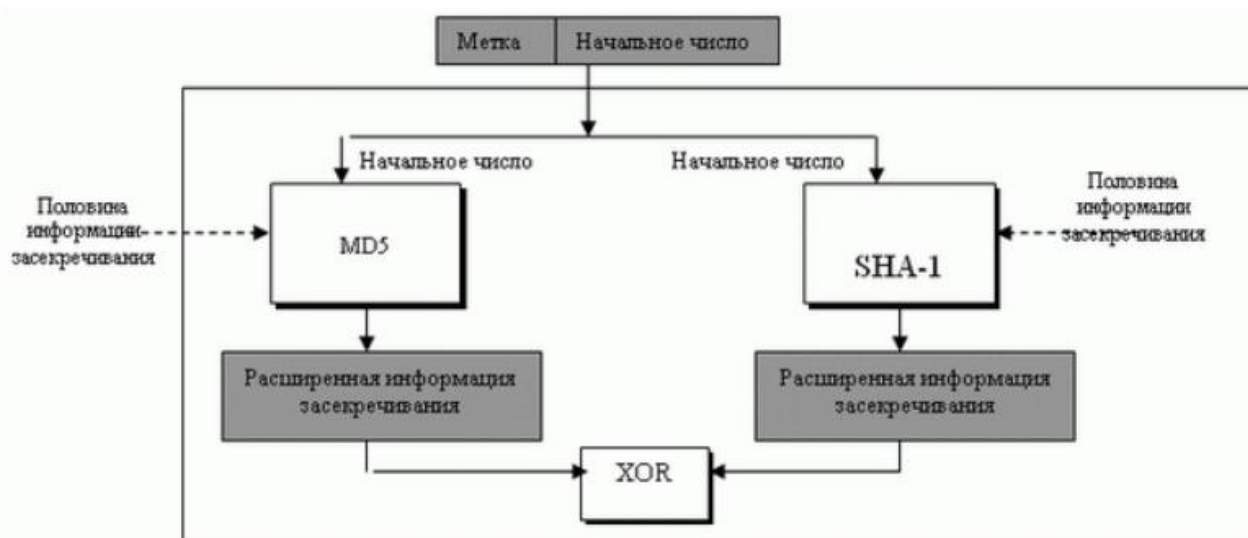


Рис. 9.12. PRF

Главный секретный код

TLS использует функцию PRF, чтобы создать главный секретный код от предварительного главного секретного кода. Это можно сделать, используя предварительный главный секретный код как информацию засекречивания, строку "главный секретный код" - как метку и последовательное соединение информации (конкатенацию) случайного числа клиента и случайное число сервера - как начальное число. Обратите внимание, что метка - фактически код ASCII строки "главного секретного кода". Другими словами, метка определяет выход для создания главного секретного кода. рис. 3.3 иллюстрирует идею.



Рис. 9.13. Генерация главного секретного ключа

Материал для ключей

TLS использует функцию PRF, чтобы создать материал для ключей от главного секретного кода. На сей раз информация засекречивания содержит: главный секретный код; метку - это строка "расширение ключа"; и начальное число – конкатенацию случайного числа сервера и случайного числа клиента, как это показано на рис. 3.4.



Рис. 9.14. Генерация материала для ключа

Протоколы

Аварийный протокол в TLS поддерживает все аварийные сигналы, определенные в SSL за исключением NoCertificate. TLS также добавляет к списку SSL некоторые новые.

Протокол установления соединения TLS вносит некоторые изменения в протокол установления соединения. Были специально изменены детали сообщения CertificateVerify и сообщения Finished.

Единственное изменение в протоколе передачи записей – использование HMAC, формируемы с помощью MAC, чтобы подписать сообщение.

4.2. Безопасность сети ПД на сетевом уровне IP SEC

IPSec является неотъемлемой частью IPv6 - Интернет-протокола следующего поколения, и расширением существующие версии Интернет-протокола IPv4. IPSec определён в RFC с 2401 по 2412.

Практически все механизмы сетевой безопасности могут быть реализованы на третьем уровне эталонной модели ISO/OSI в соответствии с рисунком 2.1. Кроме того, IP-уровень можно считать оптимальным для размещения защитных средств, поскольку при этом достигается удачный компромисс между защищенностью, эффективностью функционирования и прозрачностью для приложений.

Уровни TCP/IP	Уровни ISO/OSI
4. Прикладных программ	7. Прикладных программ 6. Представление данных
3. Транспортный	5. Сеансовый 4. Транспортный
2. Межсетевой	3. Сетевой
1. Доступа к сети	2. Канальный 1. Физический

Рис. 9.15. Модель OSI/ISO

Стандартизованными механизмами IP-безопасности могут (и должны) пользоваться протоколы более высоких уровней и, в частности, управляющие протоколы, протоколы конфигурирования и маршрутизации.

Средства безопасности для IP описываются семейством спецификаций IPSec, разработанных рабочей группой IP Security.

Протоколы IPSec обеспечивают управление доступом, целостность вне соединения, аутентификацию источника данных, защиту от воспроизведения, конфиденциальность и частичную защиту от анализа трафика.

Основополагающими понятиями IPSec являются:

- - аутентификационный заголовок (AH);

- - безопасное сокрытие данных (ESP);
- - режимы работы: туннельный и транспортный;
- - контексты (ассоциации) безопасности (SA);
- - управление ключами (IKE);

Основные составляющие архитектуры и их особенности

Архитектура средств безопасности для IP-уровня специфицирована в документе Security Architecture for the Internet Protocol. Ее основные составляющие представлены в соответствии с рисунком 2.2. Это, прежде всего протоколы обеспечения аутентичности (протокол аутентифицирующего заголовка - Authentication Header, AH) и конфиденциальности (протокол инкапсулирующей защиты содержимого - Encapsulating Security payload, ESP), а также механизмы управления криптографическими ключами. На более низком архитектурном уровне располагаются конкретные алгоритмы шифрования, контроля целостности и аутентичности. Наконец, роль фундамента выполняет так называемый домен интерпретации (Domain of Interpretation, DOI), являющийся, по сути, базой данных, хранящей сведения об алгоритмах, их параметрах, протокольных идентификаторах.

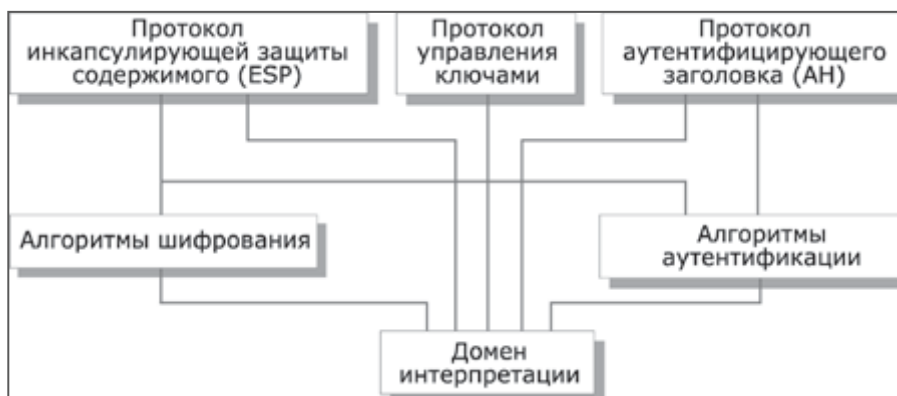


Рис. 9.16. Основные элементы архитектуры средств безопасности IP-уровня

Деление на уровни важно для всех аспектов информационных технологий. Там же, где участвует еще и криптография, важность возрастает вдвойне, поскольку приходится считаться не только с чисто техническими факторами, но и с особенностями законодательства различных стран, с ограничениями на экспорт и/или импорт криптосредств.

IPSec поддерживает две формы целостности: целостность соединения и частичную целостность последовательности. Целостность соединения является сервисом безопасности, который определяет модификацию конкретной IP датаграммы, безотносительно последовательности датаграмм в потоке трафика. Частичная целостность

последовательности является anti-reply сервисом, с помощью которого определяется получение дубликатов IP датаграм.

Эти сервисы как раз и реализуются с использованием двух протоколов обеспечения безопасного трафика, Authentication Header (AH) и Encapsulating Security Payload (ESP), и с помощью процедур и протоколов управления криптографическим ключом. Множество применяемых IPSec протоколов и метод их использования определяются требованиями безопасности.

Когда данные механизмы установлены корректно, они не мешают пользователям, хостам и другим компонентам Internet, которые не применяют данные механизмы безопасности для защиты своего трафика. Протоколы обеспечения аутентичности и конфиденциальности в IPSec не зависят от конкретных криптографических алгоритмов. (Более того, само деление на аутентичность и конфиденциальность предоставляет и разработчикам, и пользователям дополнительную степень свободы в ситуации, когда к криптографическим относят только шифровальные средства.) В каждой стране могут применяться свои алгоритмы, соответствующие национальным стандартам, но для этого, как минимум, нужно позаботиться об их регистрации в домене интерпретации. Это означает возможность выбора различного набора алгоритмов без воздействия на другие части реализации. Например, различные группы пользователей могут выбрать при необходимости различные наборы алгоритмов.

Определен стандартный набор алгоритмов по умолчанию для обеспечения интероперабельности. Использование этих алгоритмов совместно с защитой трафика на основе IPSec и протоколами управления ключа позволяет обеспечить высокую степень криптографической безопасности.

Алгоритмическая независимость протоколов имеет и обратную сторону, состоящую в необходимости предварительного согласования набора применяемых алгоритмов и их параметров, поддерживаемых общающимися сторонами. Иными словами, стороны должны выработать общий контекст безопасности (Security Association, SA) и затем использовать такие его элементы, как алгоритмы и их ключи. SA подробно рассматривается далее. За формирование контекстов безопасности в IPSec отвечает особое семейство протоколов ISAKMP, которое рассматривается также в отдельном разделе.

Безопасность, обеспечиваемая IPSec, зависит от многих факторов операционного окружения, в котором IPSec выполняется. Например, от безопасности ОС, источника случайных чисел, плохих протоколов управления системой.

Размещение и функционирование IPSec

IPSec выполняется на хосте или шлюзе безопасности, обеспечивая защиту IP- трафика. Термин «шлюз безопасности» используется для обозначения промежуточной системы, которая реализует IPSec-протоколы. Защита основана на требованиях, определенных в Базе Данных Политики Безопасности (Security Policy Database- SPD), определяемой и поддерживаемой системным администратором. Пакеты обрабатываются одним из трех способов на основании соответствия информации заголовка IP или транспортного уровня записям в SPD. Каждый пакет либо отбрасывается сервисом безопасности IPSec, либо пропускается без изменения, либо обрабатывается сервисом IPSec на основе применения определенной политики.

IPSec обеспечивает сервисы безопасности на IP-уровне, выбирая нужные протоколы безопасности, определяя алгоритмы, используемые сервисами, и предоставляя все криптографические ключи требуемым сервисам. IPSec может использоваться для защиты одного или нескольких путей между парой хостов, между парой шлюзов безопасности или между шлюзом безопасности и хостом.

IPSec использует два протокола для обеспечения безопасности трафика - Authentication Header (AH) и Encapsulating Security Payload (ESP). Хотя бы один из этих сервисов должен быть задействован при использовании ESP.

Эти протоколы могут применяться как по отдельности так и в комбинации с друг другом для обеспечения необходимого набора сервисов безопасности в IPv4 и IPv6. Каждый протокол поддерживает два режима использования: режим транспорта и режим туннелирования. В транспортном режиме протоколы обеспечивают защиту главным образом для протоколов более высокого уровня; в режиме туннелирования протоколы применяются для скрытия IP-заголовков исходных пакетов. Разница между двумя режимами рассматривается дальше.

IPSec позволяет системному администратору управлять детализацией, с которой предоставляется сервис безопасности. Например, можно создать единственный зашифрованный туннель между двумя безопасными шлюзами, или для каждого TCP соединения может быть создан зашифрованный туннель между парой хостов. IPSec позволяет указывать следующие параметры:

- а) какие сервисы используются, и в какой комбинации;
- б) необходимый уровень детализации применяемой защиты;
- в) алгоритмы, используемые для обеспечения безопасности на основе криптографии.

Существует несколько способов реализации IPSec на хосте или в соединении с роутером или firewall (для создания безопасного шлюза). Несколько общих примеров:

а) интеграция IPSec в конкретную реализацию IP, что требует доступа к исходному коду IP и применимо как к хостам, так и к шлюзам безопасности;

б) bump-in-the-stack (BITS) реализации, где IPSec действует внизу существующей реализации стека протоколов IP, между обычным IP и локальными сетевыми драйверами; доступа к исходному коду стека IP в данном контексте не требуется, что делает такой подход пригодным для встраивания в существующие системы, и реализации на хостах;

в) использование внешнего криптопроцессора (обычно в военных и в некоторых коммерческих системах), как правило, это является Bump-in-the-stack (BITS) реализацией, используется как на хостах, так и на шлюзах, обычно BITS-устройства являются IP-адресуемыми;

Транспортный режим работы

В этом варианте механизмы безопасности применяются только для протоколов, начиная с транспортного (TCP) уровня и выше, оставляя данные самого сетевого уровня (заголовок IP) без дополнительной защиты. Места размещения дополнительной информации, вставляемой протоколами в пакет, представлены в соответствии с рисунком 2.3.



Рис. 9.17. Транспортный режим

Туннельный режим работы

Этот режим интересен тем, что обеспечивает защиту также и данных сетевого уровня путем добавления нового IP-заголовка. После определения ассоциаций безопасности (например, между двумя шлюзами) истинные адреса хостов отправления и назначения (и

другие служебные поля) полностью защищаются от модификаций для АН или вообще скрываются для ESP, а в новый заголовок выставляются адреса и другие данные для шлюзов (отправления/получения). В соответствии с рисунком 2.4 видны преимущества и недостатки обоих протоколов. ESP обеспечивает сокрытие данных, но не полную аутентификацию всего пакета. АН полностью аутентифицирует, но не скрывает данные. В этом причина того, что для обеспечения высокого уровня безопасности, применение протоколов совмещается.



Рис. 9.18. Туннельный режим

Контексты безопасности и управление ключами

Формирование контекстов безопасности в IPSec разделено на две фазы. Сначала создается управляющий контекст, назначение которого - предоставить доверенный канал, т. е. аутентифицированный, защищенный канал для выработки (в рамках второй фазы) протокольных контекстов и, в частности, для формирования криптографических ключей, используемых протоколами АН и ESP.

В принципе, для функционирования механизмов IPSec необходимы только протокольные контексты; управляющий играет вспомогательную роль. Более того, явное выделение двух фаз утяжеляет и усложняет формирование ключей, если рассматривать последнее как однократное действие. Тем не менее, из архитектурных соображений управляющие контексты не только могут, но и должны существовать, поскольку обслуживают все протокольные уровни стека TCP/IP, концентрируя в одном месте необходимую функциональность. Первая фаза начинается в ситуации, когда взаимодействующие стороны не имеют общих секретов (общих ключей) и не уверены в аутентичности друг друга. Если с самого начала не создать доверенный канал, то для выполнения каждого управляющего

действия с ключами (их модификация, выдача диагностических сообщений и т.п.) в каждом протоколе (AH, ESP, TLS и т.д.) этот канал придется формировать заново.

Общие вопросы формирования контекстов безопасности и управления ключами освещаются в спецификации «Контексты безопасности и управление ключами в Internet» (Internet Security Association and Key Management Protocol, ISAKMP). Здесь вводятся две фазы выработки протокольных ключей, определяются виды управляющих информационных обменов и используемые форматы заголовков и данных. Иными словами строится протоколно-независимый каркас.

Существует несколько способов формирования управляющего контекста. Они различаются двумя показателями:

- используемым механизмом выработки общего секретного ключа;
- степенью защиты идентификаторов общающихся сторон;

В простейшем случае секретные ключи задаются заранее (ручной метод распределения ключей). Для небольших сетей такой подход вполне работоспособен, но он не является масштабируемым. Последнее свойство может быть обеспечено при автоматической выработке и распределении секретных ключей в рамках протоколов, основанных на протоколе Диффи-Хеллмана. Пример тому - «Протокол для обмена ключами в Internet» (The Internet Key Exchange, IKE).

Протокол Диффи-Хеллмана (англ. Diffie-Hellman, DH) — криптографический протокол, позволяющий двум и более сторонам получить общий секретный ключ, используя незащищенный от прослушивания канал связи. Полученный ключ используется для шифрования дальнейшего обмена с помощью алгоритмов симметричного шифрования. Алгоритм был впервые опубликован Уитфилдом Диффи (Whitfield Diffie) и Мартином Хеллманом в 1976 году.[2]

При формировании управляющего контекста идентификаторы общающихся сторон (например, IP-адреса) могут передаваться в открытом виде или шифроваться. Поскольку ISAKMP предусматривает функционирование в режиме клиент/сервер (т. е. ISAKMP-сервер может формировать контекст для клиента), сокрытие идентификаторов в определенной степени повышает защищенность от пассивного прослушивания сети.

Последовательность передаваемых сообщений, позволяющих сформировать управляющий контекст и обеспечивающих защиту идентификаторов, выглядит в соответствии с рисунком 9.19.



Рис. 9.19. Формирование управляющего контекста

В первом сообщении (1) инициатор направляет предложения по набору защитных алгоритмов и конкретных механизмов их реализации. Предложения упорядочиваются по степени предпочтительности (для инициатора). В ответном сообщении (2) партнер информирует о сделанном выборе - какие алгоритмы и механизмы его устраивают. Для каждого класса защитных средств (генерация ключей, аутентификация, шифрование) выбирается только один элемент.

В сообщениях (3) и (4) инициатор и партнер отправляют свои части ключевого материала, необходимые для выработки общего секретного ключа (опускаются детали, специфичные для алгоритма Диффи-Хеллмана). Одноразовые номера (nonce) представляют собой псевдослучайные величины, служащие для защиты от воспроизведения сообщений.

Посредством сообщений (5) и (6) происходит обмен идентификационной информацией, подписанной (с целью аутентификации) секретным ключом отправителя и зашифрованной выработанным на предыдущих шагах общим секретным ключом. Для аутентификации предполагается использование сертификатов открытых ключей. В число подписываемых данных входят одноразовые номера.

В представленном виде протокол формирования управляющего контекста защищает от атак, производимых нелегальным посредником, а также от нелегального перехвата соединений. Для защиты от атак на доступность, для которых характерно прежде всего навязывание интенсивных вычислений, присущих криптографии с открытым ключом, применяются так называемые идентифицирующие цепочки (cookies). Эти цепочки, формируемые инициатором и его партнером с использованием текущего времени (для

защиты от воспроизведения), на самом деле присутствуют во всех ISAKMP-сообщениях и в совокупности идентифицируют управляющий контекст (в первом сообщении, по понятным причинам, фигурирует только цепочка инициатора). Согласно спецификациям, заголовок ISAKMP-сообщения имеет вид в соответствии с рисунком 4.20.

Если злоумышленник пытается «завалить» кого-либо запросами на создание управляющего контекста, подделывая при этом свой IP-адрес, то в сообщении (3) он не сможет предъявить идентифицирующую цепочку партнера, поэтому до выработки общего секретного ключа и, тем более, электронной подписи и полномасштабной проверки аутентичности дело попросту не дойдет.

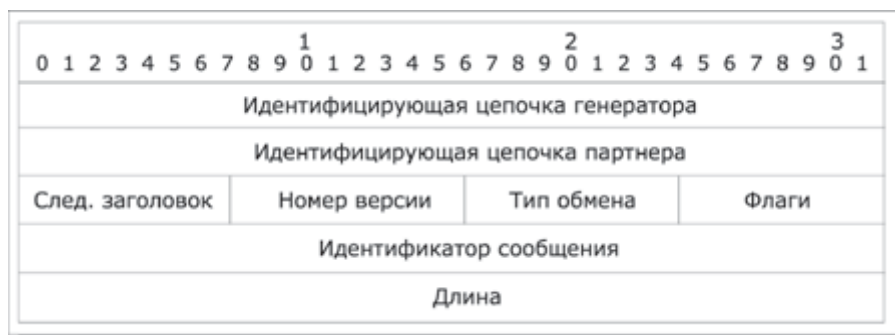


Рис. 9.20. Формат заголовка ISAKMP-сообщения

Управляющие контексты являются двунаправленными в том смысле, что любая из общающихся сторон может инициировать с их помощью выработку новых протокольных контекстов или иные действия. Для передачи ISAKMP-сообщений используется любой протокол, однако в качестве стандартного принят UDP с номером порта 500.

Протокольные контексты и политика безопасности

Системы, реализующие IPSec, должны поддерживать две базы данных:

- базу данных политики безопасности (Security Policy Database, SPD);
- базу данных протокольных контекстов безопасности (Security Association Database, SAD);

Все IP-пакеты (входящие и исходящие) сопоставляются с упорядоченным набором правил политики безопасности. При сопоставлении используется фигурирующий в каждом правиле селектор - совокупность анализируемых полей сетевого уровня и более высоких протокольных уровней. Первое подходящее правило определяет дальнейшую судьбу пакета:

- пакет может быть ликвидирован;
- пакет может быть обработан без участия средств IPSec;

- пакет должен быть обработан средствами IPSec с учетом набора протокольных контекстов, ассоциированных с правилом.

Таким образом, системы, реализующие IPSec, функционируют как межсетевые экраны, фильтруя и преобразуя потоки данных на основе предварительно заданной политики безопасности.

Далее рассматриваются контексты и политика безопасности, а также порядок обработки сетевых пакетов.

Протокольный контекст безопасности в IPSec - это однонаправленное соединение (от источника к получателю), предоставляющее обслуживаемым потокам данных набор защитных сервисов в рамках какого-то одного протокола (AH или ESP). В случае симметричного взаимодействия партнерам придется организовать два контекста (по одному в каждом направлении). Если используются и AH, и ESP, потребуется четыре контекста.

Элементы базы данных протокольных контекстов содержат следующие поля (в каждом конкретном случае некоторые значения полей будут пустыми):

- используемый в протоколе AH алгоритм аутентификации, его ключи и т.п.;
- используемый в протоколе ESP алгоритм шифрования, его ключи, начальный вектор и т.п.;
- используемый в протоколе ESP алгоритм аутентификации, его ключи и т.п.;
- время жизни контекста;
- режим работы IPSec: транспортный или туннельный;
- максимальный размер пакетов;
- группа полей (счетчик, окно, флаги) для защиты от воспроизведения пакетов.

Пользователями протокольных контекстов, как правило, являются прикладные процессы. Вообще говоря, между двумя узлами сети может существовать произвольное число протокольных контекстов, так как число приложений в узлах произвольно. В качестве пользователей управляющих контекстов обычно выступают узлы сети (поскольку в этих контекстах желательно сосредоточить общую функциональность, необходимую сервисам безопасности всех протокольных уровней эталонной модели для управления криптографическими ключами).

Управляющие контексты - двусторонние, т. е. любой из партнеров может инициировать новый ключевой обмен. Пара узлов может одновременно поддерживать несколько активных управляющих контекстов, если имеются приложения с существенно разными криптографическими требованиями. Например, допустима выработка части ключей на

основе предварительно распределенного материала, в то время как другая часть порождается по алгоритму Диффи-Хеллмана.

Протокольный контекст для IPSec идентифицируется целевым IP-адресом, протоколом (AH или ESP), а также дополнительной величиной - индексом параметров безопасности (Security Parameter Index, SPI). Последняя величина необходима, поскольку могут существовать несколько контекстов с одинаковыми IP-адресами и протоколами. Далее показано, как используются индексы SPI при обработке входящих пакетов.

IPSec обязывает поддерживать ручное и автоматическое управление контекстами безопасности и криптографическими ключами. В первом случае все системы заранее снабжаются ключевым материалом и иными данными, необходимыми для защищенного взаимодействия с другими системами. Во втором - материал и данные вырабатываются динамически, на основе определенного протокола - IKE, поддержка которого обязательна.

Протокольный контекст создается на базе управляющего с использованием ключевого материала и средств аутентификации и шифрования последнего. В простейшем случае, когда протокольные ключи генерируются на основе существующих, последовательность передаваемых сообщений выглядит в соответствии с рисунком 2.7.

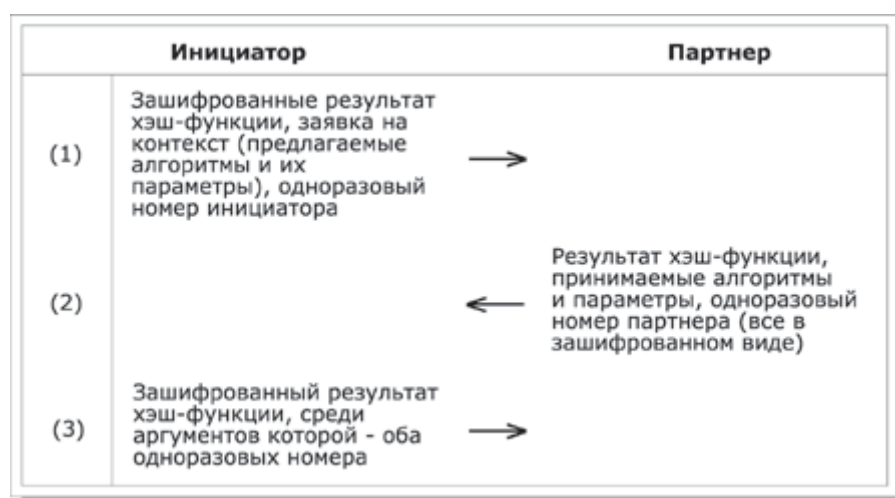


Рис. 4.21. Формирование протокольного контекста

Аутентификационный заголовок

Аутентификационный заголовок (англ. authentication header) AH предназначен для обеспечения аутентификации отправителя, контроля целостности данных и опционально для предотвращения повторной посылки (англ. replay) пакета - при условии, что принимающая сторона настроена производить проверку последовательного номера пакета. Поля IP-пакета, которые изменяются в пути следования, не подлежат контролю целостности. AH защищает

данные протоколов более высоких уровней и те поля IP-заголовков, которые не меняются на маршруте доставки или меняются предсказуемым образом (число «непредсказуемых» полей невелико - это prio. (Traffic Class), Flow Label и Hop Limit. Предсказуемо меняется целевой адрес при наличии дополнительного заголовка исходящей маршрутизации.).

Формат заголовка AH показан в соответствии с рисунком 2.8. поле «Следующий заголовок» (Next header) идентифицирует тип следующих значимых данных за аутентификационным заголовком; значения поля берутся из предопределенного множества установленных IANA (Internet Assigned Numbers Authority) номеров.

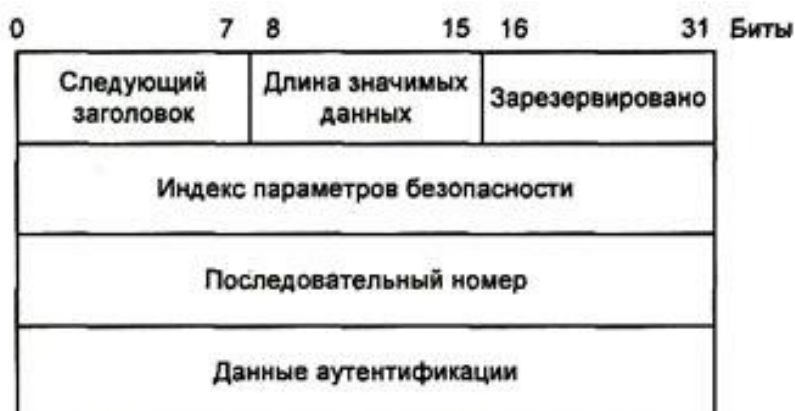


Рис. 9.22. Формат заголовка AH

Поле «Длина значимых данных» (Payload length) определяет длину самого AH в 32-битовых словах минус 2 слова, поскольку для заголовков, расширяемых для IPv6, установлено требование сокращения длины заголовка на 64 бита;

Зарезервированное поле должно быть нулевым;

Поле «Индекс параметров безопасности» (Security Parameters Index - SPI) - это значение, которое в совокупности с адресом назначения и самим протоколом (в данном случае AH) однозначно определяет ассоциацию безопасности (Security Association - SA) для данной датаграммы в виде 32-битного номера; номера с 1 по 255 зарезервированы IANA; SPI, равный нулю, означает, что SA не установлена; ассоциация безопасности - это набор параметров (версия алгоритмов шифрования и аутентификации, схема обмена ключами и т. п.), определяющих, каким образом будет обеспечиваться защита данных;

Поле «Последовательный номер» (Sequence number) - это монотонно возрастающий от 0 (при установлении SA) номер пакета. Он используется для возможности контроля получателем ситуации повторной пересылки пакетов;

Поле «Данные аутентификации» (Authentication data) содержат значение контроля целостности (Integrity check value - ICV), рассчитанное по всем данным, которые не

изменяются в пути следования пакета или предсказуемы на момент достижения им получателя. Значение ICV рассчитывается в зависимости от алгоритма, определенного в SA, например код аутентификации сообщения (Message Authentication Code - MAC) с ключом симметричного или асимметричного алгоритма или хэшфункции;

Безопасное сокрытие существенных данных

Протокол инкапсулирующей защиты содержимого (Encapsulating Security payload, ESP) предоставляет три вида сервисов безопасности:

1. обеспечение конфиденциальности (шифрование содержимого IP-пакетов, а также частичная защита от анализа трафика путем применения туннельного режима);
2. обеспечение целостности IP-пакетов и аутентификации источника данных;
3. обеспечение защиты от воспроизведения IP-пакетов;

Функциональность ESP шире, чем у AH (добавляется шифрование); ESP не обязательно предоставляет все сервисы, но либо конфиденциальность, либо аутентификация должны быть задействованы. Формат заголовка ESP выглядит в соответствии с рисунком 2.9. Это не столько заголовок, сколько обертка (инкапсулирующая оболочка) для зашифрованного содержимого. Например, ссылку на следующий заголовок нельзя выносить в начало, в незашифрованную часть, так как она лишится конфиденциальности.

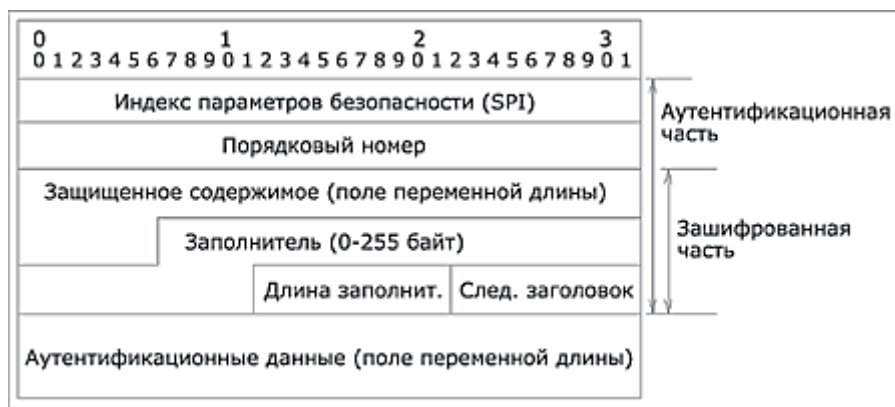


Рисунок 4.23. Формат заголовка ESP

Поля "Индекс параметров безопасности (SPI)", "Порядковый номер" и "Аутентификационные данные" (последнее присутствует только при включенной аутентификации) имеют тот же смысл, что и для AH. ESP аутентифицирует лишь зашифрованную часть пакета (плюс два первых поля заголовка).

Применение протокола ESP к исходящим пакетам можно представлять себе следующим образом. Пусть остаток пакета та его часть, которая помещается после предполагаемого места вставки заголовка ESP. При этом не важно, какой режим используется - транспортный или туннельный. Шаги протокола таковы:

1. - остаток пакета копируется в буфер;
2. - к остатку приписываются дополняющие байты, их число и номер (тип) первого заголовка остатка, с тем чтобы номер был прижат к границе 32-битного слова, а размер буфера удовлетворял требованиям алгоритма шифрования;
3. - текущее содержимое буфера шифруется;
4. - в начало буфера приписываются поля "Индекс параметров безопасности (SPI)" и "Порядковый номер" с соответствующими значениями;
5. - пополненное содержимое буфера аутентифицируется, в его конец помещается поле "Аутентификационные данные";
6. - в новый пакет переписываются начальные заголовки старого пакета и конечное содержимое буфера;

Если в ESP включены и шифрование, и аутентификация, то аутентифицируется зашифрованный пакет. Для входящих пакетов действия выполняются в обратном порядке, т. е. сначала производится аутентификация. Это позволяет не тратить ресурсы на расшифровку поддельных пакетов, что в какой-то степени защищает от атак на доступность.

Протокол обмена ключами – IKE

Поскольку основным механизмом обеспечения безопасности данных протокола являются криптографические методы, участники защищенного соединения должны наладить обмен соответствующими криптографическими ключами. Обеспечить настройку процесса такого обмена можно вручную и автоматически. Первый способ допустим для небольшого количества достаточно статичных систем, а в общем случае это производится автоматически.

Для автоматического обмена ключами по умолчанию используется Протокол управления ключами в Интернете (Internet Key Management Protocol - IKMP), иначе называемый Обмен ключами в Интернете (Internet Key Exchange - IKE). Дополнительно или альтернативно могут быть применены другие протоколы, такие как Kerberos или SKIP.

IKE совмещает в себе три основных направления (отдельных протокола):

- ISAKMP (Internet Security Association and Key Management Protocol) - протокол ассоциаций безопасности и управления ключами в интернете; это общее описание

(framework) для обеспечения аутентификации и обмена ключей без указания конкретных прикладных алгоритмов;

- Oakley (Oakley key determination protocol) - протокол определения ключей Окли; он описывает последовательности обмена ключами - моды (mode) и описывает предоставляемые ими функции;
- SKEMI (Secure Key Exchange Mechanism for Internet) - механизм безопасного обмена ключами в Интернете; он описывает многофункциональные технологии, предоставляющие анонимность, неотрекаемость (апеллируемость) и быстрое обновление ключей;

ИКЕ содержит две фазы согласования ключей. В первой фазе происходит создание защищенного канала, во второй - согласование и обмен ключами, установление SA. Первая фаза использует один из двух режимов: основной (англ. Main Mode) или агрессивный (англ. Aggressive Mode). Различие между ними в уровне защищенности и скорости работы. Основной режим, более медленный, защищает всю информацию, передаваемую между узлами. Агрессивный режим для ускорения работы оставляет ряд параметров открытыми и уязвимыми для прослушивания, его рекомендуется использовать только в случае, когда критическим вопросом является скорость работы. Во второй фазе используется быстрый режим (англ. Quick Mode), названный так потому, что не производит аутентификации узлов, считая, что это было сделано в первой фазе. Эта фаза обеспечивает обмен ключами, с помощью которых происходит шифрование данных.

Расширенный обзор безопасных ассоциаций

Понятие "безопасные ассоциации" (Security Association - SA) является фундаментальным в IPSec.

SA есть симплексное (однаправленное) логическое соединение, создаваемое для обеспечения безопасности. Весь трафик, передаваемый по SA, некоторым образом обрабатывается в целях обеспечения безопасности. И AH, и ESP используют в своей работе SAs. Одной из основных функций IKE является установление SA. Далее приводятся различные аспекты управления SA, определим требуемые характеристики управления политикой SA, обработку трафика и технологии управления SA.

SA есть совокупность параметров соединения, которые дают возможность сервисам обеспечивать безопасный трафик. SA определяет использование AH или ESP. Если к потоку трафика применяются оба протокола, AH и ESP, то создаются две SAs. При

двунаправленном соединении между двумя хостами или между двумя шлюзами безопасности требуется два SA (по одному на каждое направление).

SA однозначно определяется тройкой, состоящей из Security Parameter Index (SPI), IP Destination Address (адресом назначения) и идентификатора протокола безопасности (AH или ESP). В принципе адрес назначения может быть единственным адресом, широковещательным (broadcast) адресом или групповым (multicast) адресом. Однако механизм управления SA в настоящее время определяется только для единственной SA. Следовательно, SAs будут описаны в контексте point-to-point соединения, даже если концепция также применяется в случае point-to-multipoint.

Определены два режима SA: режим транспорта и режим туннелирования. Транспортный режим SA обеспечивает безопасную связь между двумя хостами. В IPv4 заголовок протокола безопасности транспортного режима появляется сразу после IP заголовка и всех опций и перед любыми протоколами более высокого уровня (TCP или UDP). В случае ESP транспортный режим SA обеспечивает сервисы безопасности только для протоколов более высокого уровня, но не для IP-заголовка. В случае AH защита также распространяется на отдельные части IP-заголовка.

Другим режимом SA является режим туннелирования. Если хотя бы одним из концов соединения является шлюз безопасности, то SA обязательно должна выполняться в туннелирующем режиме. SA между двумя шлюзами безопасности всегда находится в туннелирующем режиме, так же, как и SA между хостом и шлюзом безопасности. Заметим, что когда трафик предназначен для шлюза безопасности, например, в случае SNMP-команд, шлюз безопасности рассматривается как хост, и допустим транспортный режим. Два хоста могут при желании так же устанавливать туннелирующий режим.

В туннелирующем режиме SA существует "внешний" IP-заголовок, который определяет пункт назначения IPSec, и "внутренний" IP-заголовок, который определяет конечный пункт назначения для пакета. Заголовок протокола безопасности расположен после внешнего IP-заголовка и перед внутренним IP-заголовком. Если AH используется в туннелирующем режиме, части внешнего IP заголовка являются защищенными, как и весь туннелируемый IP-пакет, т.е. все внутренние заголовки защищены, как и все протоколы более высокого уровня. Если применяется ESP, защита обеспечивается только для туннелируемого пакета, а не для внешнего IP-заголовка.

9.2. Компьютерный практикум по сетевым протоколам

Программная реализация

Для создания защищенного соединения необходимо получить сертификат или создать его самому. Создать SSL сертификат самому можно средствами свободно распространяемого пакета IIS6 Resource Kit Tools. Может потребоваться установка служб. Для этого в Панели управления необходимо выбрать раздел «Программы и компоненты» и в открывшемся окне в меню слева нажать на ссылку «Включение или отключение компонентов Windows». Далее в открывшемся окне нужно включить компонент «Службы IIS».

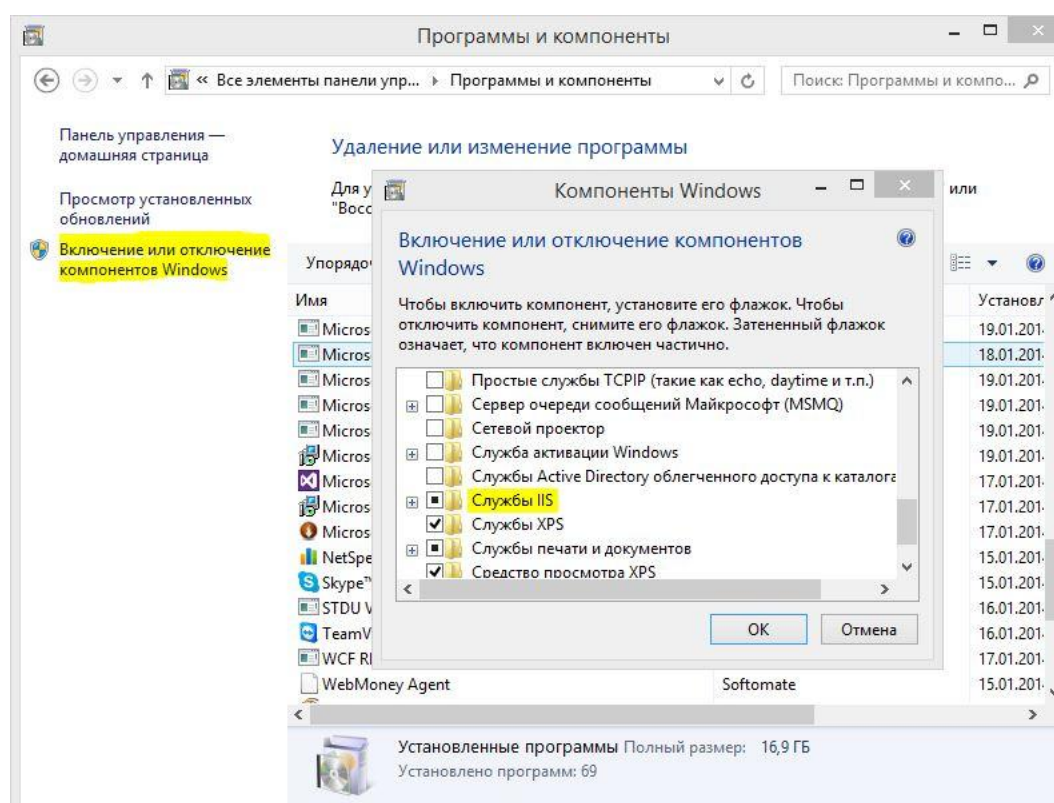


Рис. 9.24. Активация службы ISS

Выбираем службу ISS

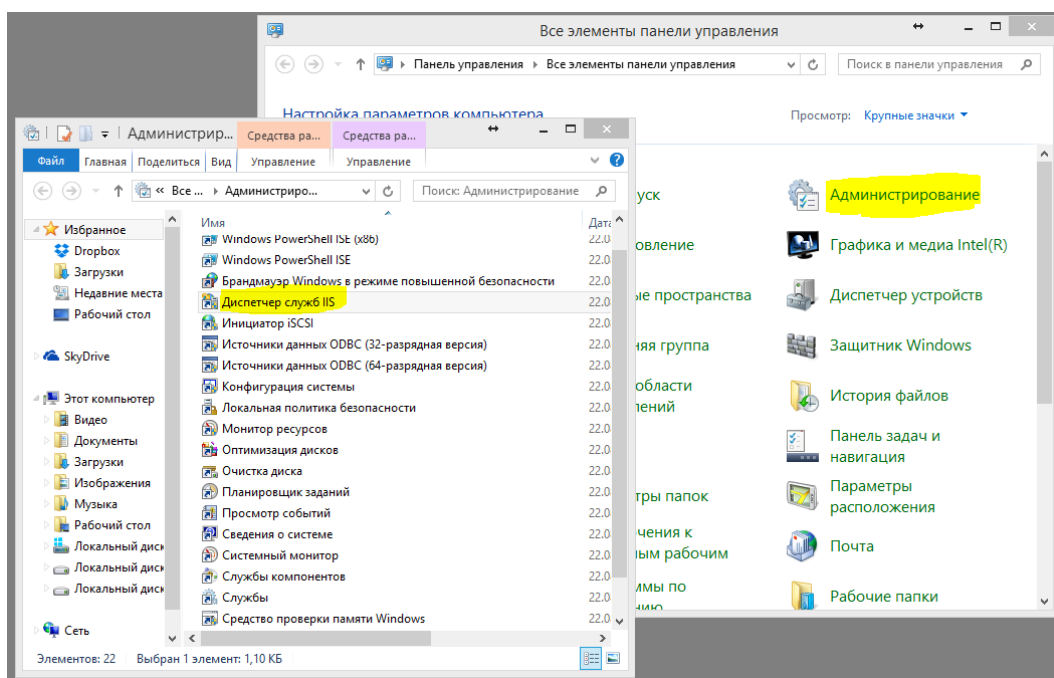


Рис. 9.25. Выбор службы ISS

В открывшемся окне жмем на «Сертификаты сервера»

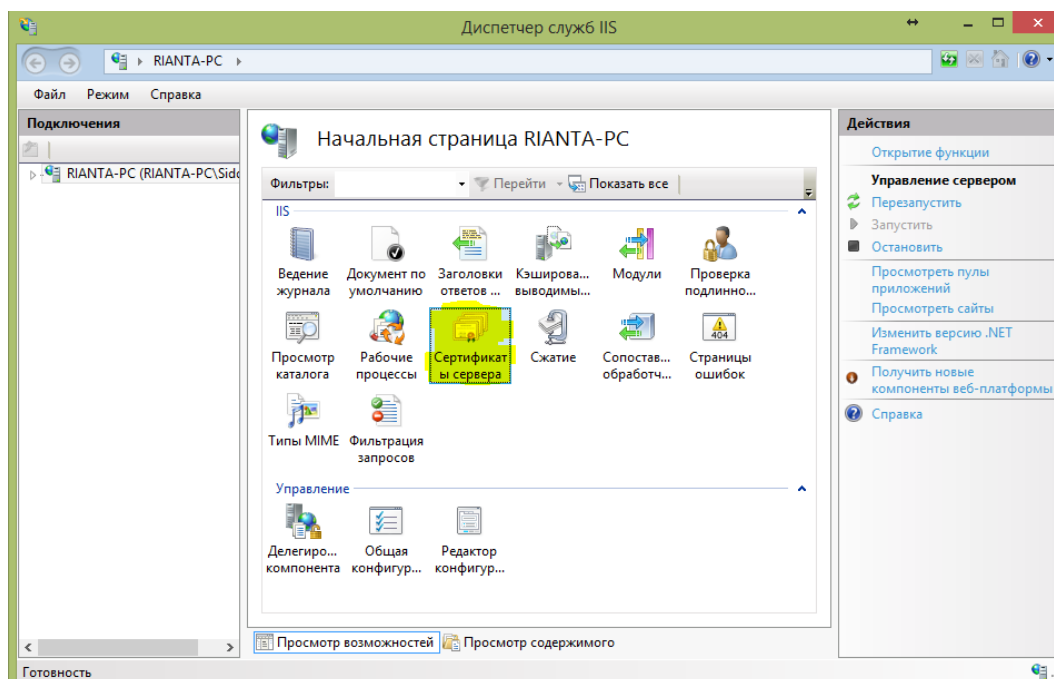


Рис. 9.26. Создание сертификата

Создаем самоподписанный сертификат

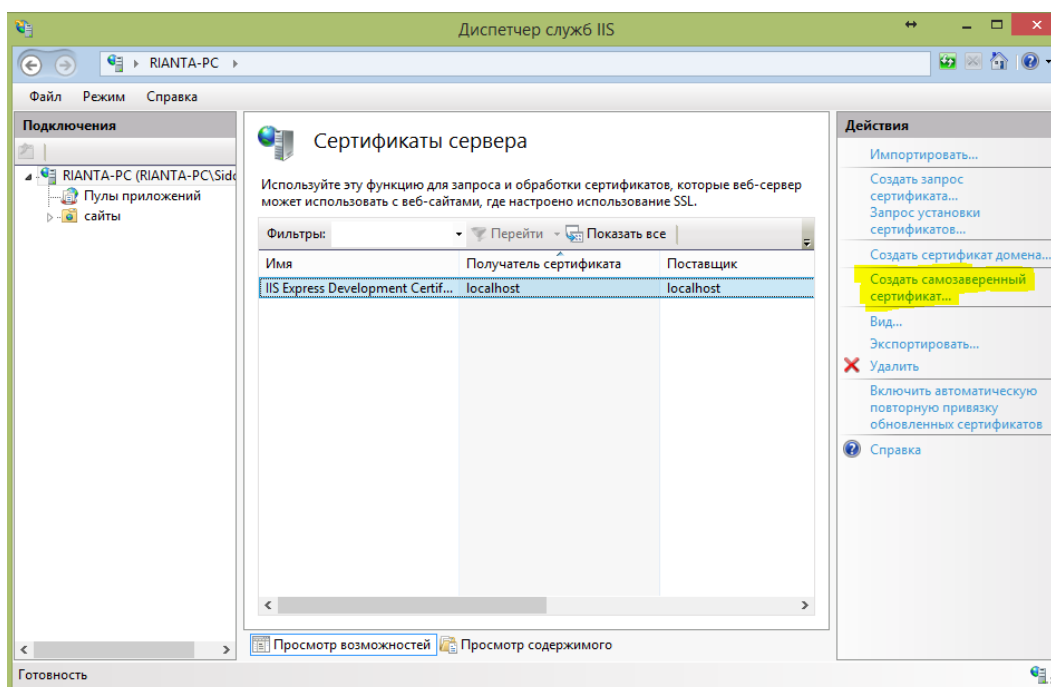


Рис. 9.27. Создание самозаверенного сертификата

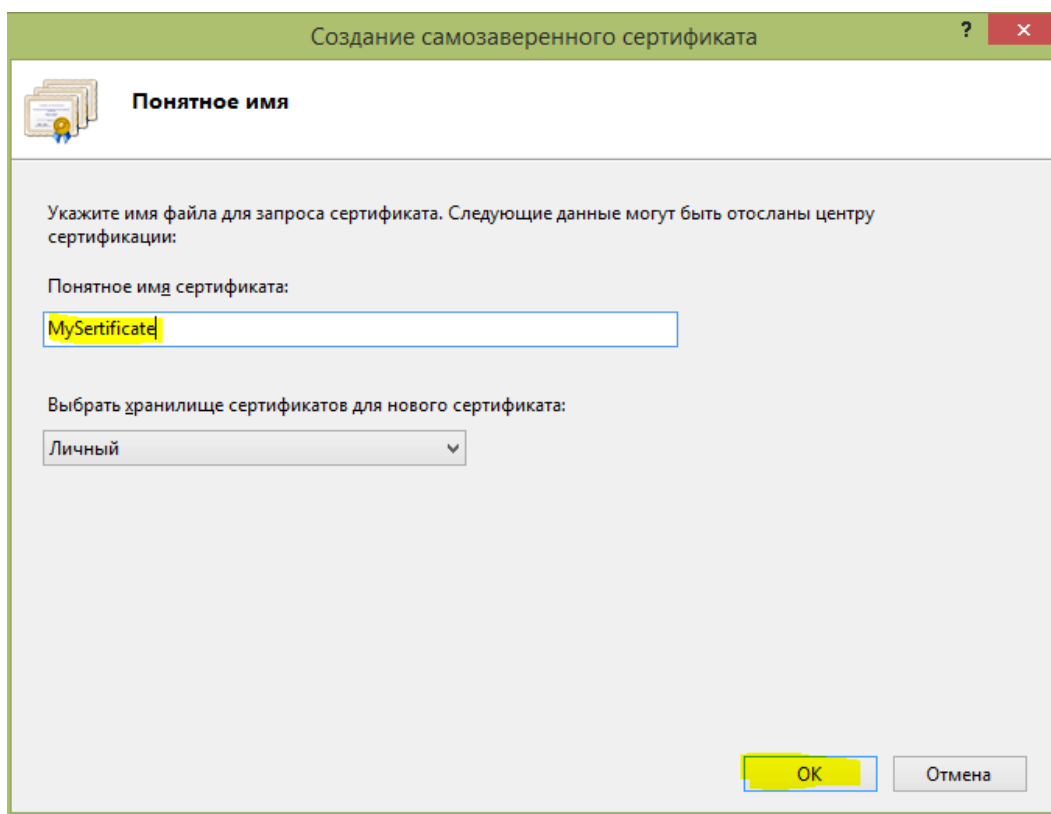


Рис. 9.28. Ввод имени сертификата

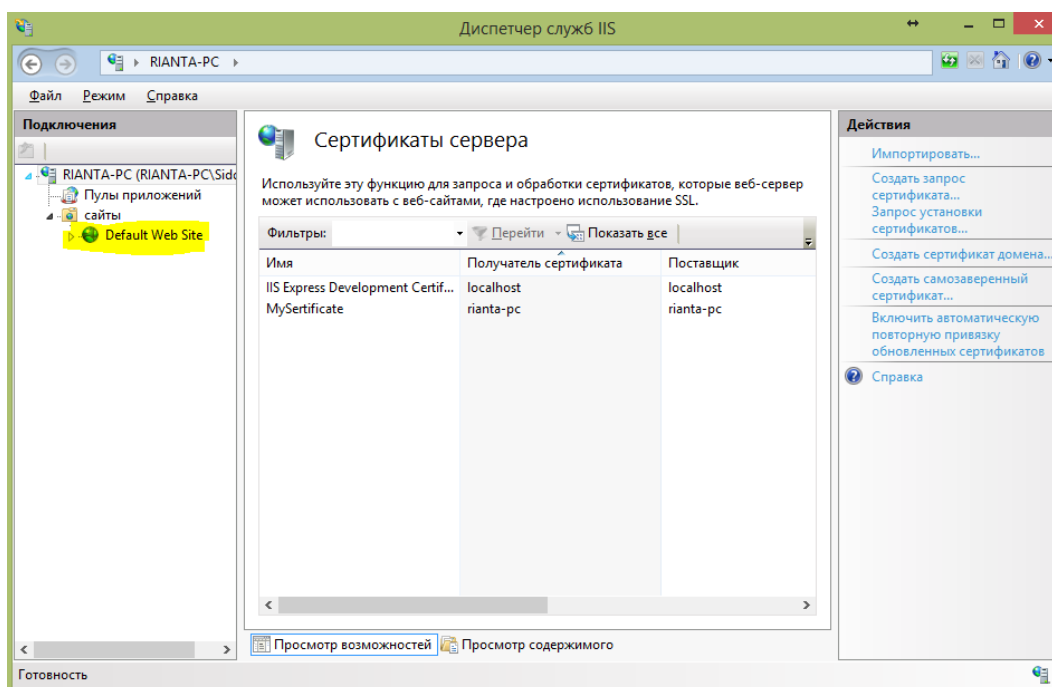


Рис. 9.29. Отображение сертификата

Привязываем сертификат к нужному IP

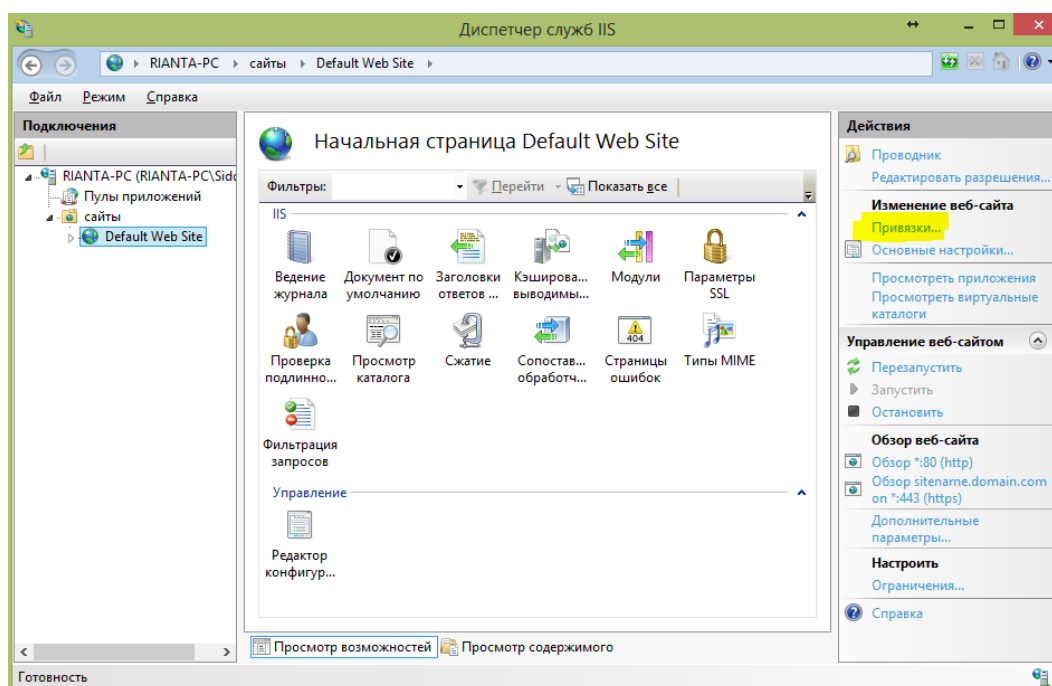


Рис. 9.30. Привязка сертификата к сайту (компьютеру)

Затем с помощью программы THEGREENBOW устанавливаем защищенное SSL – соединение.

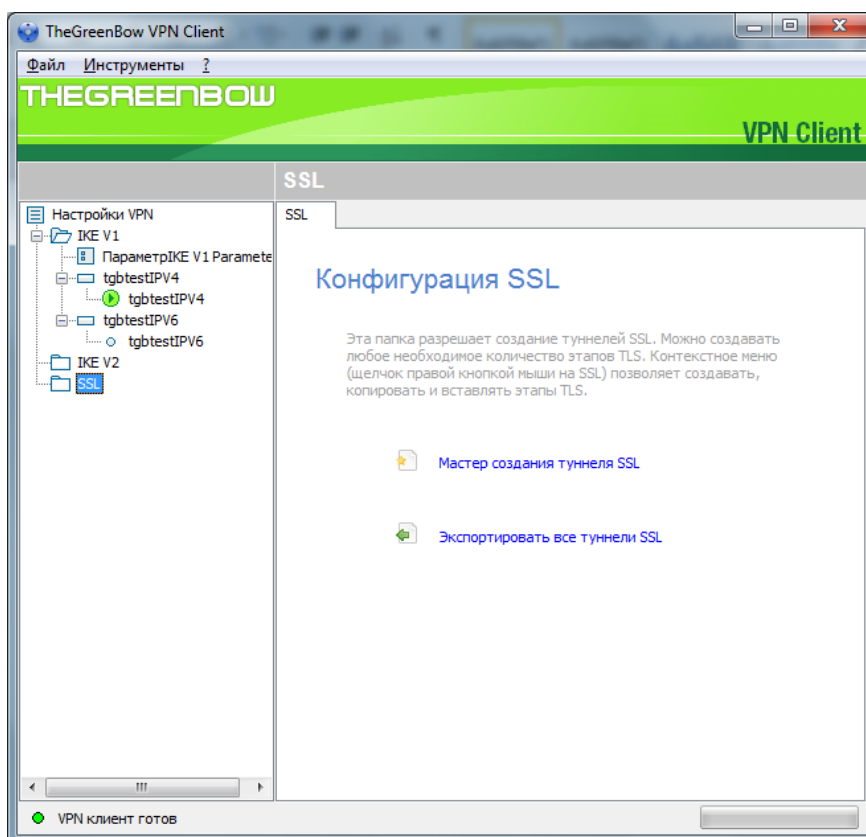


Рис. 9.31. Окно раздела SSL

Указываем путь к сертификату

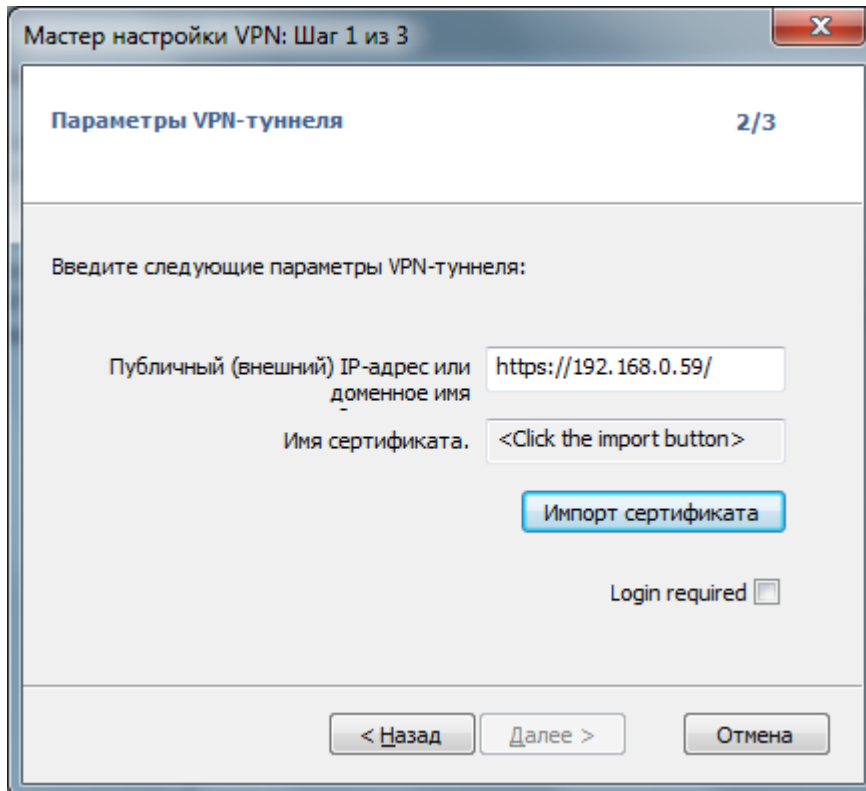


Рис. 9.32. Окно мастера настройки VPN-туннеля

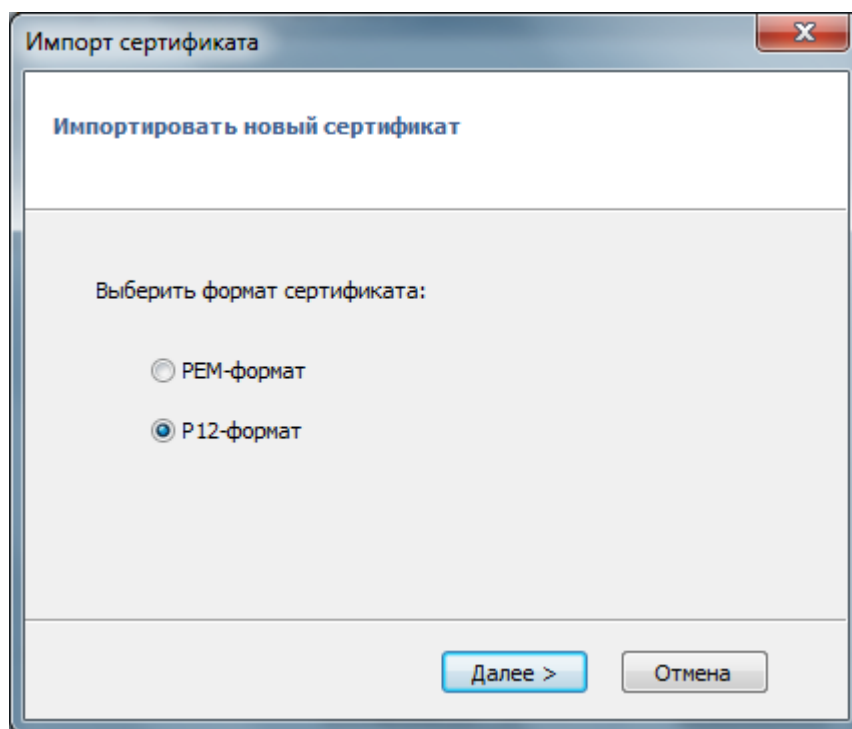


Рис. 9.33. Импорт сертификата

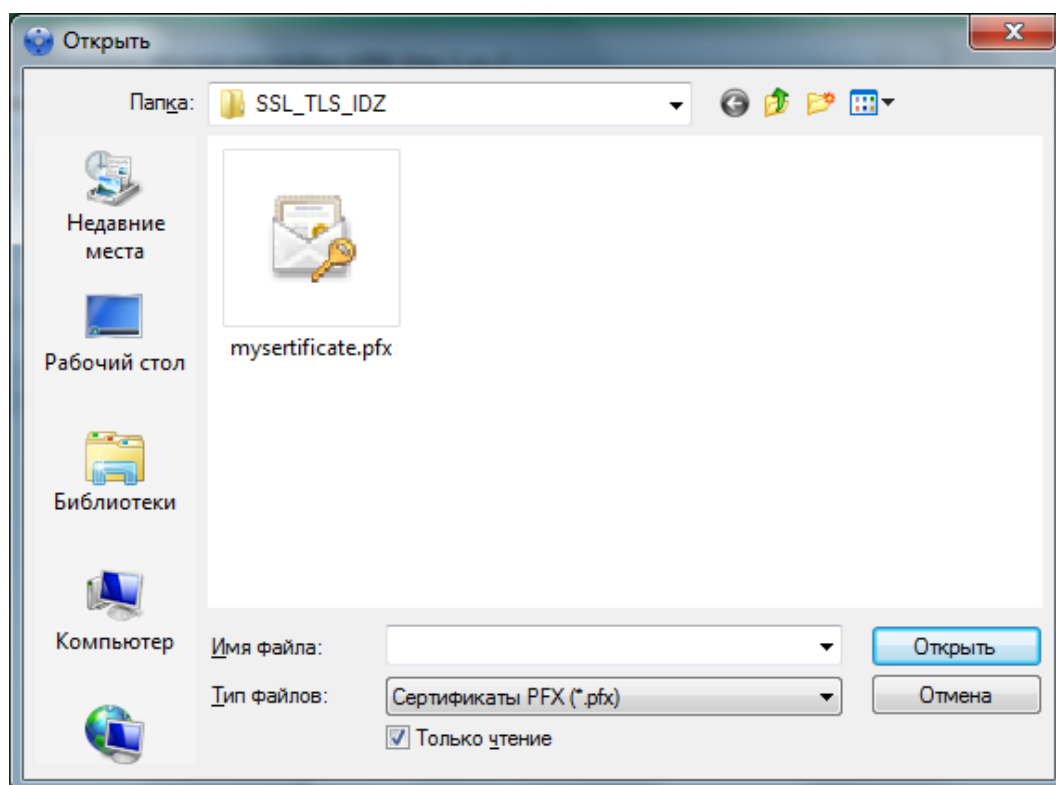


Рис. 9.34. Выбор созданного сертификата

Вводим пароль для доступа к сертификату

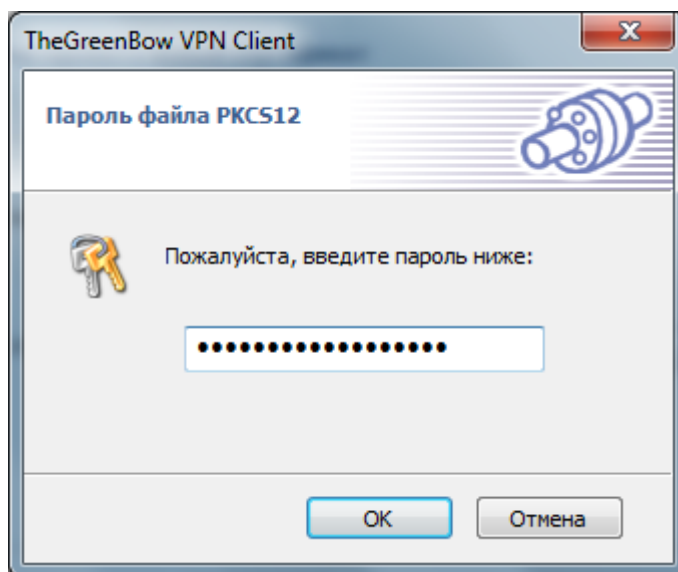


Рис. 9.35. Окно ввода пароля для доступа к сертификату

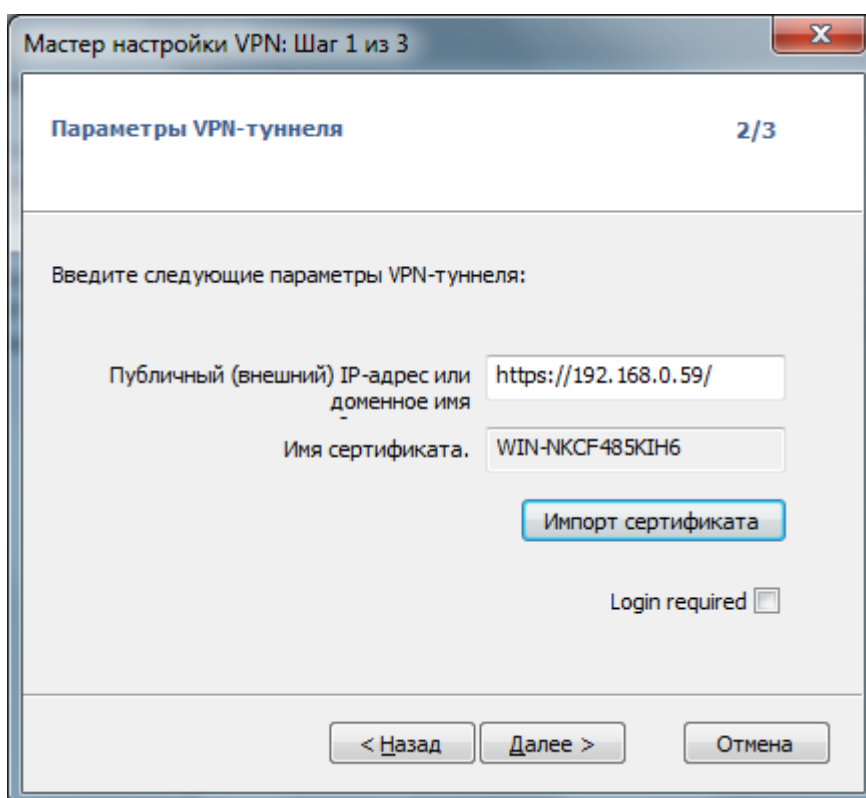


Рисунок 9.36. Окно мастера настройки VPN-туннеля

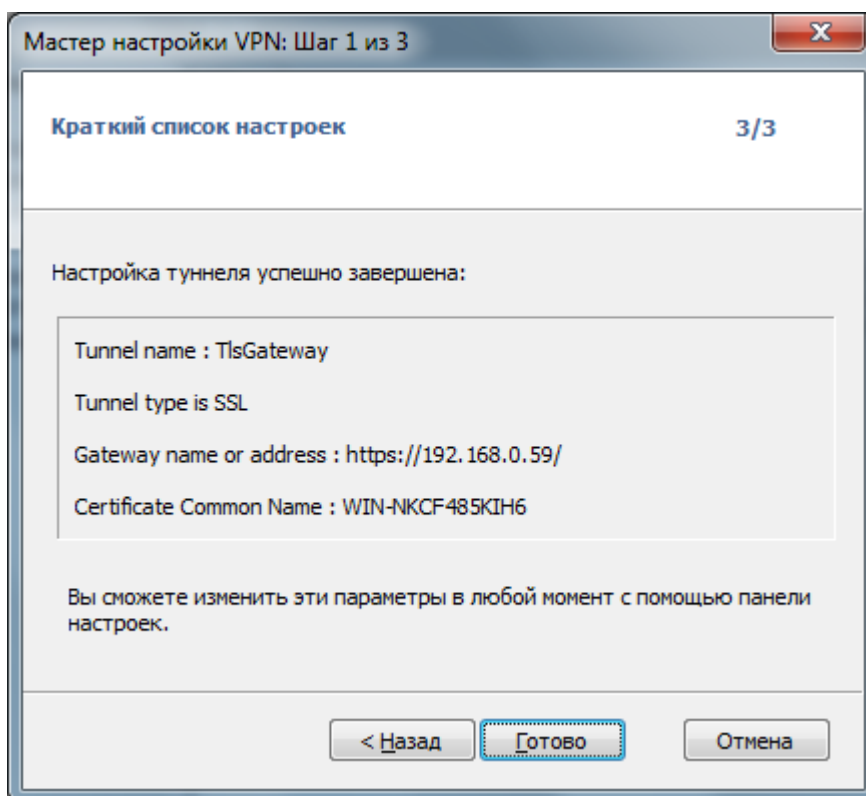


Рис. 9.37. Подтверждение об успешно созданном туннеле

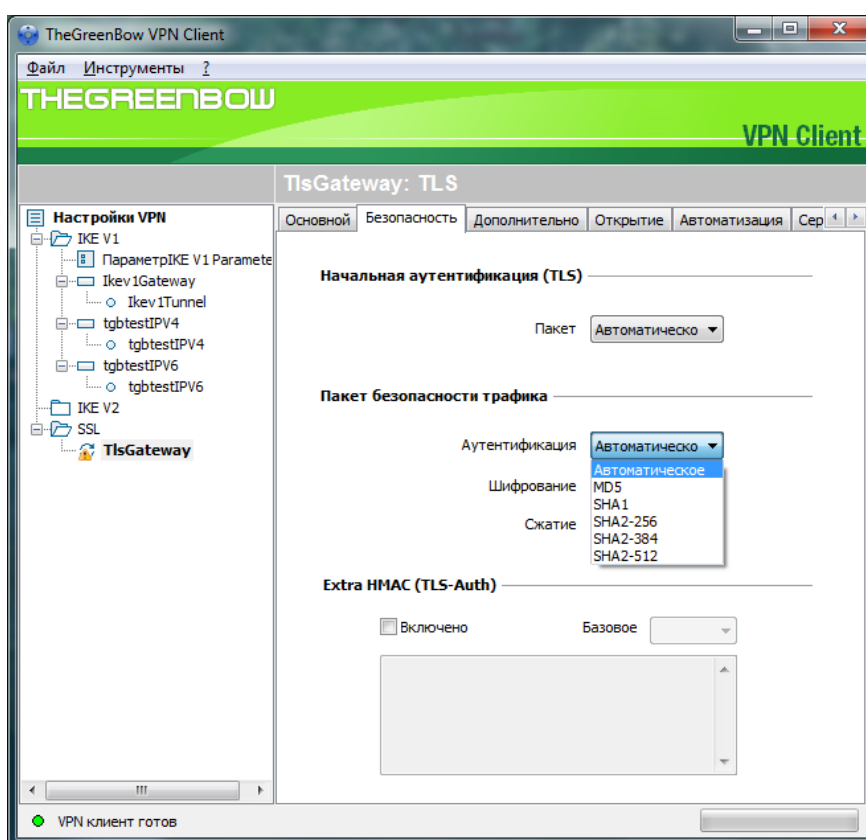


Рис. 9.38. Параметры аутентификации трафика

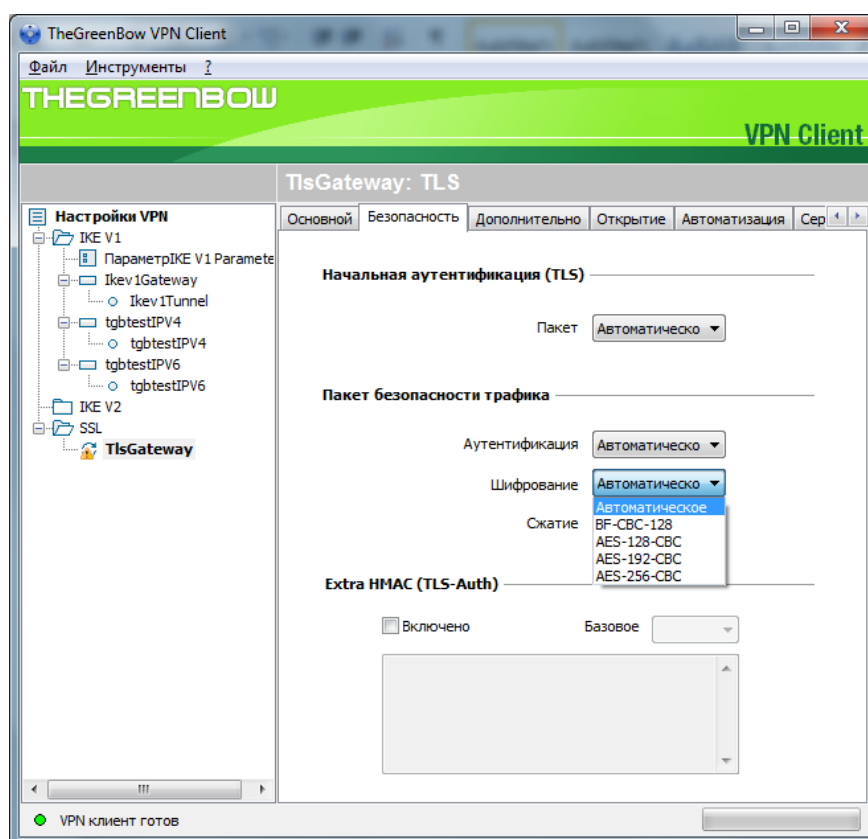


Рис. 9.39. Параметры шифрования

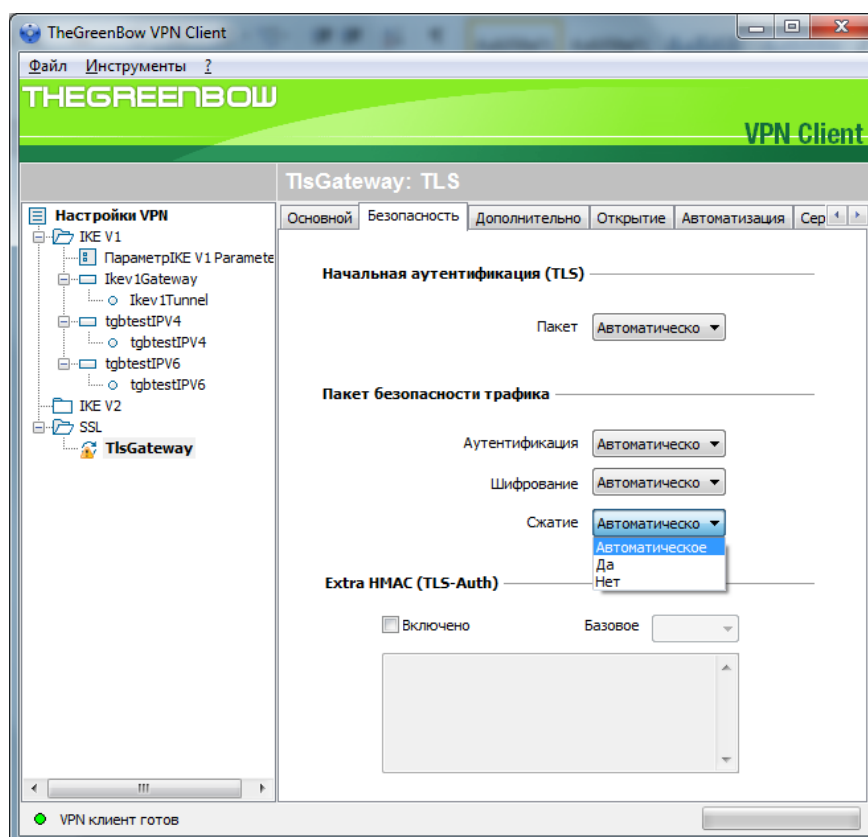


Рис. 9.40. Параметры сжатия

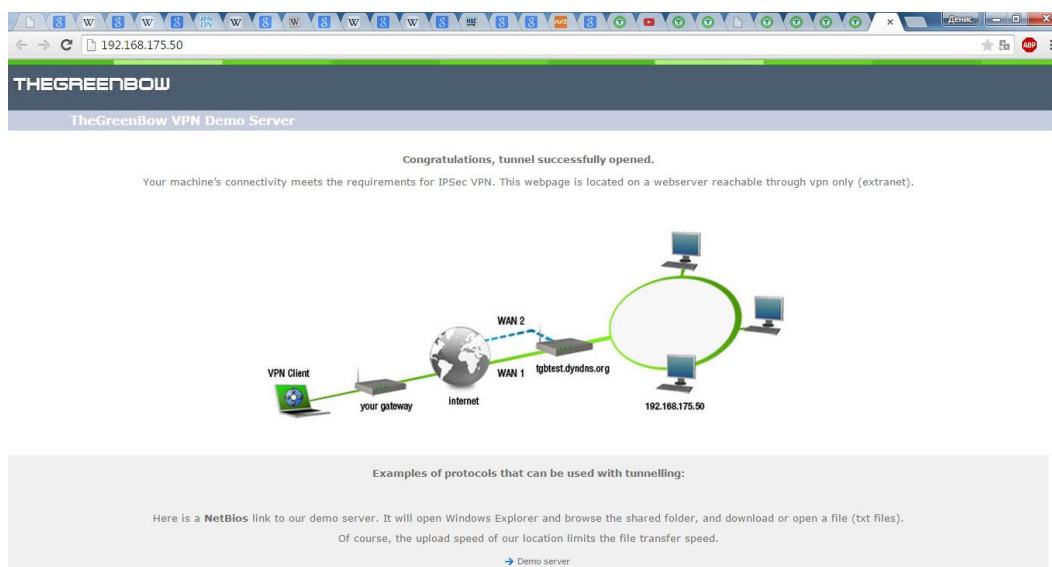


Рис. 9.41. Окно в браузере об успешном соединении

Выводы

Протокол безопасности транспортного уровня обеспечивает услуги безопасности из конца в конец для приложений, которые пользуются протоколами транспортного уровня, такими как, например, TCP. На сегодняшний день преобладает применение двух протоколов: *протокол "Уровень Безопасных Розеток" (SSL - Secure Sockets Layer)* и *протокол "Безопасность Транспортного уровня" (TLS - Transport Layer Security)*.

- SSL или TLS обеспечивают такие услуги, как фрагментация, сжатие, целостность сообщения, конфиденциальность и создание кадра данных, полученных от прикладного уровня. Как правило, SSL (или TLS) может получить прикладные данные от любого протокола прикладного уровня, но работает протокол обычно с *HTTP*.

- Комбинация алгоритмов смены ключей, хэширования и алгоритм шифрования определяют *набор шифров* для каждого сеанса.

- Для того чтобы обмениваться заверенными и конфиденциальными сообщениями, клиенту и серверу необходимо иметь шесть единиц криптографической секретности (четыре ключа и два вектора инициализации).

- В SSL (или TLS) отличают *подключение* и *сеанс*. В сеансе одна сторона играет роль клиента, а другая - роль сервера; при *подключении* обе стороны играют одинаковые роли, на равном подуровне.

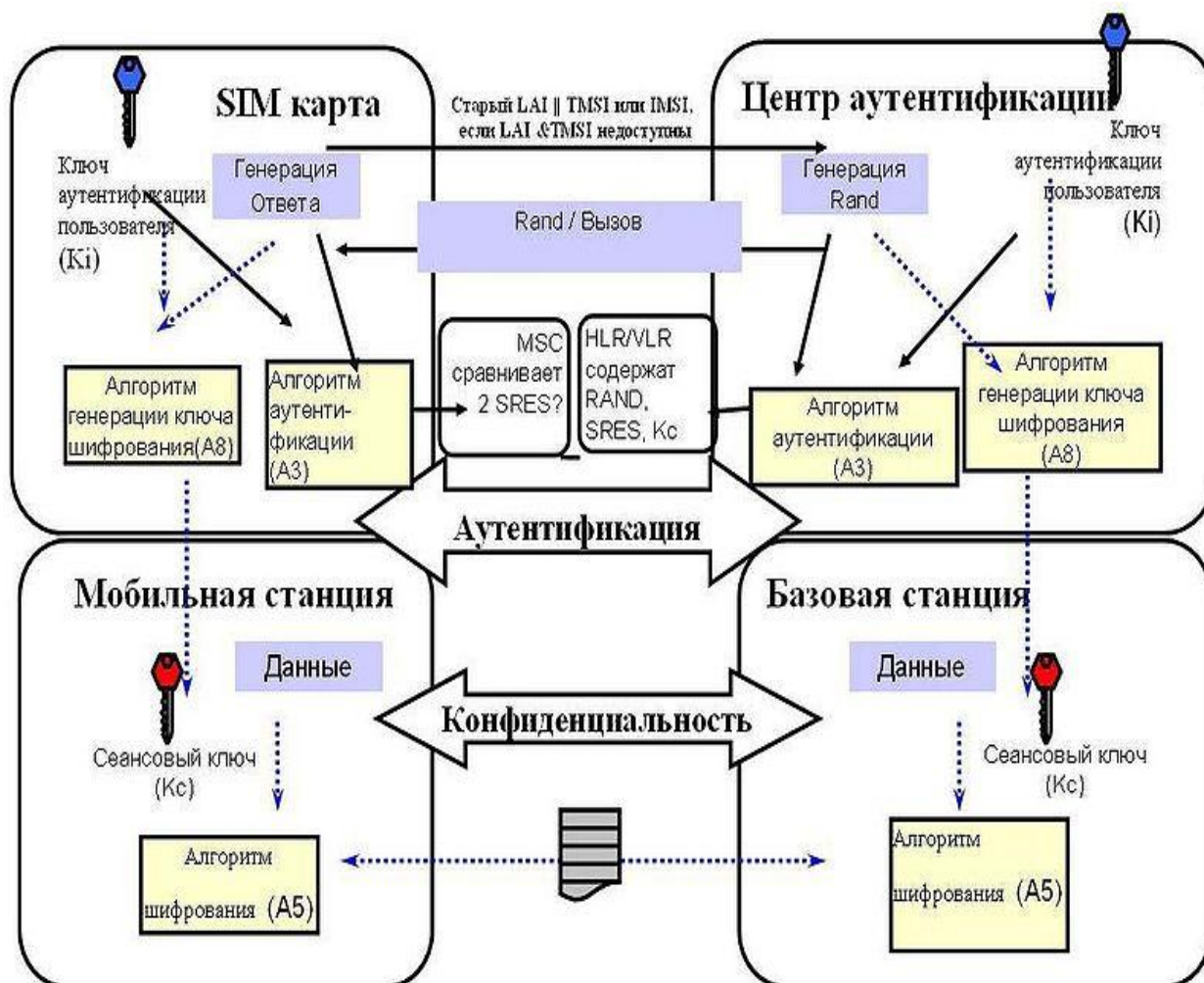
- SSL (или TLS) определяет четыре протокола на двух уровнях: *протокол установления соединения*, *протокол изменения параметров шифрования*, *аварийный*

протокол и *протокол передачи записей*. *Протокол установления соединения* использует несколько сообщений, чтобы договориться о наборе шифров, подтвердить подлинность сервера для клиента и клиента для сервера, если это необходимо, и обмениваться информацией для организации криптографической секретности. *Протокол изменения параметров шифрования* определяет процесс перемещения информации между состоянием ожидания и активным состоянием. *Аварийный протокол* передает извещения об ошибках и ситуациях, отклоняющихся от нормальных. *Протокол передачи записей* доставляет сообщения от верхнего уровня (*протокол установления соединения*, *аварийный протокол*, *ChangeCipherSpec-протокол*) или прикладного уровня.

10. ШИФРОВАНИЕ В СОВРЕМЕННЫХ СИСТЕМАХ СВЯЗИ

10.1. Безопасность GSM сетей

Прежде чем приступить к описанию алгоритма шифрования используемого в GSM сетях рассмотрим каким образом происходит аутентификация пользователя и формирования ключа шифрования. Для этого воспользуемся картинкой.



На данном рисунке схематично представлены следующие шаги:

1. Телефон оператора подключается к сети.
2. Для подтверждения своей подлинности телефон посылает специальный идентификационный код, называемый TMSI (Temporary Mobile Subscriber Identity).
3. Центр Аутентификации (ЦА) генерирует 128-битное случайное число RAND и посылает его на Мобильную Станцию (МС).
4. МС зашифровывает полученное число RAND, используя свой секретный ключ K_i и алгоритм аутентификации A3.

5. МС берет первые 32 бита из последовательности, полученной на предыдущем шаге(назовем их SRES(signed response)) и отправляет их обратно на ЦА.

6. ЦА проделывает ту же операцию и получает 32 битную последовательность XRES(expected response).

7. После чего ЦА сравнивает SRES и XRES. В случае, если оба значения равны, телефон считается аутентифицированным.

8. МС и ЦА вычисляют сессионный ключ шифрования, используя секретный ключ K_i и алгоритм формирования ключа A_8 $K_c = A_{8K_i}(\text{RAND})$

Говоря об алгоритмах аутентификации A_3 и алгоритме формирования ключа A_8 , следует отметить что на практике большинство сотовых операторов используют для этих целей один алгоритм, называемый COMP128(он имеет множество модификаций COMP128-1, COMP128-2, COMP128-3). COMP128 представляет собой обыкновенную хэш-функцию, на входе которая принимает 128-битную последовательность и на выходе возвращает 96-битную.

Как всегда в криптографии, попытка сэкономить время разработчикам обернулась полным провалом. Безопасность GSM сетей изначально основывалась на принципе «безопасность за счёт неизвестности». И когда в 1998 году алгоритм был вскрыт группой исследователей состоящих из Marc Briceno, Ian Goldberg и David Wagner обронузилась одна занятная особенность: последние 10 бит секретного ключа K_i всегда равнялись нулю. Используя это любопытное свойство, а так же уязвимость COMP128 к «атаке дней рождений» Marc Briceno, Ian Goldberg и David Wagner смогли извлечь секретный ключ K_i из SIM-карты.

Результатом этого исследования стал повсеместный отказ от алгоритма COMP128 и его замена на более надежные модификации COMP128-2 и COMP128-3, технические детали которых держатся в тайне.

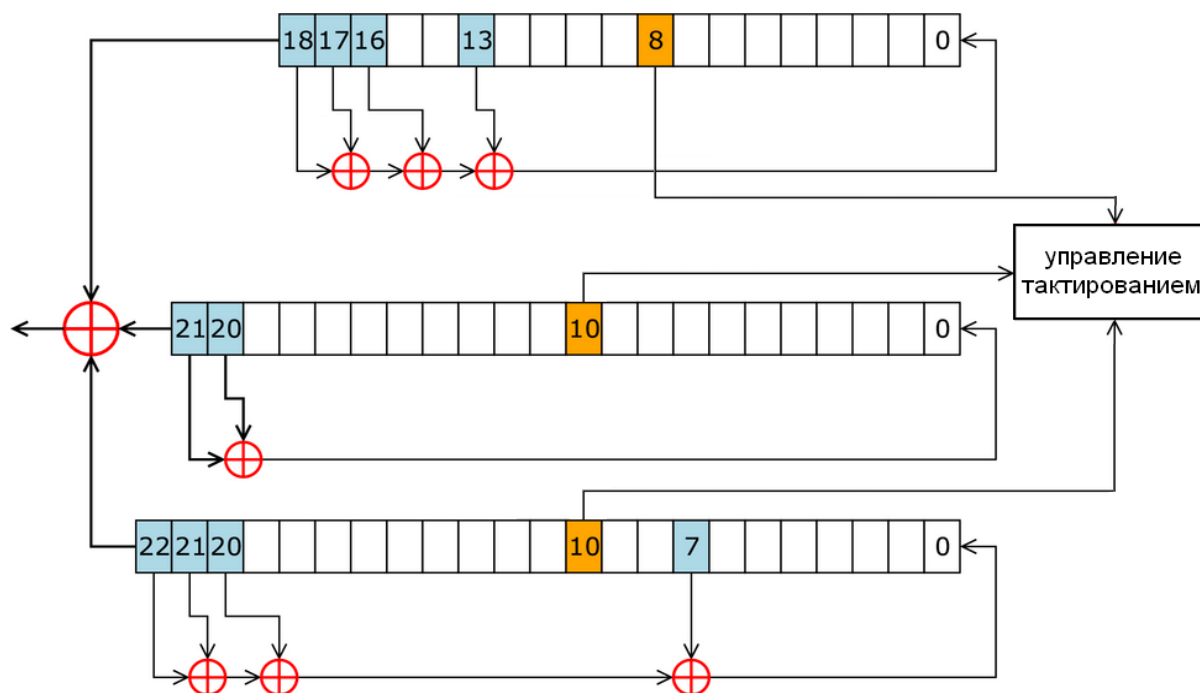
Алгоритм шифрования A5/1

В качестве алгоритма шифрования в GSM используются алгоритмы из семейства A5. На сегодняшний день их всего 3:

- **A5/1** — поточный шифр, наиболее распространенный на сегодня.
- **A5/2**-вариант предыдущего алгоритма «для бедных». Очень похож на своего «старшего брата», но изначально задумывался, как сильно ослабленная версия A5/1. В настоящее время не используется

- **A5/3-блочный шифр.** Разработан в 2002 году с целью заменить устаревший A5/1. Однако в настоящее время используется только в 3GPP сетях. У алгоритма найден ряд уязвимостей, но о практических атаках речи пока не идет.

Рассмотрим подробнее алгоритм A5/1.



Внутреннее состояние шифра A5/1 состоит из трех линейных регистров сдвига с обратной связью R1, R2, R3, длиной 19, 22 и 23 бита соответственно (всего 64 бита).

Сдвиг в регистрах R1, R2, R3 происходит только при выполнении определенного условия. Каждый регистр содержит "бит управления тактированием". В R1 это 8-й бит, а в R2 и R3 — 10-й. На каждом шаге сдвигаются только те регистры у которых значение бита синхронизации равно большинству значений синхронизирующих битов всех трех регистров.

На сегодняшний день известно большое количество успешных атак на GSM шифрование и все они относятся к атакам типа known-plaintext, т.е. для восстановления ключа атакующему помимо зашифрованных фреймов необходимо знать так же незашифрованные данные, которые соответствуют этим фреймам. На первый взгляд такое требование может показаться фантастическим, однако из-за специфики стандарта GSM, в котором помимо голосового трафика передаются различные системные сообщения, такого рода атаки из разряда теоретических переходят в разряд практических.

Системные сообщения GSM содержат повторяющиеся данные и могут использоваться злоумышленником. В частности метод, предложенный Karsten Nohl в 2010 году основан как раз таки на поиске такого рода данных в шифротексте и простом переборе различных

вариантов ключей, хранящихся в радужных таблицах, до тех пор пока не будет найден ключ, порождающий нужный шифротекст для известного заранее системного сообщения.

10.2. Криптографическая защита беспроводных сетей стандартов LTE

Стандарт сетей LTE [16] – стандарт беспроводной высокоскоростной передачи данных для мобильных телефонов и других терминалов, работающих с данными. Он основан на GSM/EDGE и UMTS/HSPA сетевых технологиях, увеличивая пропускную способность и скорость за счёт использования другого радиointерфейса вместе с улучшением ядра сети.

На рисунке 5.1 представлена структура сети стандарта LTE.

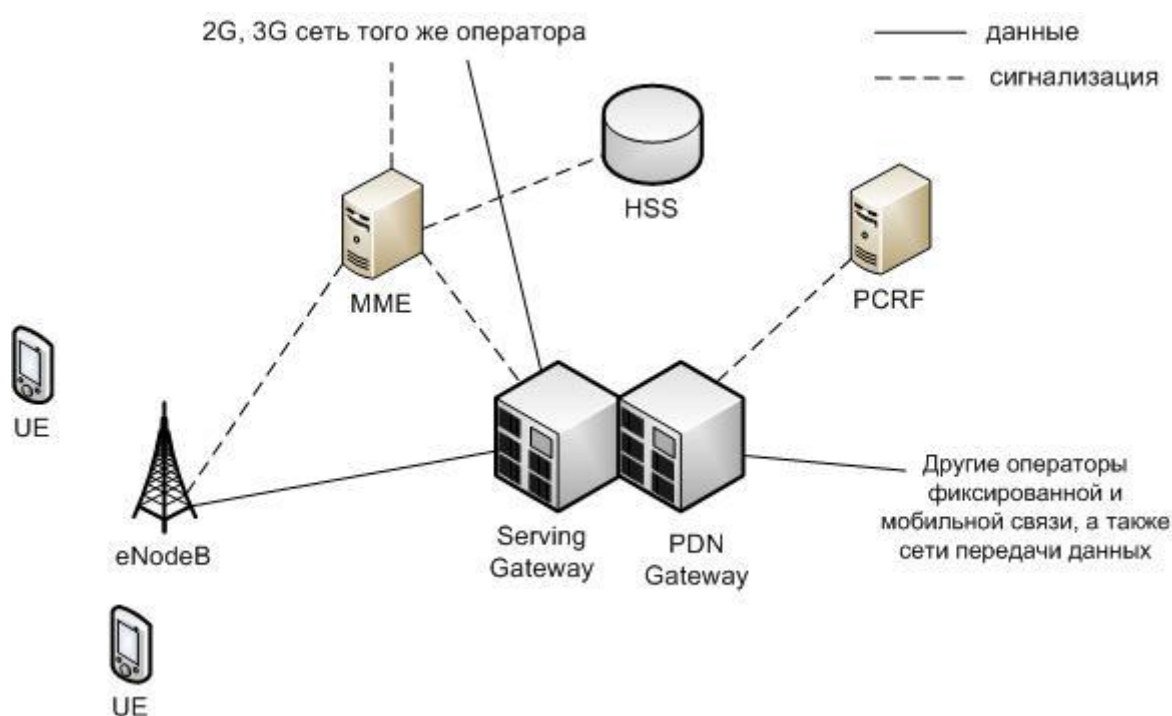


Рис. 10.1. Структура сети стандарта LTE

Из этой схемы видно, что структура сети сильно отличается от сетей стандартов 2G и 3G. Существенные изменения претерпела и подсистема базовых станций, и подсистема коммутации. Изменена технология передачи данных между оборудованием пользователя и базовой станцией. Также подверглись изменению и протоколы передачи данных между сетевыми элементами. Вся информация (голос, данные) передается в виде пакетов. Таким образом, уже нет разделения на части обрабатывающие либо только голосовую информацию, либо только пакетные данные.

Можно выделить следующие основные элементы сети стандарта LTE:

- **Serving SAE Gateway** или просто **Serving Gateway (SGW)** – обслуживающий шлюз сети LTE. Предназначен для обработки и

маршрутизации пакетных данных поступающих из/в подсистему базовых станций. SGW имеет прямое соединение с сетями второго и третьего поколений того же оператора, что упрощает передачу соединения в /из них по причинам ухудшения зоны покрытия, перегрузок и т.п. В SGW нет функции коммутации каналов для голосовых соединений, т.к. в LTE вся информация, включая голос коммутируется и передается с помощью пакетов.

- **Public Data Network SAE Gateway** или просто **PDN Gateway (PGW)** – шлюз к сетям передачи данных других операторов для сети LTE. Основная задача PGW заключается в маршрутизации трафика сети LTE к другим сетям передачи данных, таких как Интернет, а также сетям GSM, UMTS.

- **Mobility Management Entity (MME)** – узел управления мобильностью сети сотовой связи стандарта LTE. Предназначен для обработки сигнализации, преимущественно связанной с управлением мобильностью абонентов в сети.

- **Home Subscriber Server (HSS)** – сервер абонентских данных сети сотовой связи стандарта LTE. Представляет собой большую базу данных и предназначен для хранения данных об абонентах. Кроме того, HSS генерирует данные, необходимые для осуществления процедур шифрования, аутентификации и т.п. Сеть LTE может включать один или несколько HSS. Количество HSS зависит от географической структуры сети и числа абонентов.

- **Policy and Charging Rules Function (PCRF)** – элемент сети сотовой связи стандарта LTE, отвечающий за управление начислением платы за оказанные услуги связи, а также за качество соединений в соответствии с заданными конкретному абоненту характеристиками.

Для того чтобы данные могли быть транспортированы через интерфейс радио LTE, используются различные «каналы». Они используются для того, чтобы выделять различные типы данных и позволить им транспортироваться через сеть доступа более эффективно. Использование нескольких каналов обеспечивает интерфейс более высокого уровня в рамках протокола LTE и включают более чёткую и определенную сегрегацию данных.

Есть три категории, в которые могут быть сгруппированы различные каналы передачи данных:

Логические каналы – предоставляет услуги среднего уровня управления доступом MAC (*Medium Access Control*) в пределах структуры протокола LTE. Логические каналы по типу передаваемой информации делятся на логические каналы управления и логические каналы трафика. Логические каналы управления используются для передачи различных

сигнальных и информационных сообщений. По логическим каналам трафика передают пользовательские данные.

Транспортные каналы — транспортные каналы физического уровня предлагают передачу информации в MAC и выше. Информацию логических каналов после обработки на RLC/MAC уровнях размещают в транспортных каналах для дальнейшей передачи по радиоинтерфейсу в физических каналах. Транспортный канал определяет как и с какими характеристиками происходит передача информации по радиоинтерфейсу. Информационные сообщения на транспортном уровне разбивают на транспортные блоки. В каждом временном интервале передачи (Transmission Time Interval, TTI) по радиоинтерфейсу передают хотя бы один транспортный блок. При использовании технологии MIMO возможна передача до четырех блоков в одном TTI.

Физические каналы – это каналы передачи, которые переносят пользовательские данные и управляющие сообщения. Они изменяются между восходящим и нисходящим потоками, поскольку каждый из них имеет различные требования и действует по-своему.

Существующие методы и стандарты защиты беспроводных сетей LTE

Безопасность в сетях LTE заключается в нескольких видах:

- Защита абонентов.
- Защита передаваемых сообщений.
- Шифрование сообщений.
- Аутентификация абонента, и сети.

Защита абонента заключается в том, что в процессе обслуживания его скрывают временными идентификаторами.

Для закрытия данных в сетях LTE используется потоковое шифрование методом наложения на открытую информацию псевдослучайной последовательности (ПСП) с помощью оператора XOR (исключающее или). В этих сетях для обеспечения безопасности внутри сети применяется принцип туннелирования соединений. Шифрации можно подвергать пакеты S1 и X2 при помощи IPsec ESP, а также подвергаются шифрации сигнальные сообщения этих интерфейсов.

В момент подключения или активизации абонентского оборудования (UE) в сети, сеть запускает процедуру аутентификации и соглашения о ключах АКА (Authentication and Key Agreement). Целью этой процедуры является взаимная аутентификация абонента и сети и выработка промежуточного ключа K_{ASME} . Работа механизма АКА занимает доли секунды, которые необходимы для выработки ключа в приложении USIM и для установления соединения с Центром регистрации (HSS). Вследствие этого, для достижения скорости передачи данных сетей LTE необходимо добавить функцию обновления ключевой

информации без инициализации механизма АКА. Для решения этой проблемы в сетях LTE предлагается использовать иерархическую ключевую инфраструктуру. Здесь также, как и в сетях 3G, приложение USIM и Центр аутентификации (AuC) осуществляет предварительное распределение ключей. Когда механизм АКА инициализируется для осуществления двусторонней аутентификации пользователя и сети, генерируются ключ шифрования СК и ключ общей защиты, которые затем передаются из ПО USIM в Мобильное оборудование (ME) и из Центра аутентификации в Центр регистрации (HSS). ME и HSS, используя ключевую пару (СК;ИК) и ID используемой сети, вырабатывает ключ K_{ASME} . Установив зависимость ключа от ID сети, Центр регистрации гарантирует возможность использования ключа только в рамках этой сети. Далее K_{ASME} передается из Центра регистрации в устройство мобильного управления (MME) текущей сети, где он используется в качестве мастер-ключа. На основании K_{ASME} вырабатывается ключ $K_{nas-enc}$, который необходим для шифрования данных протокола NAS между мобильным устройством (UE) и MME, и $K_{nas-int}$, необходимый для защиты целостности. Когда UE подключается к сети, MME генерирует ключ $KeNB$ и передает его базовым станциям. В свою очередь, из ключа $KeNB$ вырабатывается ключ K_{up-enc} , используемый для шифрования пользовательских данных протокола U-Plane, ключ $K_{rrc-enc}$ для протокола RRC (Radio Resource Control - протокол взаимодействия между Мобильными устройствами и базовыми станциями) и ключ $K_{rrc-int}$, предназначенный для защиты целостности.

Алгоритм аутентификации и генерации ключа представлен на рис 10.2

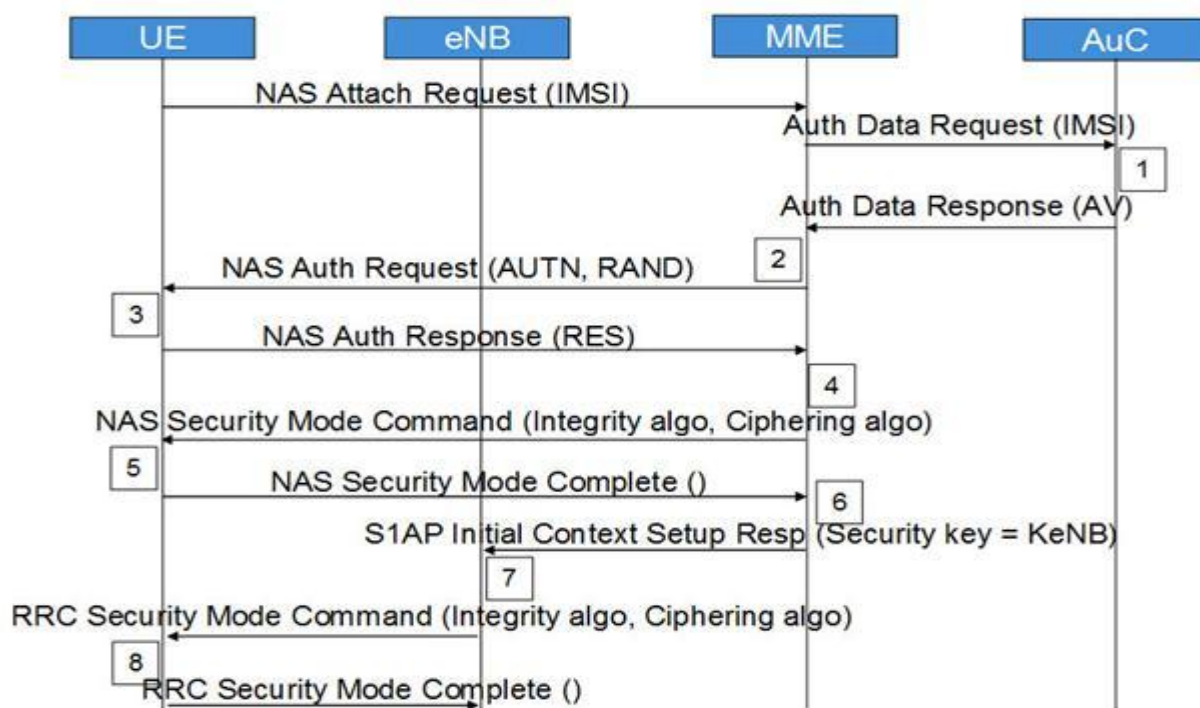


Рис. 10.2. Алгоритм аутентификации и генерации ключа

Здесь:

Шаг 1. Запрос о подключении к сети от мобильной станции (UE). MME запрашивает аутентификационные данные, относящиеся к конкретному IMSI, отправляя Authentication Data Request. AuC/HSS выбирает PSK, относящийся к конкретному IMSI и вычисляет аутентификационные данные по PSK. AuC/HSS отправляет обратно AV с Authentication Data Response.

Шаг 2. MME получает IK, CK, XRES, RAND и AUTH из AV. MME отправляет AUTH и RAND при помощи Authentication Request к UE.

Шаг 3. UE аутентифицирует NW, проверяя полученный AUTH. После чего вычисляет IK, CK, RES, XMAC из своего ключа защиты, AMF, (OP), AUTH и RAND. Она отправляет RES с Authentication response.

Шаг 4. После получения RES, MME сравнивает его с XRES и если они совпадают, то аутентификация прошла успешно, в противном случае, MME отправляет сбой аутентификации (Authentication failure) к UE. MME сбрасывает счетчик DL NAS. Рассчитывает KASME, KeNB, Knas-int, Knas-enc. Отправляет NAS команду режима безопасности (алгоритм целостности, алгоритм шифрования, NAS набор ключей ID, функцию безопасности UE) с целостностью охраняемых, но не зашифрованных, используя Knas-enc.

Шаг 5. После получения NAS команды режима безопасности, UE вычисляет KASME, KeNB, Knas-int, Knas-enc. UE отправляет NAS режима безопасности выполнен с целостностью, защищенных и зашифрованных.

Шаг 6. После получения NAS команды режима безопасности от UE, MME отправляет KeNB в eNB с S1AP первоначальная установка начального контекста (ключ защиты).

Шаг 7. После получения KeNB, eNB вычисляет Krrc-int, Krrc-enc, Krrc-enc. Затем оно отправляет RRC ключ защиты команду с AS целостностью алгоритма и AS шифрующий алгоритм.

Шаг 8. После получения RRC команды ключа защиты UE вычисляет Krrc-int, Krrc-enc, Krrc-enc. UE отправляет RRC выполненный ключ шифрования на eNB.

После всех описанных действий, все NAS и AS сообщения будут надежно защищены и зашифрованы, в отличие от пользовательских данных, которые будут только шифроваться.

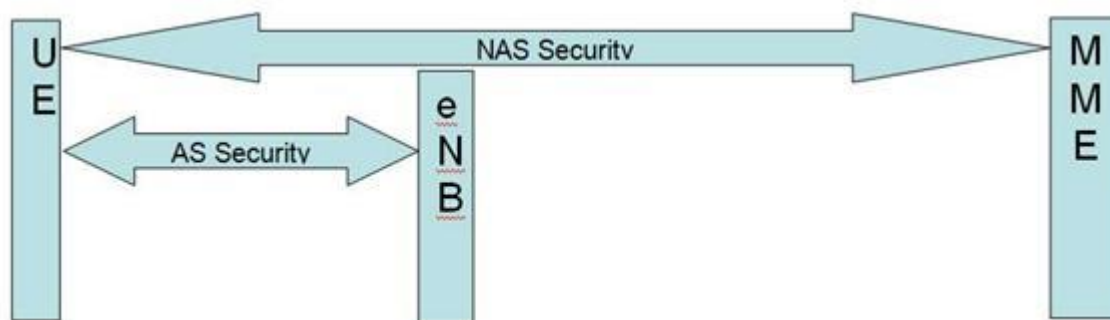


Рис. 10.3. Слои безопасности

Архитектура безопасности LTE определяет механизм безопасности и для уровня NAS и для уровня AS.

Безопасность NAS (слоя без доступа): Выполнена для NAS сообщений и принадлежит области UE и MME.

В этом случае необходима при передаче сообщений NAS между UE и MME – целостность, защищенная и зашифрованная с дополнительным заголовком безопасности NAS.

Безопасность AS (слоя с доступом): Выполнена для RRC и плоскости пользовательских данных, принадлежащих области UE и eNB. Уровень PDCP на сторонах UE и eNB отвечает за шифрование и защиту целостности.

RRC сообщения защищены целостностью и зашифрованы, однако данные U-Plane только зашифрованы.

Для генерации векторов аутентификации используется криптографический алгоритм с помощью однонаправленных функций (f_1, f_2, f_3, f_4, f_5) когда прямой результат получается путем простых вычислений, а обратный результат не может быть получен обратным путем, то есть не существует эффективного алгоритма получения обратного результата. Для этого алгоритма используется случайное 128 битное случайное число RAND, мастер-ключ K абонента, также 128 бит и порядковый номер процедуры SQN (Sequence Number). Счетчик SQN меняет свое значение при каждой генерации вектора аутентификации. Похожий счетчик SQN работает и в USIM. Такой метод позволяет генерировать каждый раз новый вектор аутентификации, не повторяя предыдущий уже использованный вектор аутентификации.

Помимо этих трех исходных величин: SQN, RAND и K в алгоритме f_1 участвует поле управления аутентификацией Authentication Management Field (AMF), а в алгоритмах $f_2 - f_5$ исходные параметры – RAND и K, что и продемонстрировано на рис. 2.3, 2.4. На выходах соответствующих функций получают Message Authentication Code (MAC) - 64 бита; XRES – eXpected Response, результат работы алгоритма аутентификации <32 – 128 бит>; ключ

шифрации СК, генерируемый с использованием входящих (K,RAND)->f3->СК; ключ целостности IK, сгенерированный с использованием входящего (K,RAND)->f4->ИК; и промежуточный ключ Anonymity Key (AK), генерируемый с помощью (K,RAND)->f5->AK - 64 бита.

При обслуживании абонента сетью E-UTRAN ключи СК и ИК в открытом виде в ядро сети не передают. В этом случае HSS генерирует K_{ASME} с помощью алгоритма KDF (Key Derivation Function), для которого исходными параметрами являются СК и ИК, а также идентификатор обслуживающей сети и SQN Δ AK. Вектор аутентификации содержит RAND, XRES, AUTN и K_{ASME} , на основе которого происходит генерация ключей шифрации и целостности, используемых в соответствующих алгоритмах.

Когда мобильная станция получает из ядра сети три параметра (RAND, AUTN и KSI_{ASME} , где KSI – Key Set Identifier, индикатор установленного ключа, однозначно связанный с K_{ASME} в мобильной станции).

После чего используя RAND и AUTN, USIM на основе алгоритмов безопасности, тождественных хранящимся в HSS, производит вычисление XMAC, RES, СК и ИК.

Затем в ответе RES UE передает в MME вычисленное RES, которое должно совпасть с XRES, полученным из HSS. Так сеть аутентифицирует абонента. Вычислив XMAC, UE сравнивает его с MAC, полученным ею в AUTN. При успешной аутентификации абонентом сети (MAC = XMAC) UE сообщает об этом в ответе RES. Если аутентификация сети не удалась (MAC \neq XMAC), то UE направляет в MME ответ CAUSE, где указывает причину неудачи аутентификации.

При успешном завершении предыдущего этапа MME, eNB и UE производят генерацию ключей, используемых для шифрации и проверки целостности получаемых сообщений. В E-UTRAN имеется иерархия ключей, которая приведена на рис. 2.5.

Векторы аутентификации (рис. 2.3, 2.4):

Ключи ИК и СК генерируются и в центре аутентификации, и в USIM;

Ключ АК генерируется только в центре аутентификации;

Ответ XRES генерируется только в центре аутентификации, а RES генерируется в USIM;

Код MAC генерируется только в центре аутентификации, а соответствующий ему параметр XMAC генерируется в USIM;

Маркер AUTH генерируется только в центре аутентификации.

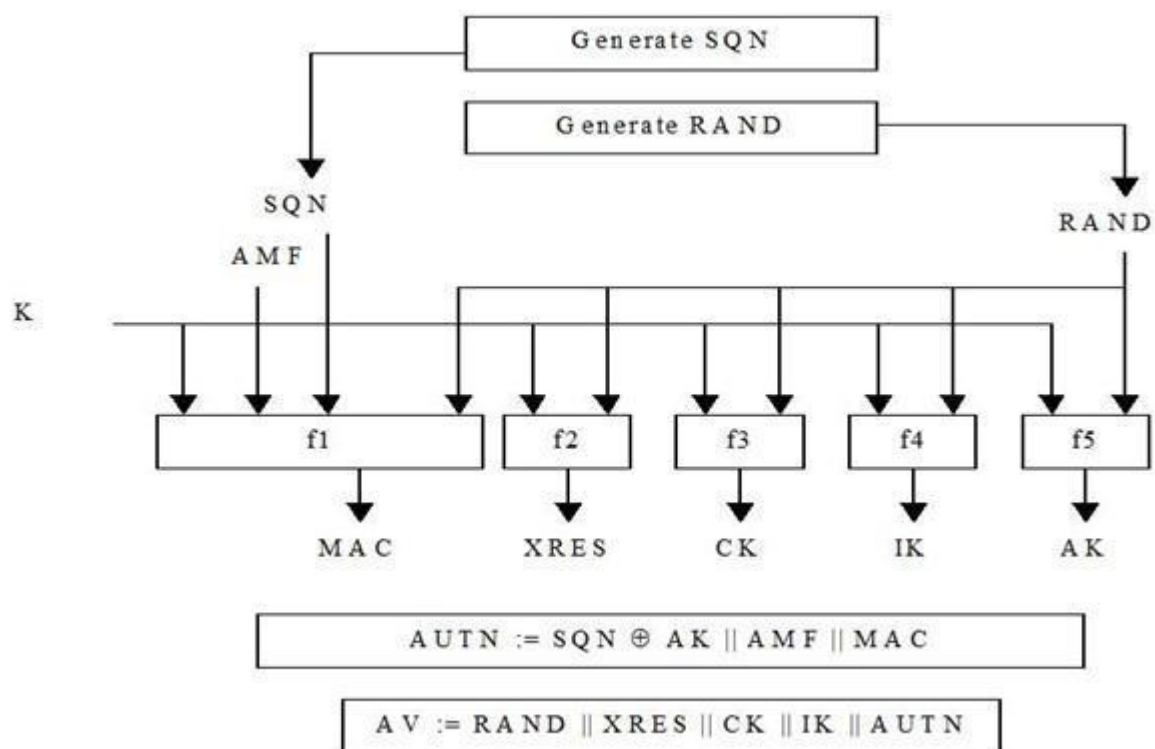


Рис. 10.4. Создание векторов на передающей стороне

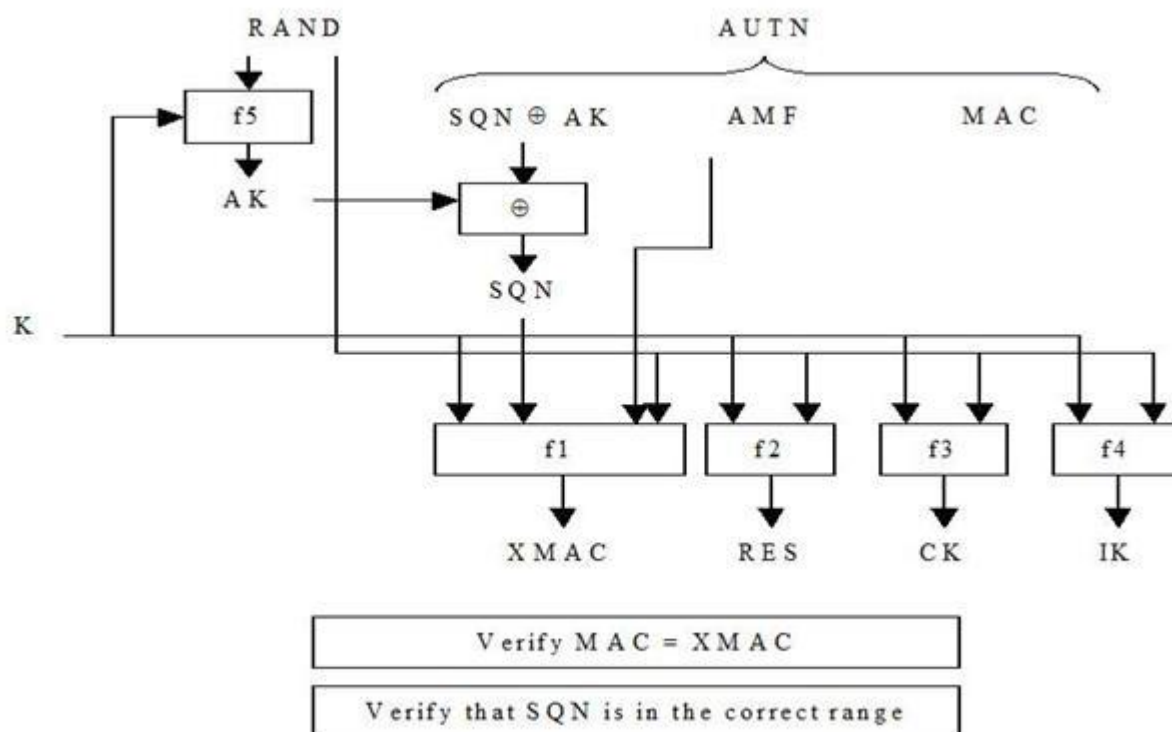


Рис. 10.5. Преобразование векторов на принимаемой стороне

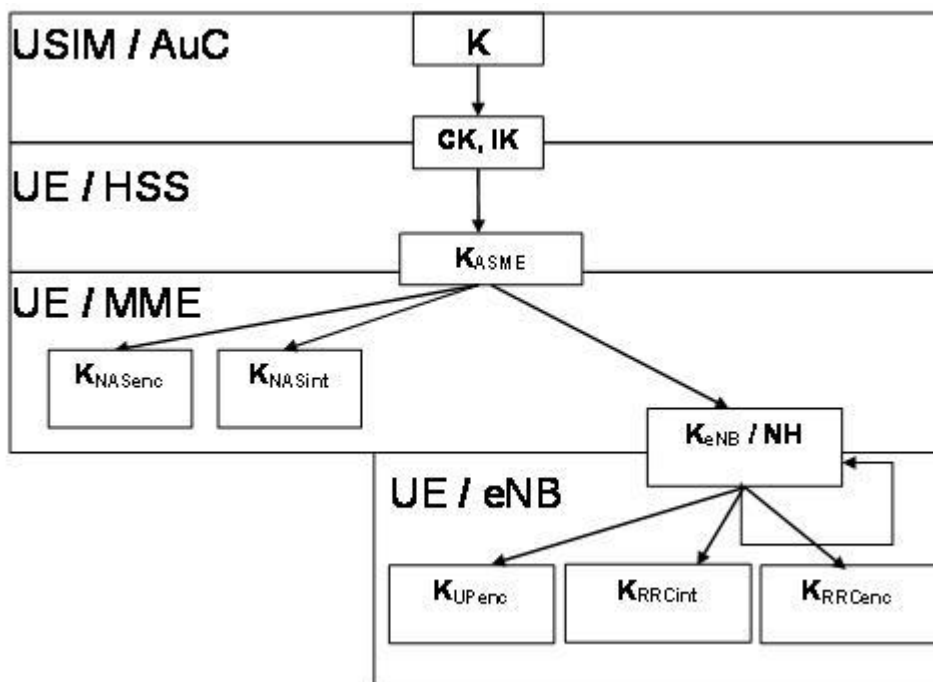


Рис. 10.6. Иерархия ключей в E-UTRAN

Исходным ключом для всей цепочки является K_{ASME} (256 бит). При передаче в радиоканале защиту обеспечивают для сигнального трафика (Control Plane) и для пользовательских пакетов (User Plane). При этом все сообщения сигнализации разделяют на сквозные сигнальные сообщения между UE и MME протоколов MM и SM (NAS – Non Access Stratum) и сигнальные сообщения между eNB протокола RRC (AS – Access Stratum). Для шифрации и защиты целостности можно использовать разные базовые алгоритмы:

- UEA2 (UMTS Encryption Algorithm 2) и UIA2 (UMTS Integrity Algorithm 2);
- разработанные для стандартов 3G, AES (Advanced Encryption Standard).

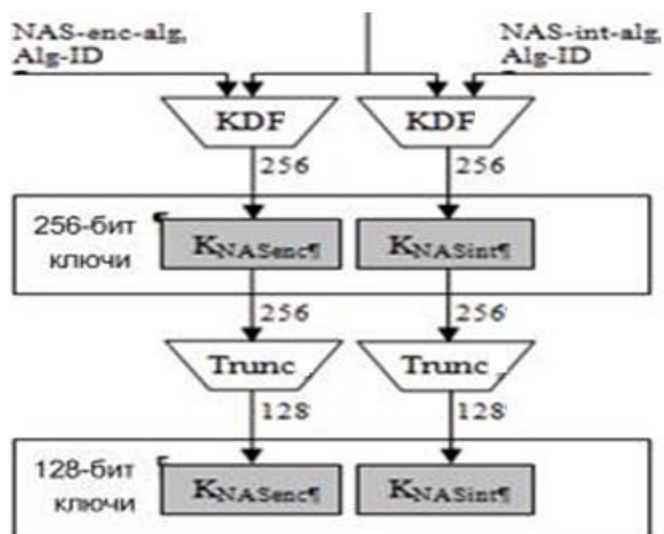


Рис. 10.7. Генерирование ключей шифрации и целостности для NAS сигнализации

Сигнальные сообщения протокола RRC (AS) также шифруют и обеспечивают их целостность. Пакеты трафика только шифруют. Эти операции производят в обслуживающей eNB и UE. Схема получения ключей шифрации и целостности (рис. 7) для AS и UP трафика отличается от предыдущего случая тем, что исходным параметром здесь служит вторичный промежуточный ключ KeNB (256 бит). Этот ключ генерируют, также используя KDF, где входными параметрами являются: KASME, счетчик сигнальных сообщений NAS вверх, прежнее значение KeNB, идентификатор соты и номер частотного канала в направлении вверх. Следовательно, при каждой периодической локализации UE происходит изменение KeNB.

Также KeNB меняется и при хэндове; при этом в алгоритме генерации нового KeNB можно использовать дополнительный параметр NH (Next Hop), фактически счетчик числа базовых станций, по цепочке обслуживающих абонента. Все реализуемые процедуры безопасности в сети E-UTRAN продемонстрированы на рис. 2.8.

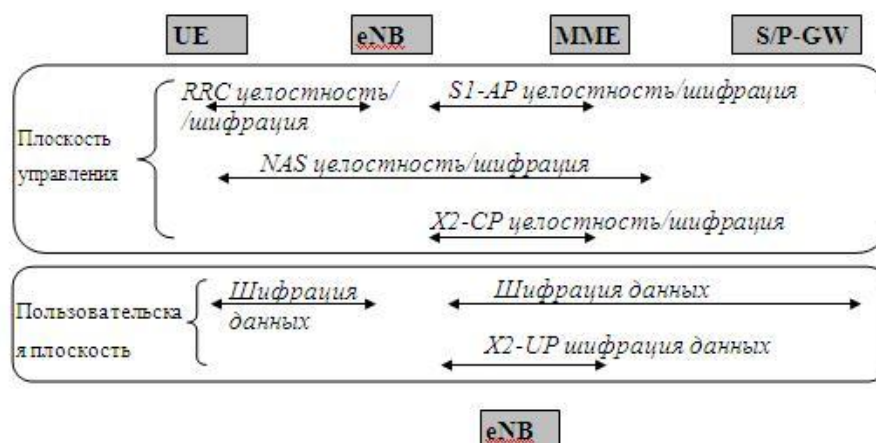


Рис. 10.8. Реализуемые процедуры безопасности в сети E-UTRAN

Алгоритм шифрации и дешифрации сообщений представлен на рис. 2.9.

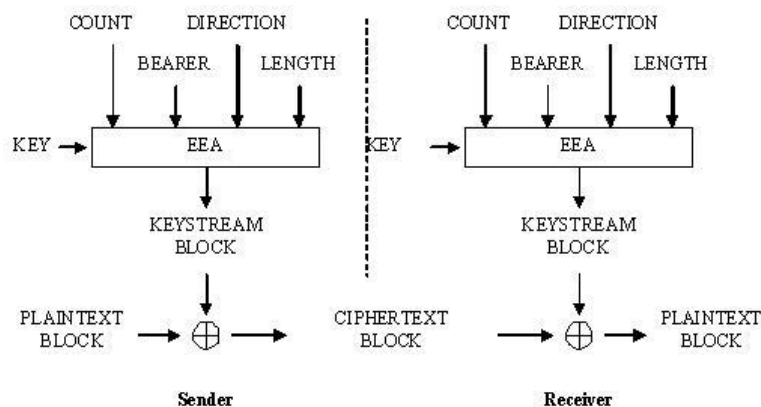


Рис. 10.9. Алгоритм шифрации в E-UTRAN

Исходными параметрами в этом алгоритме являются шифрующий ключ KEY (128 бит), счетчик пакетов (блоков) COUNT (32 бита), идентификатор сквозного канала BEARER (5 бит), указатель направления передачи DIRECTION (1 бит) и длина шифрующего ключа LENGTH. В соответствии с выбранным алгоритмом шифрации EEA (EPS Encryption Algorithm) вырабатывается шифрующее число KEYSTREAM BLOCK, которое при передаче складывают по модулю два с шифруемым исходным текстом блока PLAINTEXT BLOCK. При дешифрации на приемном конце повторно совершают эту же операцию.

Процедура защиты целостности сообщения состоит в генерации “хвоста“ MAC (Message Authentication Code) (32 бита), присоединяемого к передаваемому пакету. Алгоритм генерации MAC и проверки целостности полученного пакета путем сравнения XMAC с MAC (они должны совпасть) отображен на рис. 5.10.

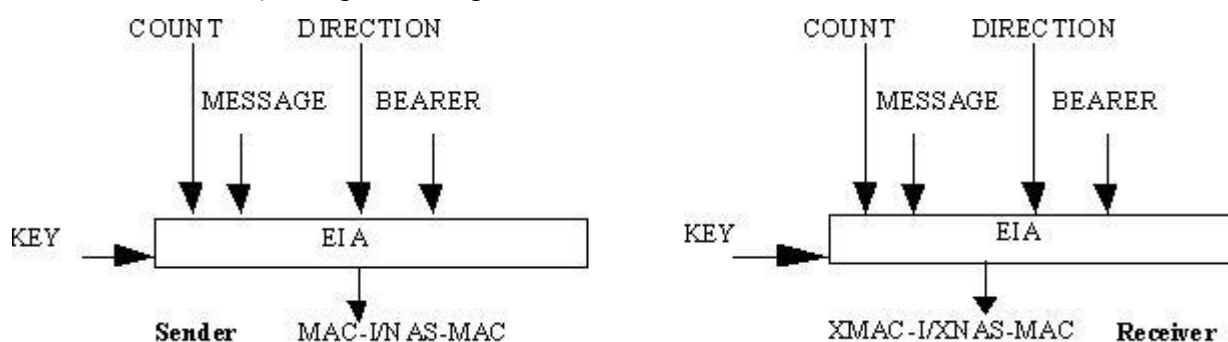


Рис. 10.10. Алгоритм проверки целостности E-UTRAN

В алгоритме EIA (EPS Integrity Algorithm) использован ключ целостности KEY (128 бит), счетчик сообщений COUNT (32 бита), идентификатор сквозного канала BEARER (5 бит), указатель направления передачи DIRECTION (1 бит) и само сообщение MESSAGE.

Моделирование технологии LTE в среде MATLAB с использованием встроенного пакета LTE System Toolbox

LTE System Toolbox™ предоставляет соответствующие стандарту функции и приложения для проектирования, моделирования и проверки коммуникационных систем стандартов LTE и LTE-Advanced. Данный инструмент ускоряет разработку LTE-алгоритмов и физического уровня (PHY), предоставляет эталонный образец для проверки и тестирования на соответствие стандарту, а также позволяет генерировать тестовые сигналы. С помощью LTE System Toolbox можно производить настройку, моделирование, измерения и анализ канала связи. Также можно создавать и повторно использовать сценарии тестов для подтверждения того, что проекты, прототипы и разработки соответствуют стандарту LTE.

Основные особенности:

- Модели, соответствующие стандартам LTE и LTE-Advanced (Release 8, 9, 10 и 11).
- Функции обработки на канальном уровне, поддержка от 1 до 10 режимов передачи по нисходящему каналу, образцы проектов, в том числе CoMP.
- Тестовые модели (E-TM), эталонный измерительный канал (RMC) для LTE, LTE-A, а также генератор UMTS-сигналов.
- Интерактивные инструменты для проверки на соответствие стандарту и BER тестов.
- Передача и приём сигналов при помощи радиоустройств для тестирования систем в реальном эфире.
- Выделение системных и контрольных параметров из принятого сигнала, в том числе cell ID, MIB и SIB1.
- Оценка канала связи.

Сквозное моделирование LTE

System Toolbox даёт возможность моделировать и имитировать физический уровень стандарта LTE. Моделирование системы на канальном уровне позволяет добиться требуемых значений характеристик системы, в том числе пропускной способности и BER, а также определить конкретные реализации системы на основе производимых измерений.

LTE System Toolbox также позволяет улучшить планирование системы, облегчая моделирование канального уровня, которое предоставляет некоторые параметры, необходимые для проектирования базовых станций с заданной геометрией и характеристиками распространения сигнала.

Набор поддерживаемых функций для моделирования режимов передачи и приёма, а также канала связи, включает в себя:

- режимы FDD и TDD на несущих частотах;
- все полосы передачи LTE-сигналов от 1,4 до 20 МГц, LTE-A до 100 МГц с агрегацией несущей;
- различные типы LTE-сигналов, включая нисходящие и восходящие опорные сигналы и сигналы синхронизации;
- физические LTE-каналы, в том числе каналы управления и каналы общего доступа;
- готовую процедуру обработки нисходящего канала, в том числе формирование нисходящего общего канала и канала управления, все возможные MIMO-режимы и генерацию OFDM-сигналов;
- готовую процедуру обработки восходящего канала, в том числе формирование восходящих общего канала и канала управления, SU-MIMO и MU-MIMO режимы и генерацию SC-FDMA сигналов;

- адаптацию к каналу связи, включая схемы выбора типов модуляции и кодирования (MCS) в соответствии с оценкой качества канала связи (CQI), индикатора ранга (RI) и индикации матрицы прекодера (PMI);
- возможности и примеры построения LTE-Advanced, в том числе приём и передача с несколькими eNB (CoMP) и с агрегацией несущей;
- модели распространения LTE-сигналов, в том числе модель пешехода (EPA), модель автомобиля (EVA), модель типичной городской застройки (ETU), модель распространения в движении, а также модели MIMO-каналов в скоростном поезде. LTE System Toolbox даёт возможность создавать тесты, измеряющие пропускную способность PDSCH-канала в соответствии с указанными в стандарте LTE (TS 36.101) условиями испытаний. Структуры данных в LTE System Toolbox позволяют удобно отображать все параметры системы. Функции данного инструмента отражают любые возможные комбинации режимов работы передатчиков, моделей каналов и приемников. Используя этот инструмент для тестирования на соответствие стандарту и BLER-тестирования, вы можете измерять характеристики системы и сравнивать их с указанными в спецификации к стандарту

На рисунке 5.11 изображена структурная схема программного комплекса.

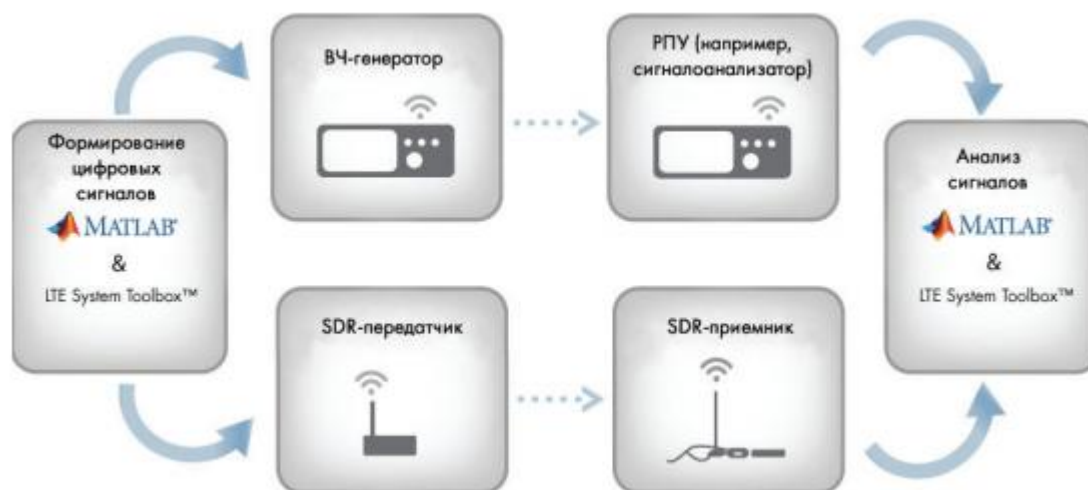


Рис. 10.12. Структурная схема программного комплекса

Для генерации тестового сигнала от базовой станции к абоненту используется генератор LTE-Downlink E-TM Generator.

LTE Downlink RMC Generator

Generate PDSCH reference measurement channel (RMC) waveforms. These are specified in TS36.101 Annex A.3 for UE performance testing.

Reference channel	R.12	RMC parameter summary
Duplex mode	FDD	
Transmission scheme	TxDiversity	Number of downlink resource blocks 6
Cell identity	0	Transmit antenna ports 4
RNTI	1	Modulation QPSK
RV sequence	[0 1 2 3]	Transmission layers 4
Rho (dB)	0	Total info bits per frame per codeword 3416 bits
OCNG	Off	Codeword input data
Number of subframes	10	
Number of codewords	1	Transport info bit stream (codeword 1) User defined
PMI set	[1]	[1; 0; 0; 1]
Number of HARQ processes	8	Transport info bit stream (codeword 2) User defined
Windowing (samples)	0	[1; 0; 0; 1]
Waveform output variable	rmcwaveform	
Resource grid output variable	rmcgrid	
RMC configuration output variable	rmcconfig	

Generate waveform

Рис. 10.13. Окно генератора от базовой станции к абоненту

В данном окне задаются параметры генерируемого сигнала на базовой станции, такие как: Количество каналов, вид модуляции, количество обслуживаемых абонентов в секунду, число кодовых слов, вектор инициализации.

На рисунке 5.14 показан вид сгенерированного сигнала, а на рисунке 5.15 трех мерный спектр полученного сигнала.

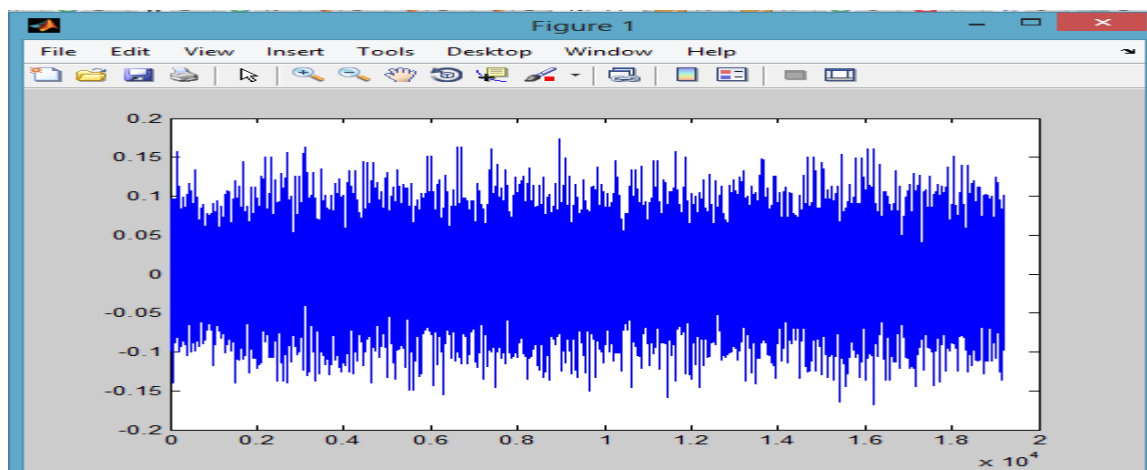


Рисунок 10.14. Сгенерированный сигнал станцией

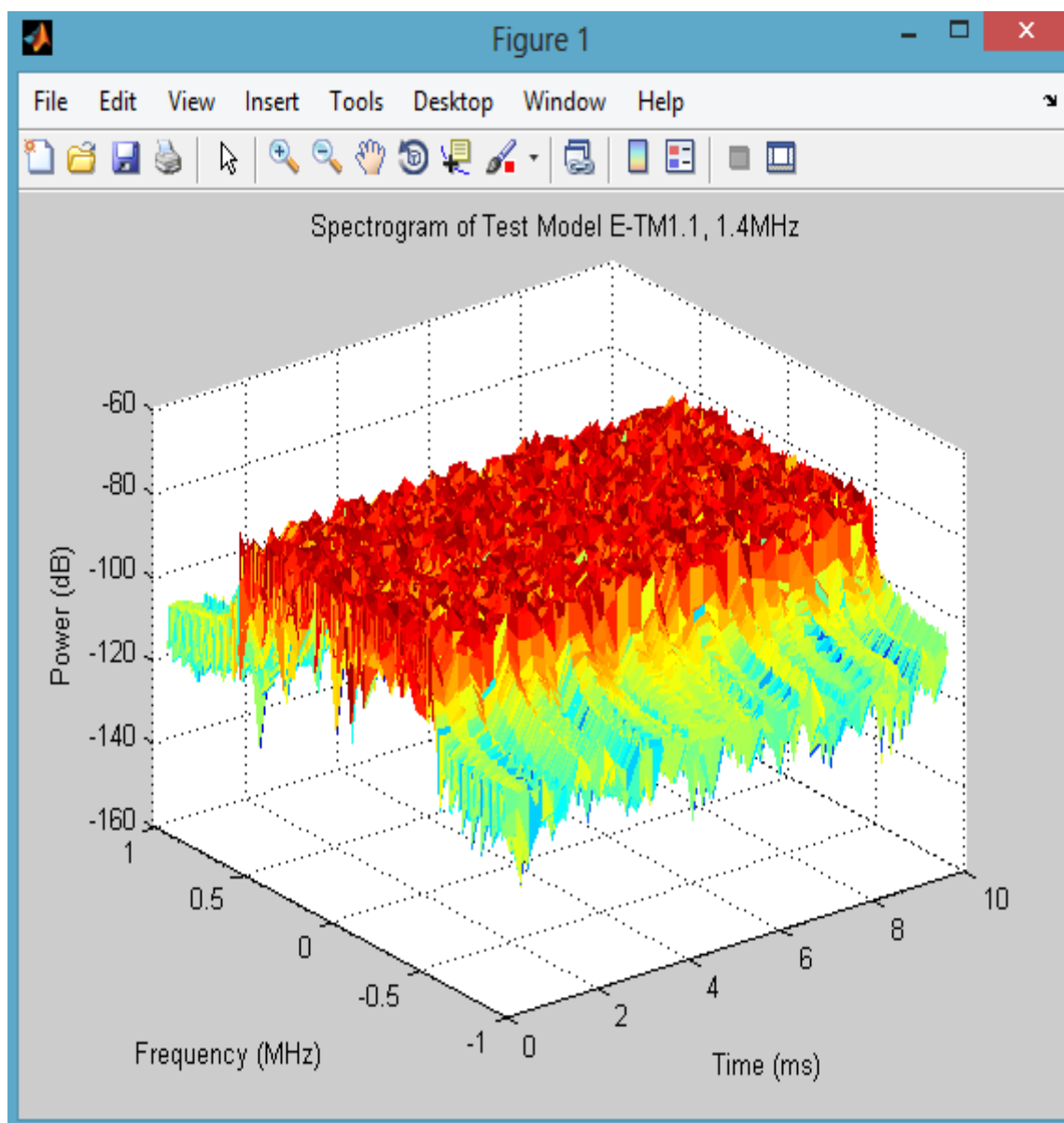


Рис. 10.15. Спектр сгенерированного сигнала станцией в течении 10 секунд

Для генерации тестового сигнала от абонента к базовой станции используется генератор LTE-Uplink RMS Generator.

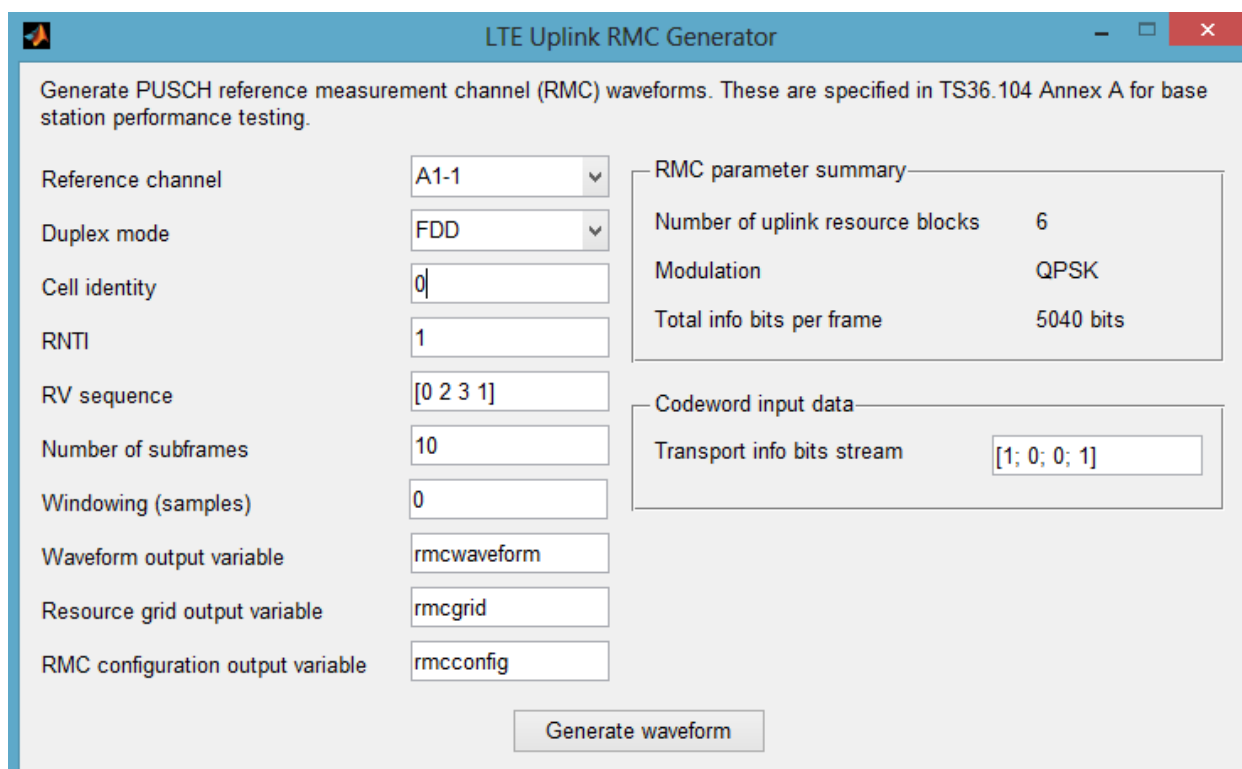


Рис. 10.16. Окно генератора LTE-Uplink RMS Generator

В окне на рисунке 10.16 производятся настройки параметров генерирования от абонента к базовой станции, такие как: Количество каналов для передачи потока пакетов, тип модуляции, вектор инициализации, количество передаваемых пакетов в секунду.

Сигнал, генерируемый генератором LTE-Uplink RMS Generator, представлен на рисунке 10.17.

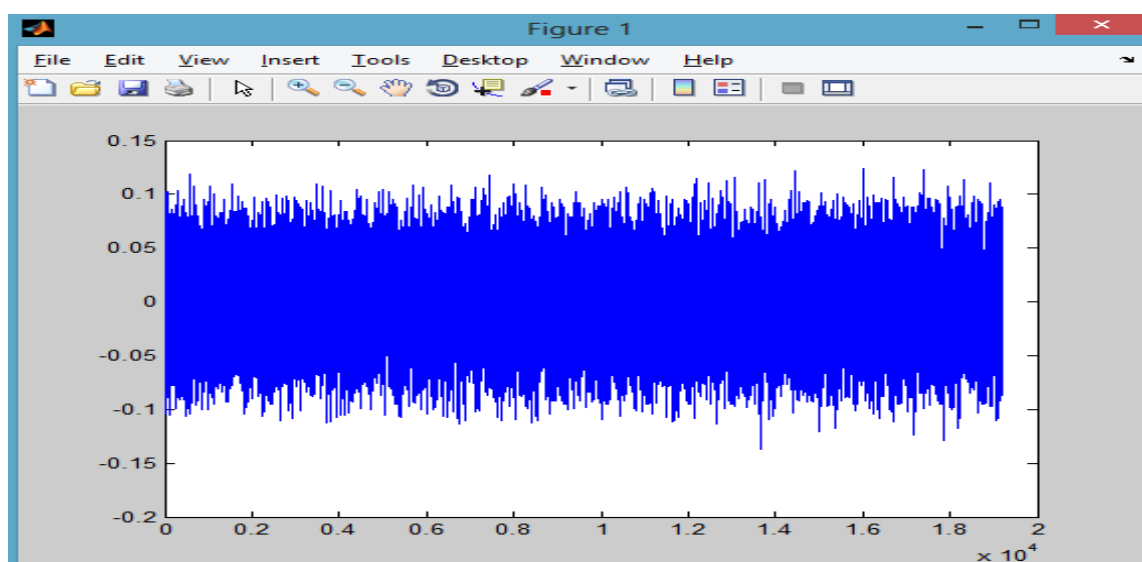


Рис. 10.17. Сигнал генерируемый генератором LTE-Uplink RMS Generator

Для вычисления потерь и пропускной способности системы передачи можно использовать блок LTE PDSCH Conformance Testing.

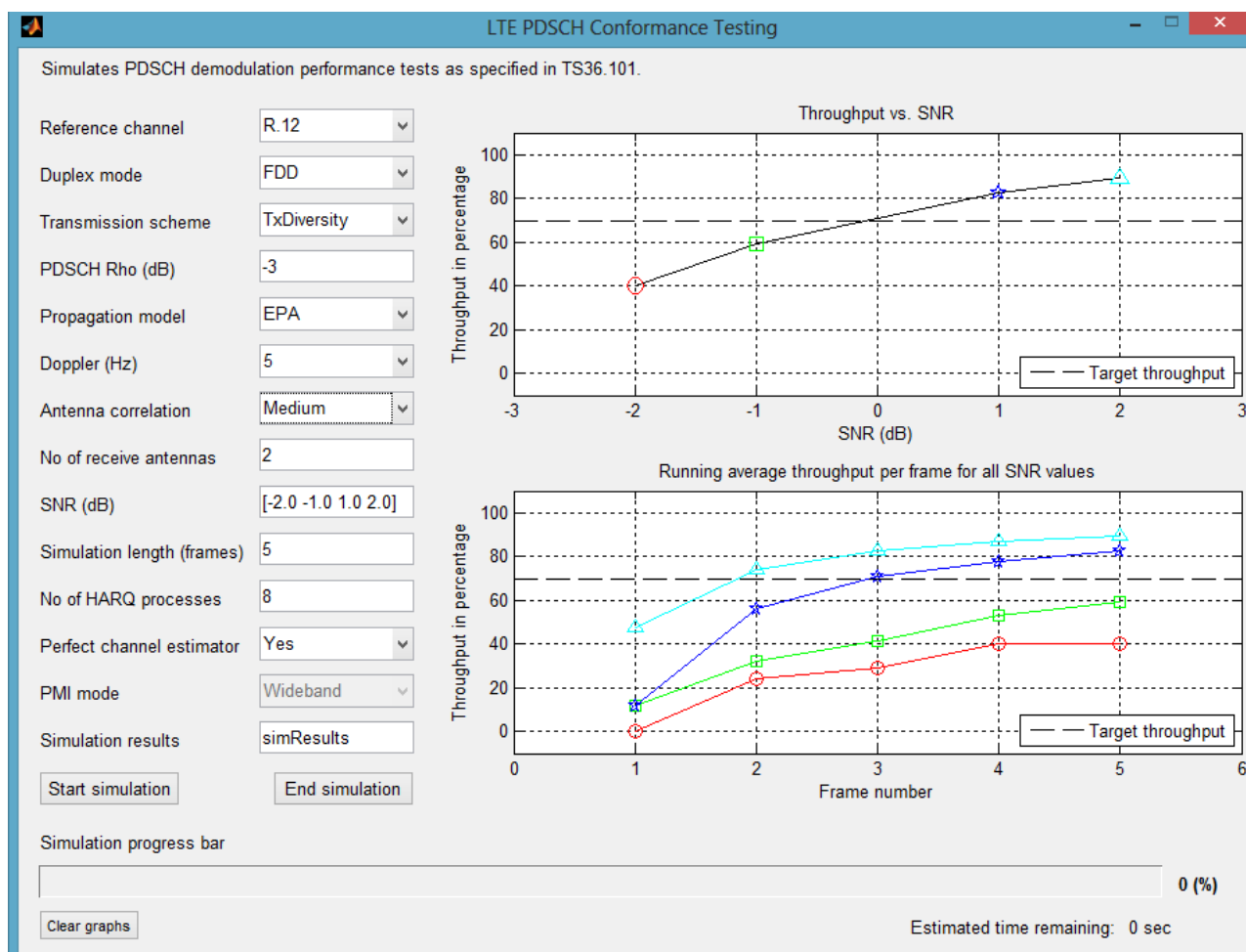


Рис. 10.18. Окно LTE PDSCH Conformance Testing

В данном окне можно произвести настройку параметров линии передачи, таких как: количество каналов, тип модуляции, доплеровскую частоту, уровень шума, настройки антенны и другие. А так же наглядно наблюдать изменение количества потерь в линии передачи и пропускную способность системы.

В результате проделанной работы были изучены основные понятия беспроводных сетей LTE, физическая структура построения беспроводной сети LTE. Изучены существующие методы и средства защиты беспроводных сетей LTE. Построена программная структурная схема беспроводной сети LTE среде Matlab с использованием встроенного пакета LTE System Toolbox и проведены исследования основных узлов сети LTE.

ЗАКЛЮЧЕНИЕ

В представленном учебном пособии впервые рассмотрены основные виды модуляции, кодирования и сигнально-кодовые конструкции. В учебном пособии рассмотрена теория и практика моделирования модемов сотовой связи FSK, MSK GMSK и модемы спутниковых систем связи M-QAM, M-PSK. Для анализа использовано программное обеспечение NI LabVIEW. Показана возможность создания созвездий различных видов модуляции, спектры сигналов с использованием различных фильтров, получены глазковые диаграммы и фазовые зависимости от отношения сигнал/шум. Для каждого вида модуляции проведен численный анализ вероятности символьной ошибки от отношения сигнал/шум.

Рассмотрены вопросы пропускной способности канала связи, рассмотрены методы кодирования источника (Шеннона-Фано, алгоритмы сжатия Лемпеля - Зива, вейвлет-фрактальное сжатие). Проведены численные оценки коэффициентов сжатия, полученные с использованием современного программного обеспечения.

Проведено имитационное моделирование и исследование помехоустойчивых кодов Хемминга, БЧХ (Боуза-Чоудхури-Хоквенгема), Рида-Соломона на базе MATLAB 2015 Simulink, исследовано сверточные кодирование и декодирование. Проведено моделирование декодирования сверточных кодов по методу Витерби с использованием ПО MATLAB. Исследование турбокодирования, рассмотрена обобщенная схема турбокодера с параллельным каскадированием, сверточные турбокоды, декодирование турбокодов. Характеристики помехоустойчивости сверточных турбокодов исследованы с использованием ПО MATLAB 2015 Simulink. Впервые использовано ПО NI LabVIEW для исследования низкоплотностных кодов - разработан программный комплекс для визуализации и исследования LDPC-кодов. Проведена оценка помехоустойчивости LDPC-кодов в зависимости от отношения сигнал/шум и параметров кода.

Рассмотрены сигнально-кодовые конструкции на основе Треллис кодовой модуляции (TCM) и их анализ с использованием MATLAB. Исследование сигнально-кодовой конструкции на базе системы с ортогональным частотным мультиплексированием и пространственно-временным кодированием OFDM - MIMO с использованием NI LabVIEW.

В главе 5. представлено задание на самостоятельную работу "Оптимизация методов помехоустойчивого кодирования для телекоммуникационных систем" - 32 варианта для самостоятельного расчета и примеры расчета для нескольких вариантов.

Рассмотрены классические шифры, теория классических шифров, компьютерный практикум для классических шифров и задания на самостоятельную работу по классическим шифрам. Представлены современные шифры с секретным ключом, теория шифров с секретным ключом, компьютерный практикум для шифров с секретным ключом и задания на самостоятельную работу по шифрам с секретным ключом. Рассмотрены отечественные и зарубежные шифры с открытым ключом, теория шифров с открытым ключом, компьютерный практикум для шифров с открытым ключом и задания на самостоятельную работу по шифрам с открытым ключом. Представлены криптографические протоколы в сетях передачи данных и компьютерный практикум для исследования протоколов SSL и TLS. Рассмотрено шифрование в современных системах связи стандартов GSM и LTE и компьютерный практикум для исследования стандарта LTE в MATLAB.

ЛИТЕРАТУРА

1. Банкет В.Л. Помехоустойчивое кодирование в телекоммуникационных системах: учебн. пособие. - Одесса: ОНАС им А.С. Попова, 2011. - 104 с.
2. Зюко А.Г., Фалько А.И., Панфилов И.П., Банкет В.Л., Иващенко П.В. Помехоустойчивость и эффективность систем передачи информации. М.: Радио и связь. 1985.
3. Методы повышения энергетической и спектральной эффективности цифровой радиосвязи: учеб. пособие / В. А. Варгаузин, И. А. Цикин. — СПб.: БХВ-Петербург, 2013. — 352 с.
4. Банкет В.Л. Сигнально-кодовые конструкции в телекоммуникационных системах. - Одесса: Фешкс, 2009. - 180 с.
5. Мелихов С.В. Аналоговое и цифровое радиовещание: Учебное пособие. Издание второе, исправленное. - Томск: Томск. гос. ун-т систем управления и радиоэлектроники, 2012. – 233 с.
6. Голиков А.М., Уваровский В.Д. Исследование многоуровневых методов модуляции сигналов, используемых в космических системах связи, на базе аппаратуры и ПО labVIEW 2010. Методические указания по лабораторным работам – Томск: Том. гос. ун-т систем управления и радиоэлектроники, 2011. – 50 с.
7. Галкин В.А. Цифровая мобильная радиосвязь. Учебное пособие для вузов. – М.: Горячая линия-Телеком, 2007. – 432 с..
8. Федосов В. П., Нестеренко А. К. Цифровая обработка сигналов в LabVIEW: учеб. пособие / под ред. В. П. Федосова. – М.: ДМК Пресс, 2007. – 456 с.
9. Теория и техника передачи информации : учебное пособие /Ю. П. Акулиничев, А. С. Бернгардт. — Томск: Эль Контент, 2012. — 210 с.

10. Скляр Б. Цифровая связь. — М.: Издательский дом Вильямс. 2003 — 1104с
11. Феер К.: Беспроводная цифровая связь. М.: Радио и связь, 2000. - 520 с.
12. Крейнделин В.Б., Колесников А.В. Оценивание параметров канала в системах связи с ортогональным частотным мультиплексированием. Учебное пособие / МТУСИ.-М., 2010. -29 с.
13. Д. Ватолин, М. Смирнов «Методы сжатия данных: Сжатие изображений»
// http://www.compression.ru/book/part2/part2__3.htm
14. С. Уэлстид. “Фракталы и вейвлеты для сжатия изображений в действии”. Москва. “Издательство ТРИУМФ” 2003. 360 .
15. <https://sites.google.com/site/szatieinformacii/lekcii/tema13>
16. Дворкович В.П., Дворкович А.В. Цифровые видеоинформационные системы (теория и практика) Москва: техносфера, 2012. – 1008 с.
17. LabVIEW. Справочник по функциям. [Электронный ресурс]. – Режим доступа:http://chaos.sgu.ru/library/programms/progr/labVIEW/LabVIEW_suranov.pdf
18. Майков, Д.Ю. Оценка сдвига частоты для процедуры Initial Ranging в системе «мобильный WiMax» / Д.Ю. Майков, А.Я. Демидов, Н.А. Каратаева, Е.П. Ворошилин // Доклады ТУСУРа. – 2011. – №2 (24). – 59-63 с.
19. Серов А. В. Эфирное цифровое телевидение DVB-T/H. - БХВ-Петербург, 2010 – 465 с.
- 20 . Стандарт DVB-H. Система мобильного ТВ вещания. [Электронный ресурс] – Режим доступа: <http://www.konturm.ru/tech.php?id=dvvh>
- 21.http://www.mathworks.com/examples/simulink-communications/mw/comm_product-LTEDownlinkExample-lte-phy-downlink-with-spatial-multiplexing
22. J. H. Yuen, et. al. Modulation and Coding for Satellite and Space Communications. Proc. IEEE, vol. 78., n. 7, July, 1990, pp. 1250-1265.
23. Forney G. Concatenated Codes. Cambridge, Massachusetts: M. I. T. Press, 1966.
24. Скляр Б. Цифровая связь. Теоретические основы и практическое применение. [2-е изд., испр.]: пер. с англ. - М.: Изд. дом "Вильямс", 2003. - 1104 с.
25. Алферов А.П., Зубов А.Ю., Кузьмин А.С., Черемушкин А.В. Основы криптографии – М.: Гелиос АРВ, 2001. – 480 с.
26. Метлицкий Ю.В. Разработка программного комплекса для визуализации и анализа стандарта криптографической защиты AES, МИФИ 2003 г.
27. Зензин О.С., Иванов М.А. Стандарт криптографической защиты – AES. Конечные поля. – М: КУДИЦ-ОБРАЗ, 2002. -176 с.

28. Б. Шнайер «Прикладная криптография. 2-е издание. Протоколы, алгоритмы и исходные тексты на языке С». - М.: Изд-во "Триумф", 2002. - 816 с.
29. <http://www.des-crypto.ru/cryptography/rc4/>
30. Асосков А.В. и др. Поточные шифры. - М.: КУДИЦ-ОБРАЗ, 2003. - 336 с.
31. <http://www.wisdom.weizmann.ac.il/~itsik/RC4/rc4.html>
32. Романец Ю.В. Защита информации в компьютерных системах и сетях / Под ред. В.Ф. Шаньгина. - 2-е изд., перераб. и доп. - М.: Радио и связь, 2001. - 376 с.
33. RSA Laboratories // <http://www.rsa.com/rsalabs/node.asp?id=2009>
34. Алгоритм RSA : метод. указания к выполнению лабораторных работ для студентов спец. 090105 «Комплексное обеспечение информационной безопасности автоматизированных систем» очной формы обучения / сост.: О. Н. Жданов, И. А. Лубкин ; Сиб. гос. аэрокосмич. ун-т. – Красноярск, 2007. – 38 с.
35. Diffie, D. New directions in Cryptography / D. Diffie, M.Hellman // IEEE Transactions on information theory. November. 1976.
36. Rivest, R. A Method for obtaining digital signatures and public keyCryptosystems / R. Rivest, A. Shamir, L. Adleman // Communications of the ACM. February. 1978.
37. Столингс, В. Криптография и защита сетей: принципы и практика / В. Столингс ; пер. с англ. . – 2-е изд. – М. : Изд. дом «Вильямс», 2001. – 672 с.
38. Коблиц, Н. Курс теории чисел и криптографии / Н. Коблиц ; пер. с англ. М. А. Михайловой и В. Е. Тараканова ; под. ред. А. М. Зубкова. – М. : Науч. изд-во ТВП, 2001. – 254 с.
39. Безопасность на транспортном уровне: SSL и TLS | Лекция | НОУ ИНТУИТ [Электронный ресурс]. – Режим доступа <http://www.intuit.ru/studies/courses/553/409/lecture/9387> (дата обращение 13.04.2015).
40. http://matlab.ru/products/LTE-System-Toolbox/lte-system-toolbox_rus_web.pdf (дата запроса 05.04.2016)