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

Кафедра электронных приборов

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В ЭЛЕКТРОНИКЕ

Учебное пособие
для студентов направления «Электроника и микроэлектроника»
(специальность «Электронные приборы и устройства»

2012

Колегов Алексей Анатольевич

Информационные технологии в электронике: учебное пособие для студентов направления «Электроника и микроэлектроника» (специальность «Электронные приборы и устройства» / А.А. Колегов; Министерство образования и науки Российской Федерации, Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования Томский государственный университет систем управления и радиоэлектроники, Кафедра электронных приборов. - 2-е изд. - Томск: ТУСУР, 2012. - 206 с.

Целью дисциплины является обучение студентов теоретическим и прикладным основам информационной технологии, которая служит фундаментом информационной подготовки инженеров всех видов деятельностей.

Курс основывается на трактовке информационной технологии как совокупности технологических элементов (устройств или методов), используемых для обработки информации.

В результате изучения студенты должны знать: возможности, методы и средства информационных технологий в производстве, научных исследованиях, управленческой и других сферах деятельности.

На основе полученных знаний студент должен уметь: использовать модели, методы и средства информационных технологий при создании автоматизированных систем обработки информации и управления различного назначения, ориентироваться в типовых инструментальных средствах и областях их эффективного применения.

Предназначено для студентов очной и заочной форм, обучающихся по направлению «Электроника и микроэлектроника» (специальность «Электронные приборы и устройства») по дисциплине «Информационные технологии в электронике».

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

Кафедра электронных приборов

УТВЕРЖДАЮ
Зав.кафедрой ЭП
_____С.М. Шандаров
«___» _____ 2012 г.

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ В ЭЛЕКТРОНИКЕ

Учебное пособие
для студентов направления «Электроника и микроэлектроника»
(специальность «Электронные приборы и устройства»

Разработчик
_____А.А.. Колегов
«____»_____2012 г

2012

ОГЛАВЛЕНИЕ

ВВЕДЕНИЕ

Целью дисциплины является обучение студентов теоретическим и прикладным основам информационной технологии, которая служит фундаментом информационной подготовки инженеров всех видов деятельностей.

Курс основывается на трактовке информационной технологии как совокупности технологических элементов (устройств или методов), используемых для обработки информации.

В результате изучения студенты должны знать: возможности, методы и средства информационных технологий в производстве, научных исследованиях, управленческой и других сферах деятельности.

На основе полученных знаний студент должен уметь: использовать модели, методы и средства информационных технологий при создании автоматизированных систем обработки информации и управления различного назначения, ориентироваться в типовых инструментальных средствах и областях их эффективного применения.

1 ЗНАКОМСТВО С ОСНОВНЫМИ ПОНЯТИЯМИ

1.1 Технология

Технология (от греч. techne — искусство, мастерство, умение) — совокупность методов обработки, изготовления, изменения состояния, свойства, формы сырья, материала или полуфабриката, осуществляемых в процессе производства продукции. Задача технологии как науки — выявление физических, химических, механических и других закономерностей с целью определения и использования на практике наиболее эффективных и экономичных производственных процессов. Понятие технологии непрерывно развивается, дополняется все новыми признаками, что является главной причиной многозначности его трактовки.

В узком смысле под технологией понимаются конкретные технологические приемы (например, установка резца под тем или иным углом к обрабатываемой заготовке). В широком смысле технология есть способ освоения человеком материального мира посредством социально организуемой деятельности, которая включает, по Г.И. Марчуку, три компонента: информационный (научные принципы), материальный (орудия труда), социальный (специалисты, владеющие профессиональными навыками) [1].

Технология — это не только некоторая организация естественных процессов, направленная на создание искусственных объектов (например, рецептура приготовления какого-либо материала), но и наука о лучших способах этой организации. Следовательно, необходимо различать технологию как науку и как практику [1].

Поскольку в ходе технологического прогресса определяющая роль принадлежит предметным структурам практики, то деятельность по их проектированию, совершенствованию и организации постепенно обособляется в соответствии с требованиями общественной практики (появляются профессии технологов, организаторов производства, управляющих).

Задача технолога заключается в проектировании и организации предметных структур, составляющих основу того или иного вида деятельности, будь то архитектурный комплекс или технико-технологическая оснастка производственного процесса.

Во всех определениях сущность понятия технологии обычно связывают с понятиями процедуры и операции. При этом под процедурой понимают набор действий (операций), посредством которых осуществляется тот или иной главный процесс (или его отдельный этап), выражающий суть конкретной технологии. А под операцией понимается непосредственное практическое решение задачи в рамках данной процедуры, т.е. однородная логически неделимая часть конкретного процесса [1].

При проектировании любой технологии следует учитывать, что технология, определяя путь расчленения сложного процесса на составляющие его более простые этапы, должна показывать, как необходимо действовать конкретному исполнителю, чтобы добиться максимально эффективного выполнения стоящей перед ним цели. Цель является важным признаком любой технологии, она не только характеризует технологию с сущностной стороны, но и является «стержнем» любого процесса, задавая определенный порядок его осуществления и режим развертывания [1].

Также при проектировании и внедрении любой конкретной технологии возникает необходимость рассматривать ее как структуру (систему), включающую три компонента [1].

1. Технические средства — «твердая» часть технологических систем (hardware). Например, в технологических комплексах АСОУ — это ЭВМ и средства организационной техники.

2. Знания и профессиональные навыки ведения соответствующего процесса — «мягкая» часть технологических систем (software). В АСОУ — это информационно-методическое обеспечение персонала и программно-алгоритмическое обеспечение ЭВМ.

3. Специализированное организационное обеспечение, т.е. организация (организованность), соответствующая уровню и специфике реализуемой, данной технологией и в данных условиях применения принципов и функций (orgware). В АСОУ — это специально спроектированные организационные структуры управления (состав служб, процедуры подготовки и принятия решения, контроль исполнения, экономические показатели, система материального стимулирования и т.д.).

1.2 Информатика, информация и информационная технология (ИТ)

Понятие информатики является более общим и емким, чем технология. По определению А.П. Ершова информатика — это наука, изучающая закономерности и методы накопления, передачи и обработки информации на основе ЭВМ.

Трактовка понятия информатики очень многозначна. Под информатикой понимается раздел науки, и раздел техники, и вид человеческой деятельности. В то же время важно различать понятие информатики как науки, техники и человеческой деятельности.

Определение понятия информатики очень важно, так как с его помощью осуществляется связь между философией и частными науками (кибернетикой, теорией информации, системотехникой и др.), между теоретической и практической деятельностью в области информационной технологии. Информатика стремительно развивается в научном и практическом плане. Это развитие воплощается в становлении новой технологии сбора, обработки, хранения, передачи информации, технологии, которая переводит практику управления, регулирования материального производства, научных исследований, образования и других областей человеческой деятельности на принципиально новый индустриальный уровень.

Информационная технология — это, во-первых, совокупность процессов циркуляции и переработки информации и, во-вторых, описание этих процессов. Объектами переработки и циркуляции является информация, данные. Основными свойствами ИТ являются: целесообразность, наличие компонентов и структуры, взаимодействие с высшей средой, целостность, развитие во времени [2].

За последнее время в развитии ИТ наблюдается период кардинальных изменений, связанных главным образом со сменой технологической базы автоматизированных систем обработки данных — переходом от создания отдельных ВЦ к внедрению локальных и распределенных сетей ЭВМ. Такие системы в совокупности обеспечивают комплексную автоматизацию производства, создание маневренных (гибких) технологических систем, мобильно перестраиваемых на выпуск различной продукции [2].

Эволюция ИТ, основанная на применении сначала простейших счетных приборов, а затем ЭВМ, представляет собой непрерывное вытеснение механических способов обработки информации немеханическими. Сначала механические счетные приборы (арифмометры) были вытеснены электрическими счетными машинами, а электрические — электронными. Электронные ЭВМ начинают вытесняться магнитными и оптическими, т.е., по существу, электромагнитными. В настоящее время мы говорим уже о фотонных и биологических компьютерах. Растет быстродействие ЭВМ, увеличивается

оперативная память, уменьшаются габариты. Информационная технология стремится в соответствии с закономерностями развития технологий к своим предельным параметрам [2].

Процесс развития ИТ можно условно разделить на четыре больших этапа, началом каждого из которых явилось подлинно революционное событие в жизни общества: возникновение речи, изобретение письма, книгопечатания, создание ЭВМ.

Наиболее яркие примеры применения ИТ: персональные компьютеры, локальные сети, средства речевого ввода информации и графопостроители, автоматизированные системы научных исследований (АСНИ), системы автоматизации проектирования (САПР), автоматизация систем управления (АСУ), автоматизация систем обработки информации (АСОИ), автоматизированные системы организационного управления (АСОУ) и т.д.

Новая информационная технология — совокупность внедряемых («встраиваемых») в системы организационного управления принципиально новых средств и методов обработки данных, представляющих собой целостные технологические системы и обеспечивающих целенаправленное создание, передачу, хранение и отображение информационного продукта (данных, идей, знаний) с наименьшими затратами и в соответствии с закономерностями той социальной среды, где развивается НИТ[1].

Информация (от лат. Information — разъяснение, изложение) — первоначально сведения, передаваемые людьми устным, письменным или другим способом (с помощью условных сигналов, технических средств и т.д.); с середины XX в. — общенаучное понятие, включающее обмен сведениями между людьми, человеком и автоматом, автоматом и автоматом; обмен сигналами в животном и растительном мире; передачу признаков от клетки к клетке, от организма к организму; одно из основных понятий кибернетики.

Слово информация означает сообщение, осведомление о чем-либо. Однако такое переводческое толкование далеко не всегда может служить определением понятия информации [2]. Поэтому и в науке, и на практике используется множество определений этого понятия: от наиболее общего, философского — информация есть отражение реального мира — до наиболее узкого, практического — информация есть все сведения, являющиеся объектом хранения, передачи и преобразования. Развитие ИТ привело к возникновению и такого чисто технологического понятия, как машинная информация, под которой подразумевается информация, зафиксированная в виде, непосредственно доступном обработке на ЭВМ, включая ее передачу с электронными скоростями без пространственного перемещения носителя. К категории машиной информации относятся информационные фонды в виде баз данных и программных средств.

В управленческой деятельности понятие информации может быть истолковано как некоторая совокупность сведений (сообщений),

определяющих меру наших знаний о тех или иных событиях, явлениях, фактах и их взаимосвязи [2].

В настоящее время много говорят о ценности информации, об экономической эффективности системы информационных потоков, их интенсификации и упорядочивании путем внедрения средств современной информатики. Однако важно учитывать, что материальные ценности создает только предприятие, а не система информационных потоков. Все, что можно сделать с помощью информационной системы, это представить руководителям информацию, требующуюся для организации производственного процесса. Но чтобы организовать этот процесс рациональным образом, информация должна умело использоваться в системе управления.

Постоянно возрастающие мощность и разнообразие средств обработки информации вызвали к жизни термин «информатика», но уже в его качественно новом и более емком понимании, чем прежде [2].

2 ОБРАБОТКА ИНФОРМАЦИИ

2.1 Понятие сообщения

Сведения о состоянии объекта в ИС формируются в виде сообщений. Под сообщением понимается все то, что подлежит передаче. Независимо от содержания сообщение обычно представляется в виде электрического, звукового, светового, механического или других сигналов. Таким образом, сообщение отображает некоторые исходные сигналы любого вида и по свойствам зависит от исходных сигналов. Изменение некоторого физического процесса во времени, обеспечивающее передачу сообщения (а тем самым и информации), называется *сигналом*. Характеристика сигнала, которая служит для представления сообщения, называется *параметром сигнала* [3].

В ИС все исходные сигналы, поступающие от объекта, можно разделить на две большие группы: сигналы оптические, которые отображают устойчивые состояния некоторых объектов и могут быть представлены, например, в виде определенного положения элемента, системы, текста в документе, определенного состояния электронного устройства и т.д., и сигналы динамические, для которых характерно быстрое изменение во времени, отображающее, например, изменения электрических параметров системы.

Динамические и статические сигналы имеют свои области использования. Статические сигналы существенное место занимают при подготовке, регистрации и хранении информации. Динамические используются в основном для передачи информации. Однако заметим, что это не всегда является обязательным.

По характеру изменения сигналов во времени различают сигналы непрерывные и дискретные. Непрерывный сигнал отображается некоторой непрерывной функцией и физически представляет собой непрерывно

изменяющиеся значения колебаний. Дискретный сигнал характеризуется конечным множеством значений и в зависимости от исходного состояния принимает значения, связанные с определенным состоянием системы. Исходя из физической сущности процесса, свойственного объекту управления, можно выделить некоторые разновидности непрерывных и дискретных функций, отображающих реальные сигналы:

1) непрерывную функцию непрерывного аргумента. Функция имеет вид $f(t)$, непрерывна на всем отрезке и может описать реальный сигнал в любой момент времени. При этом не накладывается никаких ограничений на выбор момента времени и значения самой функции;

2) непрерывную функцию дискретного аргумента. Обычно такие сигналы возникают при квантовании непрерывных величин по времени. В этом случае задаются некоторые фиксированные моменты времени t_J, отсчитываемые через интервал Dt. который обычно определяется спектральными свойствами исходного физического процесса. Функция $f(t_J)$ может принимать любые мгновенные значения, но она определяется лишь для дискретных значений времени. Этот вид сигналов и связанных с ним функций имеет место при формировании исходных сообщений из непрерывных величин;

3) дискретную функцию непрерывного аргумента $f_J(t)$. В этом случае функция имеет ряд конечных дискретных значений, однако определена на всем отрезке времени t для любого мгновенного значения времени. Дискретизация самой функции связана с созданием шкалы квантования по уровню, что свойственно различным датчикам, при этом шаг квантования определяется требуемой точностью воспроизведения исходной величины;

4) дискретную функцию дискретного аргумента $f_J(t_J)$. В этом случае функция принимает одно из возможных дискретных значений, общее число которых является конечным, и определяется для окончательного набора дискретных значений времени. Имеем дискретизацию как по уровням, так и по моментам времени.

2.2 Интерпретация информации

Полученные из информационных кодов данные интерпретируются объектом. Что это означает? Прежде всего, устанавливается их значение для этого объекта. Значения данных определяются их сопоставлением с комплексом целей объекта и выделением тех из них, к которым объект может приблизиться, реализуя полученную в итоге информацию. Для этого объект должен обладать сформированной к моменту начала обработки данных структурой текущих целей. Эта структура может быть представлена многоуровневым комплексом элементов, каждый из которых соответствует необходимости достижения объектом какой- либо одной цели. Связи между элементами определяются зависимостью достижения одних целей от достижения других. Каждый элемент ассоциирован с набором возможных действий объекта, влияющим на достижение

соответствующей цели, и характером тех данных, которые могут дать ему информацию, способствующую выбору целесообразных действий. Структура целей может иметь частью статический, а частью динамический характер, что зависит от свойств самого объекта. Это касается состава элементов, их внутреннего содержания и связей между ними. Эту структуру можно назвать памятью целей объекта.

Данные, не соответствующие никаким целям объекта, не несут для него информацию и потому пропадают, возвращая объект в то состояние, в котором он был до получения этих данных. Бесцельное использование данных означает нарушение целесообразности функционирования объекта, и если таковые становятся значительными, то это ведет к прекращению его существования.

Вторым шагом после определения значимости данных для объекта является либо непосредственное их восприятие как информации и безусловная реализация (рефлекторная дуга), либо они сохраняются в элементах памяти, связанных с установленными на предыдущем шаге целями объекта. Комплекс ранее сохраненных и вновь поступивших данных, связанных по цели их хранения, оценивается на достаточность их совокупности для выбора действий объекта, приближающих его к соответствующей цели. Процесс оценки может иметь различную природу в зависимости от свойств объекта, но в его основе лежит сопоставление имеющегося комплекса данных с построенными ранее для данной цели информационными шаблонами действий. Информационные шаблоны действий объекта могут быть врожденными (статическими) или построенными им в результате предыдущих актов информационных взаимодействий (динамическими).

Информационные шаблоны действий обеспечивают сопоставление характеристик наборов данных, действий и результатов приближения к цели. Другими словами, с их помощью оценивается возможный результат действий по достижению соответствующей цели при наличии определенных данных. Способность строить динамические шаблоны определяется наличием возможности у объекта изменять некоторые элементы своей памяти в соответствии с тем, какие его действия при наличии какой информации приводили к какому результату.

Здесь мы подошли к тому, что при определенном уровне развития объектов им становятся присущи свойства информационного моделирования своих взаимодействий с внешней средой, которое используется для выбора наиболее целесообразного для них поведения. Таким образом, правомерно говорить о наличии внутри объекта информационной модели внешней среды и его взаимодействии с ней.

Информационная модель внешней среды объекта – это структурированная совокупность трех компонент:

1) воспринятой объектом информации, запомненной в виде данных;

2) информационных шаблонов действий объекта;

3) методов сопоставления первых двух компонент в соответствии с комплексом целей объекта.

Конкретные реализации этой модели у разных объектов могут иметь различную элементную базу, но концептуально они строятся и действуют по общим принципам, которые вытекают из общего их назначения и общности свойств информационных процессов в природе.

2.3 Кодирование информации

Кодирование информации – это процесс формирования определенного представления информации.

В более узком смысле под термином «кодирование» часто понимают переход от одной формы представления информации к другой, более удобной для хранения, передачи или обработки.

Компьютер может обрабатывать только информацию, представленную в числовой форме. Вся другая информация (например, звуки, изображения, показания приборов и т.д.) для обработки на компьютере должна быть преобразована в числовую форму. Например, чтобы перевести в числовую форму музыкальный звук, можно через небольшие промежутки времени измерять интенсивность звука на определенных частотах, представляя результаты каждого измерения в числовой форме. С помощью программ для компьютера можно выполнить преобразования полученной информации, например «наложить» друг на друга звуки от разных источников.

Аналогичным образом на компьютере можно обрабатывать текстовую информацию. При вводе в компьютер каждая буква кодируется определенным числом, а при выводе на внешние устройства (экран или печать) для восприятия человеком по этим числам строятся изображения букв. Соответствие между набором букв и числами называется кодировкой символов.

Как правило, все числа в компьютере представляются с помощью нулей и единиц (а не десяти цифр, как это привычно для людей). Иными словами, компьютеры обычно работают в двоичной системе счисления, поскольку при этом устройства для их обработки получаются значительно более простыми. Ввод чисел в компьютер и вывод их для чтения человеком может осуществляться в привычной десятичной форме, а все необходимые преобразования выполняют программы, работающие на компьютере.

Среди всего разнообразия информации, обрабатываемой на компьютере, значительную часть составляют числовая, текстовая, графическая и аудиоинформация. Познакомимся с некоторыми способами кодирования этих типов информации в ЭВМ.

2.3.1 Кодирование чисел

Существуют два основных формата представления чисел в памяти компьютера. Один из них используется для кодирования целых чисел, второй (так называемое представление числа в формате с плавающей точкой) используется для задания некоторого подмножества действительных чисел.

Множество целых чисел, представимых в памяти ЭВМ, ограничено. Диапазон значений зависит от размера области памяти, используемой для размещения чисел. В k-разрядной ячейке может храниться 2^k различных значений целых чисел.

Чтобы получить внутреннее представление целого положительного числа N, хранящегося в k-разрядном машинном слове, необходимо:

1) перевести число N в двоичную систему счисления;

2) полученный результат дополнить слева незначащими нулями до k разрядов.

Пример

Получить внутреннее представление целого числа 1607 в 2-байтовой ячейке.

Переведем число в двоичную систему:

$1607_{10} = 11001000111_2$.

Внутреннее представление этого числа в ячейке будет следующим:

0000 0110 0100 0111.

Для записи внутреннего представления целого отрицательного числа ($–N$) необходимо:

1) получить внутреннее представление положительного числа N;

2) инвертировать код этого числа;

3) к полученному числу прибавить 1.

Пример

Получим внутреннее представление целого отрицательного числа –1607. Воспользуемся результатом предыдущего примера и запишем внутреннее представление положительного числа 1607: 0000 0110 0100 0111. Инвертированием получим обратный код: 1111 1001 1011 1000. Добавим единицу: 1111 1001 1011 1001 — это и есть внутреннее двоичное представление числа –1607.

Формат с плавающей точкой использует представление вещественного числа R в виде произведения мантиссы m на основание системы счисления n в некоторой целой степени p, которую называют порядком:

$$R = m \cdot n^p.$$

Представление числа в форме с плавающей точкой неоднозначно. Например, справедливы следующие равенства:

$12.345 = 0.0012345 \cdot 10^4 = 1234.5 \cdot 10^{-2} = 0.12345 \cdot 10^2$

Чаще всего в ЭВМ используют нормализованное представление числа в форме с плавающей точкой. Мантисса в таком представлении должна удовлетворять условию:

$$0.1_p <= m < 1_p.$$

Иначе говоря, мантисса меньше 1 и первая значащая цифра — не ноль (p — основание системы счисления).

В памяти компьютера мантисса представляется как целое число, содержащее только значащие цифры (0 целых и запятая не хранятся); так, для числа 12.345 в ячейке памяти, отведенной для хранения мантиссы, будет сохранено число 12345. Для однозначного восстановления исходного числа остается сохранить только его порядок, в данном примере — это 2.

2.3.2 Кодирование текста

Множество символов, используемых при записи текста, называется *алфавитом*. Количество символов в алфавите называется его *мощностью*.

Для представления текстовой информации в компьютере чаще всего используется алфавит мощностью 256 символов. Один символ из такого алфавита несет 8 бит информации, т.к. $2^8 = 256$. Но 8 бит составляют один байт, следовательно, двоичный код каждого символа занимает 1 байт памяти ЭВМ.

Все символы такого алфавита пронумерованы от 0 до 255, а каждому номеру соответствует 8-разрядный двоичный код от 00000000 до 11111111. Этот код является порядковым номером символа в двоичной системе счисления.

Для разных типов ЭВМ и операционных систем используются различные таблицы кодировки, отличающиеся порядком размещения символов алфавита в кодовой таблице. Международным стандартом на персональных компьютерах является уже упоминавшаяся таблица кодировки ASCII.

Принцип последовательного кодирования алфавита заключается в том, что в кодовой таблице ASCII латинские буквы (прописные и строчные) располагаются в алфавитном порядке. Расположение цифр также упорядочено по возрастанию значений.

Стандартными в этой таблице являются только первые 128 символов, т.е. символы с номерами от нуля (двоичный код 00000000) до 127 (01111111). Сюда входят буквы латинского алфавита, цифры, знаки препинания, скобки и некоторые другие символы. Остальные 128 кодов, начиная со 128 (двоичный код 10000000) и кончая 255 (11111111), используются для кодировки букв национальных алфавитов, символов псевдографики и научных символов.

2.3.3 Кодирование графической информации

В видеопамяти находится двоичная информация об изображении, выводимом на экран. Почти все создаваемые, обрабатываемые или просматриваемые с помощью компьютера изображения можно разделить на две большие части — растровую и векторную графику.

Растровые изображения представляют собой однослойную сетку точек, называемых пикселями (pixel, от англ. picture element). Код пикселя содержит информации о его цвете.

Для черно-белого изображения (без полутонов) пиксель может принимать только два значения: белый и черный (светится — не светится), а для его кодирования достаточно одного бита памяти: 1 — белый, 0 — черный.

Пиксель на цветном дисплее может иметь различную окраску, поэтому одного бита на пиксель недостаточно. Для кодирования 4-цветного изображения требуются два бита на пиксель, поскольку два бита могут принимать 4 различных состояния. Может использоваться, например, такой вариант кодировки цветов: 00 — черный, 10 — зеленый, 01 — красный, 11 — коричневый.

На RGB-мониторах все разнообразие цветов получается сочетанием базовых цветов — красного (Red), зеленого (Green), синего (Blue), из которых можно получить 8 основных комбинаций:

R	G	B	Цвет
0	0	0	Черны
0	0	1	Синий
0	1	0	Зеленый
0	1	1	Голубой
1	0	0	Красный
1	0	1	Розовый
1	1	0	Коричневый
1	1	1	Белый

Разумеется, если иметь возможность управлять интенсивностью (яркостью) свечения базовых цветов, то количество различных вариантов их сочетаний, порождающих разнообразные оттенки, увеличивается. Количество различных цветов — K и количество битов для их кодировки — N связаны между собой простой формулой:

$$2^N = K.$$

В противоположность растровой графике векторное изображение многослойно. Каждый элемент векторного изображения — линия, прямоугольник, окружность или фрагмент текста — располагается в своем собственном слое, пиксели которого устанавливаются независимо от

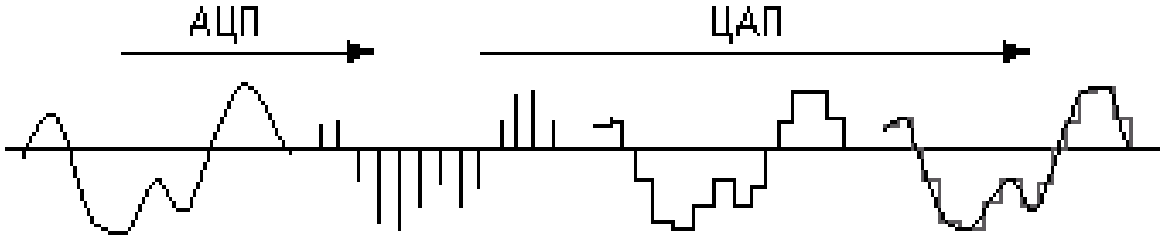
других слоев. Каждый элемент векторного изображения является объектом, который описывается с помощью специального языка (математических уравнений, линий, дуг, окружностей и т.д.). Сложные объекты (ломаные линии, различные геометрические фигуры) представляются в виде совокупности элементарных графических объектов.

Объекты векторного изображения, в отличие от растровой графики, могут изменять свои размеры без потери качества (при увеличении растрового изображения увеличивается зернистость).

2.3.4 Кодирование звука

Из курса физики вам известно, что звук — это колебания воздуха. Если преобразовать звук в электрический сигнал (например, с помощью микрофона), мы увидим плавно изменяющееся с течением времени напряжение. Для компьютерной обработки такой — аналоговый — сигнал нужно каким-то образом преобразовать в последовательность двоичных чисел.

Поступим следующим образом. Будем измерять напряжение через равные промежутки времени и записывать полученные значения в память компьютера. Этот процесс называется дискретизацией (или оцифровкой), а устройство, выполняющее его, — аналого-цифровым преобразователем (АЦП):



Для того чтобы воспроизвести закодированный таким образом звук, нужно выполнить обратное преобразование (для него служит цифроаналоговый преобразователь — ЦАП), а затем сгладить получившийся ступенчатый сигнал.

Чем выше частота дискретизации (т.е. количество отсчетов за секунду) и чем больше разрядов отводится для каждого отсчета, тем точнее будет представлен звук. Но при этом увеличивается и размер звукового файла. Поэтому в зависимости от характера звука, требований, предъявляемых к его качеству и объему занимаемой памяти, выбирают некоторые компромиссные значения.

Описанный способ кодирования звуковой информации достаточно универсален, он позволяет представить любой звук и преобразовывать его самыми разными способами. Но бывают случаи, когда выгодней действовать по-иному.

Человек издавна использует довольно компактный способ представления музыки — нотную запись. В ней специальными символами указывается, какой высоты звук, на каком инструменте и как сыграть. Фактически, ее можно считать алгоритмом для музыканта, записанным на особом формальном языке. В 1983 г. ведущие производители компьютеров и музыкальных синтезаторов разработали стандарт, определивший такую систему кодов. Он получил название MIDI.

Конечно, такая система кодирования позволяет записать далеко не всякий звук, она годится только для инструментальной музыки. Но есть у нее и неоспоримые преимущества: чрезвычайно компактная запись, естественность для музыканта (практически любой MIDI-редактор позволяет работать с музыкой в виде обычных нот), легкость замены инструментов, изменения темпа и тональности мелодии.

Заметим, что существуют и другие, чисто компьютерные, форматы записи музыки. Среди них следует отметить формат MP3, позволяющий с очень большим качеством и степенью сжатия кодировать музыку. При этом вместо 18—20 музыкальных композиций на стандартный компакт-диск (CDROM) помещается около 200. Одна песня занимает примерно 3,5 Mb, что позволяет пользователям сети Интернет легко обмениваться музыкальными композициями.

2.4 Устройства связи и передачи сообщений

Письмо и газета относятся к самым старым и до сих пор не устаревшим способа передачи сообщений посредством записи на долговременном носителе сообщений. В случае передачи информации с помощью недолговременного носителя сообщений человек использует также различные физические устройства в соответствии с уровнем развития техники на данный момент. Примерами таких устройств связи служат телефон, радио и телевидение, предназначенные как для случайной, так и для регулярной передачи сообщений [3].

Внешнее устройство связи состоит из приемника (получателя) и передатчика (отправителя). О внутреннем строении устройств связи никаких общих утверждений сделать нельзя, при более внимательном рассмотрении многие из них оказываются составленными из нескольких более мелких устройств связи. Один и тот же носитель может использоваться для сообщений на входе и на выходе. Такие устройства служат для усиления или регенерации сообщения, связанной с устранением помех, и называются релейными линиями. Примерами таких устройств служат рупор, слуховая трубка, а также их современные электронные варианты — мегафон и слуховой аппарат. Устройства связи называют преобразователем, если на входе и на выходе устройства используются различные физические носители [3].

Если устройство предназначено для связи между людьми, то сообщения на входе и выходе должны быть производимы или

воспринимаемы людьми, т.е. носители должны соответствовать человеческим эффекторам и рецепторам. В качестве примеров могут служить клавишный музыкальный инструмент (физический носитель на входе — давление, на выходе — звуковые волны) и осциллограф, управляемый через микрофон (на входе — звуковые волны, на выходе — световые волны) [3].

Если приемником составного устройства является приемник первого устройства, участвующего в соединении, а передатчиком — передатчик последнего устройства, то между передатчиком одного устройства связи и приемником другого могут использоваться и такие носители, которые не доступны человеческим эффекторам и рецепторам. Примером является телефонная связь по проводам или радио. Протяженную в пространстве среду, через которую носитель сообщения передается от передатчика к приемнику, называют каналом связи [3].

Для сообщений, которыми обмениваются люди, в большинстве случаев имеются соглашения относительно их формы. О таких сообщениях говорят, что они передаются в языковой форме, что они составлены на некотором языке. Когда мы говорим о языковых сообщениях, мы имеем в виду то общее, что присуще каждому из этих случаев; способ передачи — письменно, устно, посредством осязания или еще как-то — не имеет здесь никакого значения [3].

3 ТЕХНОЛОГИЯ ХРАНЕНИЯ ИНФОРМАЦИИ

3.1 Типы запоминающих устройств

3.1.1 ОЗУ, ПЗУ

Компактная микроэлектронная «память» широко применяется в современной электронной аппаратуре самого различного значения. Память определяют как функциональную часть ЭВМ, предназначенную для записи, хранения и выдачи команд и обрабатываемых данных. Комплекс технических средств, реализующих функцию памяти, называют *запоминающим устройством* (ЗУ).

Для обеспечения работы процессора (микропроцессора) необходимы программа, т.е. последовательность команд, и данные, над которыми процессор производит предписываемые командами операции. Команда и данные поступают в основную память ЭВМ через устройство ввода, на выходе которого они получают цифровую форму представления, т.е. форму кодовых комбинаций (0 и 1). Основная память, как правило, состоит из ЗУ двух видов — *оперативного* (ОЗУ) и *постоянного* (ПЗУ).

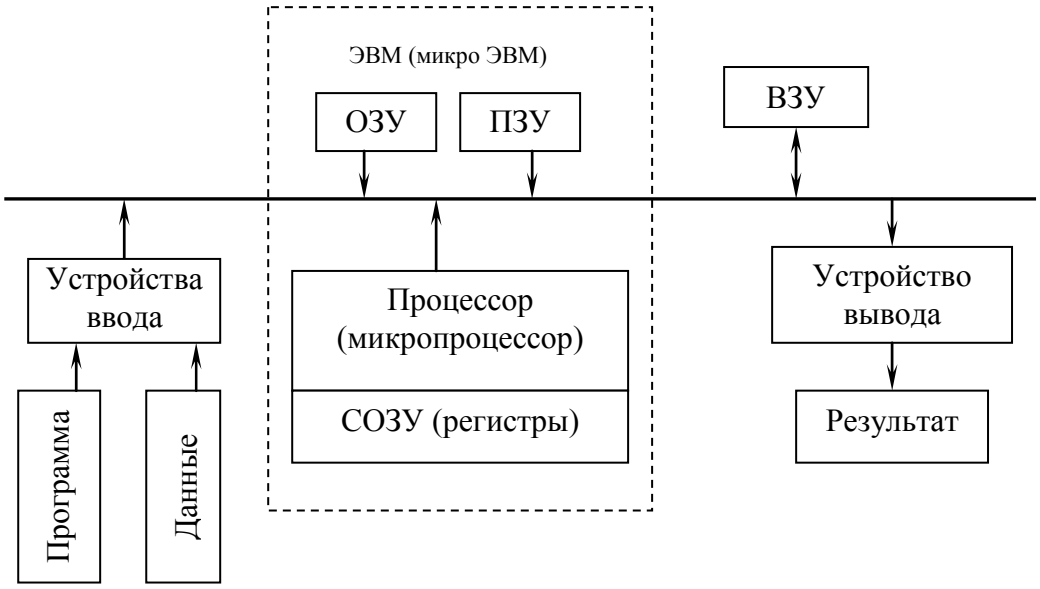


Рисунок 3.1 — Организация работы процессора

ОЗУ предназначено для хранения переменной информации, оно допускает изменение своего содержимого в ходе выполнения процессором вычислительных операций с данными. Это значит, что процессор может выбрать (режим считывания) из ОЗУ код команды и данные и после обработки поместить в ОЗУ (режим записи) полученный результат. Причём возможно размещение в ОЗУ новых данных на местах прежних, которые в этом случае перестают существовать. Таким образом, ОЗУ может работать в режимах записи, считывания и хранения информации.

ПЗУ содержит информацию, которая не должна изменятся в ходе выполнения процессором программы. Такую информацию составляют стандартные подпрограммы, табличные данные, коды физических констант и постоянных коэффициентов и т.п. И эта информация заносится в ПЗУ предварительно, например путём пережигания легкоплавких перемычек в структуре ПЗУ, и в ходе работы процессора может только считываться. Таким образом, ПЗУ работает только в режимах хранения и считывания.

Функциональные возможности ОЗУ шире, чем ПЗУ: ОЗУ может работать в качестве ПЗУ, т.е. в режиме многократного считывания однократно записанной информации, а ПЗУ в качестве ОЗУ работать не может, т.к. не позволяет изменить однократно записанную в ней информацию. Далее коснёмся разновидности ПЗУ, которая допускает перепрограммирование, однако и это ПЗУ не может заменить ОЗУ.

В свою очередь, ПЗУ обладает преимуществом перед ОЗУ в свойстве сохранять информацию при сбоях и отключении питания. Это свойство получило название *энергозависимость*. ОЗУ является энергозависимым, т.к. информация, записанная в ОЗУ, утрачивается при сбоях питания.

Для обеспечения надёжной работы ЭВМ при отказах питания нередко ПЗУ используют и в качестве памяти программ. В таком случае

программа заносится в ПЗУ предварительно и уже не может быть заменена в данном ПЗУ другой программы. Очевидно, использование ПЗУ таким образом целесообразно прежде всего в специализированных автоматических устройствах, работающих по постоянной программе.

Запоминающее устройство, реализующее функции основной памяти, размещают рядом с процессором на одной плате, в одном блоке в зависимости от типа ЭВМ, и такое ЗУ в этом смысле является внутренним. Быстродействие внутреннего ЗУ должно быть соизмеримо с быстродействием процессора. Практически это требование не всегда удаётся выполнить: по временным параметрам ОЗУ и ПЗУ отстают от процессора. По этому внутри ЭВМ размещают ещё и вспомогательную (буферную) память на быстродействующих регистрах, которые используются в качестве сверхоперативного ЗУ (СОЗУ) с небольшой информационной ёмкостью.

Наряду с внутренней памятью существует и внешняя память, реализуемая обычно на магнитных носителях: лентах или дисках.

Перейдём к вопросу о реализации внутренней и внешней памяти ЭВМ на основе микроэлектронной элементной базы. В современных вычислительных средствах и в электронной аппаратуре различного функционального назначения для построения ОЗУ и ПЗУ, а также регистровых ЗУ широко применяют полупроводниковые интегральные микросхемы. Микросхемы памяти изготавливают по полупроводниковой технологии на основе хранения с высокой степенью интеграции компонентов на кристалле, что определяет их принадлежность к большим интегральным схемам (БИС). Конструктивно БИС памяти представляет собой полупроводниковый кристалл с площадью в несколько десятков квадратных миллиметров, заключённый в корпус.

Для самой общей характеристики памяти принимают в расчёт информационную ёмкость, быстродействие, энергопотребление. Информационную ёмкость определяют числом единиц информации в битах и байтах, которые БИС памяти может хранить одновременно. Быстродействие характеризуют временными параметрами, в частности временем цикла записи и считывания. Энергопотребление определяют произведением тока по потреблению и напряжения источников питания.

Для общего представления о микросхемах памяти как функциональных узлах электронной аппаратуры, рассмотрим их наиболее характерные свойства, отражающие принцип построения и управления работы.

Основной частью ОЗУ является массив элементов памяти, объединенных в матрицу накопителя. ЭП может хранить один бит (0 или 1) информации. Каждый ЭП имеет свой адрес. Для обращения к ЭП надо его «выбрать» с помощью кода адреса, сигналы которого подводят к соответствующим выводам микросхемы. ЗУ, ОЗУ или ПЗУ, которые допускают обращение по адресу к любому ЭП в произвольном порядке, называют запоминающим устройством с произвольной выборкой (ЗУПВ).

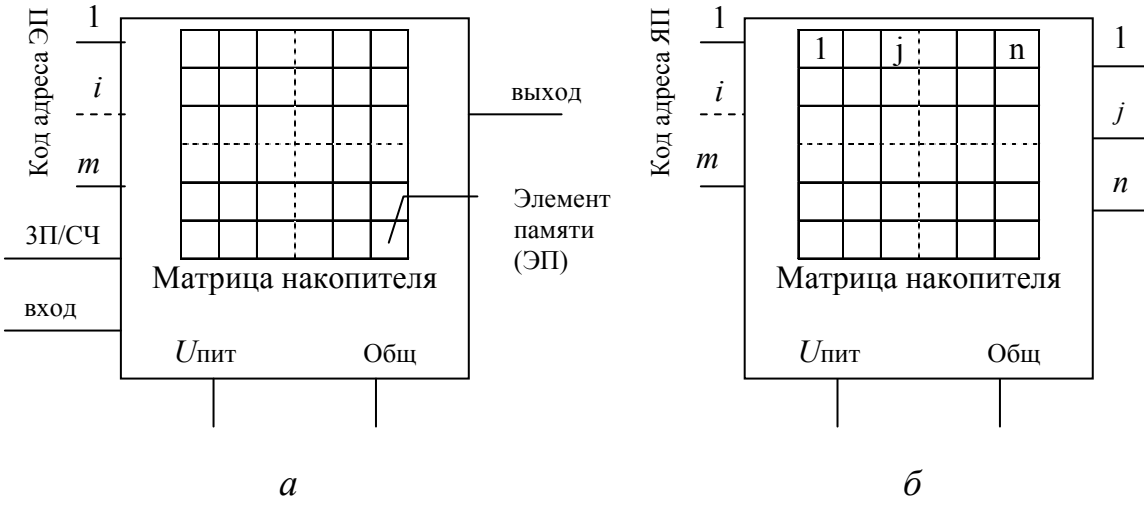


Рисунок 3.2 — Микросхема памяти как функциональный узел:
а — ОЗУ; *б* — ПЗУ

Разрядность кода адреса «*m*» определяет информационную ёмкость микросхемы ОЗУ, т.е. число ЭП в матрице накопителя.

Для ввода и вывода информации служит вход и выход микросхемы. Для управления режимом микросхемы памяти необходим сигнал «Запись-считывание», значение 1 которого определяет режим записи бита информации в ЭП, 0 — режим считывания бита информации из ЭП. Такую организацию матрицы накопителя, при которой можно записывать или считывать один бит, называют одноразрядной.

Микросхемы ОЗУ по типу ЭП разделяют на статические и динамические. В статических ОЗУ в качестве ЭП применены статистические триггеры на биполярных или МДП-транзисторах (МДП-структура «металл — диэлектрик — полупроводник»). Как известно, статистический триггер способен сохранять свое состояние неограниченное время. Число состояний, в которых может находиться триггер, равно двум, что и позволяет использовать его для хранения двоичной единицы информации. В динамических ОЗУ ЭП выполнены на основе электрических конденсаторов, сформированных внутри полупроводникового кристалла. Такие ЭП не могут долго сохранять свое состояние, определяемое наличием или отсутствием электрического заряда, и потому нуждаются в периодическом восстановлении (регенерации). Динамические ОЗУ отличаются от статистических ОЗУ большей информационной емкостью, что обусловлено меньшим числом компонентов в одном ЭП и, следовательно, более плотным их размещением в полупроводниковом кристалле, поэтому динамические ОЗУ сложнее в применении, поскольку нуждаются в организации принудительной регенерации, и в дополнительном оборудовании, и в усложнении устройств управления.

Микросхемы ПЗУ построены также по принципу матричной структуры накопителя (рис. 3.2, *б*). Функции ЭП в микросхемах ПЗУ выполняют перемычки в виде проводников, диодов или транзисторов между шинами строк и столбцов в накопителе. В матрице наличие перемычки соответствует, например, 1 а ее отсутствие — 0. Микросхемы ПЗУ имеют словарную организацию, и поэтому информация считывается в форме многоразрядного кода, т.е. словам. Совокупность ЭП в матрице накопителя, в которой размещается слово, называют *ячейкой памяти* (ЯП). Число ЭП в ЯП определяет ее разрядность n. Каждая ЯП имеет свой адрес, и для обращения к определенной ЯП для считывания из нее информации необходимо к адресным выводам микросхемы подвести сигналы кода, соответствующего данной ячейке адреса. Число ячеек памяти равно 2^m, а информационная емкость микросхемы — $2^{m}*^{n}$ бит.

Занесение информации в микросхемах ПЗУ, т.е. их программирование, осуществляют в основном двумя способами. Один способ заключается в формировании в накопителе перемычек в местах пересечения строк и столбцов матрицы через маску на заключительной технологической стадии изготовления микросхемы ПЗУ. Такие микросхемы ПЗУ называют *масочными*. Другой способ программирования микросхемы ПЗУ основан на пережигании легкоплавких перемычек в тех пересечениях шин строк и столбцов, куда должны быть записаны 0 или 1, в зависимости от принятого кодирования. В исходном состоянии такая микросхема имеет в матрице перемычки во всех пересечениях строк и столбцов. Программирование осуществляет пользователь электрическими импульсами с помощью устройства для программирования, называемого *программатором*.

Микросхемы ПЗУ, масочные (ПЗУМ) и программируемые (ППЗУ), допускают однократное программирование, поскольку оно осуществляется формированием или разрушением соединений в матрице. Один из вариантов реализации ПЗУ ориентирован на программирование заданных логических функций. Такие ПЗУ называют программируемыми логическими матрицами (ПЛМ).

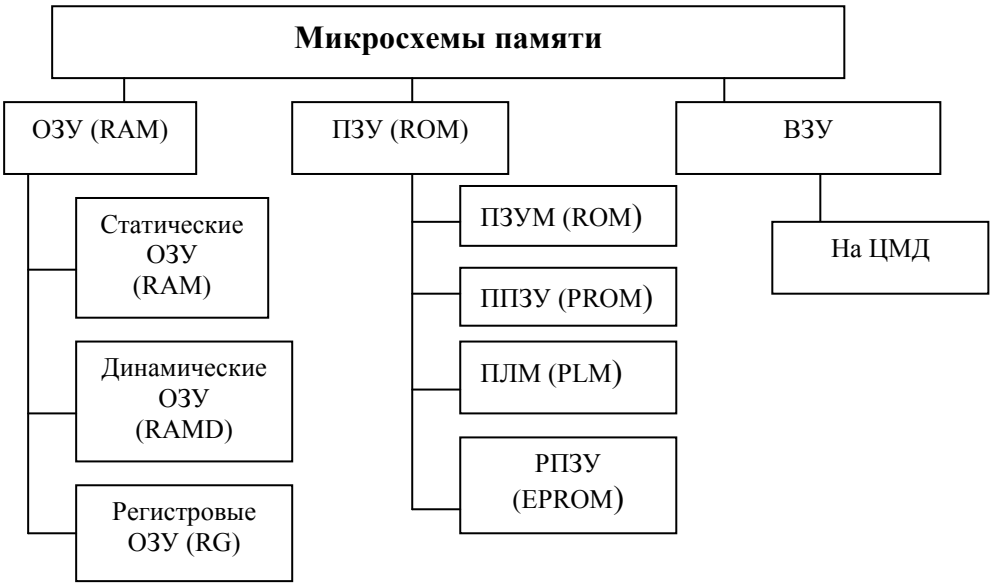


Рисунок 3.3 — Классификация микросхем памяти

Существует разновидность микросхем ПЗУ, допускающая неоднократное (сотни и тысячи циклов) программирование (репрограммирование). ЭП в микросхемах, репрограммируемых ПЗУ, (РПЗУ) является МДП-транзистор, обладающий свойством переходить в состояние проводимости под воздействием импульса программирующего напряжения и сохранять это состояние длительное время (тысяча часов). Данный эффект обусловлен накоплением электрического заряда в подзатворном диэлектрике. Если на транзистор не воздействовать импульсом программирующего напряжения, то он сохраняет закрытое для электрического тока состояние. Для стирания информации перед новым циклом программирования необходимо вытеснить накопленный под затвором заряд. В зависимости от способа выполнения этой операции микросхемы РПЗУ разделяют на два вида: со стиранием электрическим сигналом (РПЗУ — ЭС) и ультрафиолетовым светом (РПЗУ — УФ), которым полупроводниковый кристалл облучают через специальное окно в крышке корпуса. Микросхемы РПЗУ сохраняют информацию без питания, т.е. являются энергонезависимыми.

В соответствии с принятой системой (ОСТ 11 073. 915-80) обозначение микросхемы содержит четыре обязательных элемента.

Первый элемент – цифра, указывающая группу микросхемы по конструктивно-технологическому признаку: 1, 5, 6, 7 – полупроводниковые, 2, 4, 8 – гибридные, 3 – прочие (пленочные, пьезокерамические). Второй элемент — две-три цифры, указывающие номер разработки данной серии. В сочетании указанные два элемента составляют номер серии, к которой принадлежит микросхема. Третий элемент – две буквы, обозначающие функциональную подгруппу и вид микросхемы: РУ — ОЗУ с управлением, РМ — матрицы ОЗУ, РЕ — масочные ПЗУ, РФ — репрограммируемое ПЗУ со стиранием информации ультрафиолетовым

светом, РТ — программируемое ПЗУ, РР — репрограммируемое ПЗУ со стиранием информации электрическим сигналом, РЦ — ЗУ на ЦМД, ИР — регистры. Четвёртый элемент — порядковый номер разработки микросхемы в серии микросхем одного вида. Перед первым элементом для характеристики условий применения, материала и типа корпуса могут размещаться: К – общетехнического применения, Э – экспортное исполнение, Р – пластмассовый корпус типа 2, Е – металлополимерный корпус типа 2, М – керамический, металло- или стеклокерамический корпус типа 2, А – пластмассовый корпус типа 4, И – стеклокерамический корпус типа 4, Н – керамический кристаллоноситель, Б – бескорпусное исполнение. После четвёртого элемента может быть размещена дополнительная группа: А, Б, В и т.д., определяющая условие разбраковки микросхем по одному из функциональных параметров: быстродействию, потребляемому току и др.

3.1.2 Однократно программируемые ЗУ

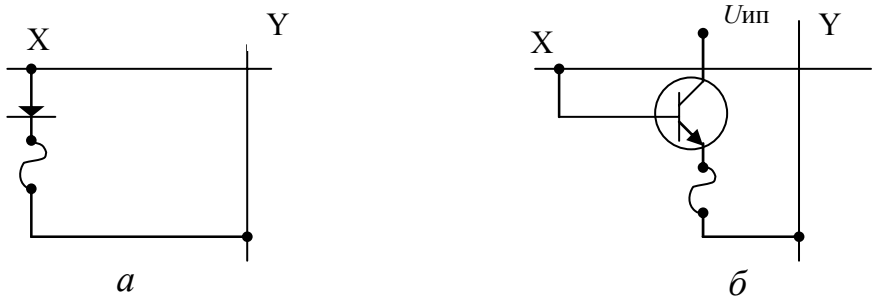
ППЗУ получили широкое распространение среди всех электрически программируемых устройств памяти. Принцип действия ячейки ППЗУ основан на физических процессах, позволяющих необратимо изменить электрическое сопротивление двухполюсника. По принципу действия различают два типа однократно программируемых запоминающих элементов (ЗЭ): резисторный и диодный, в которых программирование осуществляется соответственно пережиганием плавких перемычек и пробоем p-n-переходов.

Бит информации, хранящийся в ЗЭ резисторного типа, определяется наличием или отсутствием плавкой перемычки. В режиме считывания на ЗЭ подают напряжение и хранимое значение бита определяют по значению тока, протекающего через перемычку. В состоянии после изготовления ЗЭ хранит 1 (сопротивление перемычки мало), а после пережигания плавкой перемычки — 0. В качестве плавких перемычек широко применяют тонкие плёнки из нихрома или поликристаллического кремния. Сопротивление перемычки составляет около 10 Ом. В результате программирования через перемычку пропускают импульс тока, плотностью около 10^7 А/см^2, в результате чего она необратимо разрушается. Вследствие малых размеров перемычки и большой энергии, выделяемой при пережигании, физические процессы в плёнке достаточно сложны.

Работа ЗЭ диодного типа основана на необратимых явлениях, происходящих при пробое обратносмещенного p-n-перехода. В исходном состоянии ЗЭ диодного типа хранит 0, а его обратное сопротивление очень велико. При программировании к диоду прикладывается запирающее напряжение повышенного уровня, под действием которого p-n-переход пробивается, т.е. происходит короткое замыкание (состояние логического 0).

В схеме ячейки на рис. 3.4, *а* в режиме считывания подается положительное напряжение на шину X, а выходной сигнал снимается с

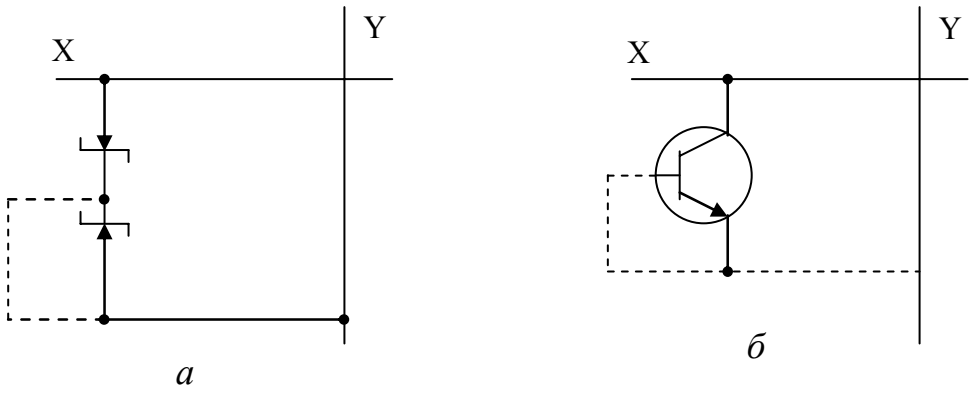
нагрузки, включенной последовательно в шину Y. В режиме программирования (запись 0) на шину X подается импульс более высокого напряжения, под действием которого пережигается перемычка и нарушается электрическое соединение между катодом диода и шиной Y. Обычно для пережигания нихромовых перемычек необходимо пропустить ток 50—100, а кремниевых — примерно 20 мА.



Рискнок 3.4 — Запоминающие ячейки на основе резисторного ЗЭ:
а — с диодной; *б* — с транзисторной развязкой

Вследствие того, что диод является пассивным элементом, для получения высокого быстродействия формирователи возбуждения выходных шин выборки строки должны иметь малое выходное сопротивление, т.к. в момент подачи напряжения на шину происходит зарядка паразитных емкостей матрицы входным током. Использование в качестве элементов развязки транзисторов, включенных по схеме с общим коллектором (рис. 3.4, *б*), позволяет существенно снизить ток выборки для шин X, благодаря усилительным свойствам транзистора дешифратор при программировании может задавать в выбранную шину значительно меньший ток, чем необходимо для пережигания перемычек. Обычно запоминающие матрицы строят на основе n-p-n-транзис-торов, что позволяет достичь наивысшего быстродействия и наибольшей плотности упаковки на кристалле для биполярных транзисторов.

Ячейки памяти на основе диодного ЗЭ (рис. 3.5) состоят из двух встречновключенных p-n-переходов, при считывания состояния ЗЭ на шину X подается положительное напряжение, а с нагрузки, подключенной к шине Y, снимается выходной сигнал. В исходном состоянии элемент хранит 0, а его сопротивление очень велико. Для записи 1 к встречновключенной паре p-n-переходов прикладывается повышенное напряжение, при котором запертый переход пробивается и замыкается накоротко. Сравнительно перспективным считается выполнение ЗЭ на встречновключенных диодах Шотки (рис. 3.5, *а*), которые в настоящее время получают все большее распространение. В исходном состоянии встречновключенные диоды не проводят ток (состояние логического 0), но при программировании вследствие пробоя происходит закорачивание обратносмещенного диода (состояние логической 1).



а *б*

Рисунок 3.5 — Запоминающие ячейки на основе диодного ЗЭ:
а — на диодах Шотки; *б* — на транзисторе

Встречновключенные переходы часто создаются на основе транзистора с отключенной базой (рис. 3.5, *б*). Эмиттер транзистора соединяется с разрядной шиной. В режиме программирования при подаче достаточного потенциала на эмиттер (при заземлении коллектора) происходит необратимый пробой эмиттерного перехода и транзистор превращается в диод, образованный переходом коллектор-база.

На основе рассмотренных эффектов запоминающих ячеек созданы микросхемы и блоки ППЗУ ёмкостью более 64 Кбит и временем выборки 15—100 нс со встроенными схемами обрамления.

У ППЗУ есть некоторый недостаток, который заключается в том, что микросхемы невозможно подвергнуть полному контролю по записи на заводе-изготовителе, поэтому процент выхода годных микросхем по результатам программирования составляет 50—70 %.

3.1.3 Цилиндрические магнитные домены

Проведенный анализ базировался на предположении о плоскопараллельной форме доменов. Такие структуры наблюдаются в тонких пленках и пластинках. Однако в реальных ферромагнитных образцах нередки и другие виды доменных структур. В одноосных кристаллах часто наблюдаются так называемые лабиринтные доменные структуры. Их возникновение объясняется тем, что направление доменных границ в плоскости пластины ничем не фиксировано (в плоскости пластины нет анизотропии). Изгиб доменных границ может быть обусловлен малыми неоднородностями пленки, случайностью в момент зарождения доменной структуры или эффектами тепловой хаотизации. Такая структура остается выгодной и при помещении в малое внешнее магнитное поле, перпендикулярное поверхности пленки.

При увеличении магнитного поля в такой ситуации возникает очень интересное явление. Очевидно, что при увеличении поля растут домены, в

которых вектор параллелен вектору индукции магнитного поля B и, наоборот, уменьшается размер доменов, в которых M антипараллелен B. При этом данный полосовой домен распадается на отдельные цилиндрические домены кругового сечения (рис. 3.6). Благодаря магнитодипольному взаимодействию они отходят друг от друга и равномерно распределяются по всей поверхности пластины, образуя, как правило, правильную гексагональную решетку. Плотность доменов зависит от величины индукции B. Интересно отметить, что при уменьшении B решетка цилиндрических магнитных доменов (ЦМД) может сохраняться и в слабых полях, даже при B = 0.

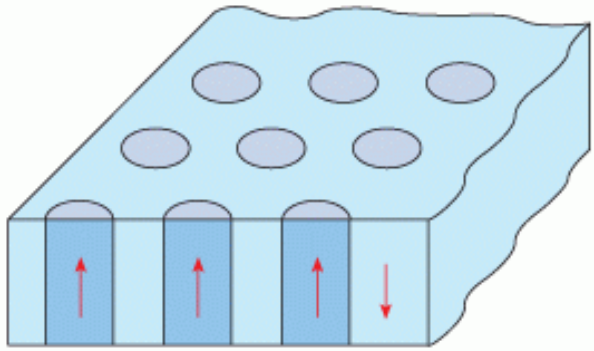


Рисунок 3.6 — Цилиндрические магнитные домены

ЦМД обладают интересными, присущими только им, свойствами. Если в пластинке с полосовой доменной структурой внутреннее магнитное поле должно быть равно нулю, то в образцах с ЦМД из-за наличия кривизны доменных границ это поле должно быть отлично от нуля. Иначе ЦМД не будут устойчивыми. Ситуация здесь аналогична поведению пузырька газа в жидкости. Для существования пузырька в жидкости необходимо, чтобы давление внутри пузырька отличалось от давления в жидкости. Также и в случае ЦМД: для их устойчивого существования необходимо наличие внутреннего магнитного поля, которое будет создавать дополнительное давление на искривленную доменную границу. Приведенная аналогия как раз объясняет английское название ЦМД — magnetic bubble (магнитный пузырек).

ЦМД в настоящее время применяются в устройствах памяти ЭВМ. Логическим элементом 1 в этих устройствах является сам ЦМД и элементом 0 — пространство между ЦМД.

Впервые ЦМД обнаружили группы чехословацких и голландских физиков в 1961 году. После открытия в 1965 году высокой подвижности ЦМД в ряде магнитных пленок (до 10—15 км/с) американский физик Э. Бобек выдвинул идею отнести ЦМД к числу перспективных кандидатов на роль носителей информации в запоминающих устройствах нового типа — без механических частей.

Почему выгодно использовать ЦМД в качестве носителей информации ЭВМ? Для этого надо вспомнить основные желаемые требования к носителям информации. Во-первых, необходимо, чтобы плотность записи информации была высока. Это позволяет сделать носители информации и сами ЭВМ как можно меньшими. Во-вторых, желательно, чтобы скорость записи и считывания информации была высокой. Это налагает очень жесткие требования к механической части запоминающих устройств.

Элементы памяти на основе ЦМД как раз позволяют решить указанные проблемы. Повышение плотности записи информации на основе ЦМД может быть достигнуто за счет уменьшения диаметра ЦМД. В настоящий момент диаметр ЦМД доведен до значений ~1 мкм, что позволяет создать устройства с плотностью записи 0,1 Гбит/см^2. Повышение скорости записи и считывания достигается, как указывалось выше, большой подвижностью ЦМД. А самое главное — в устройствах с ЦМД нет механических частей. Запись и считывание информации в них осуществляются за счет движения ЦМД по магнитной пленке. Для создания и перемещения ЦМД используются технологические схемы, например метод магнитных аппликаций и переменного вращающегося магнитного поля, проводников с током, локального разогрева пленки лазерным лучом и т.д. В основном распространены первые два метода.

3.1.4 Магнитные носители

Технология записи информации на магнитные носители появилась сравнительно недавно — примерно в середине 20-го века (40-е — 50-е годы). Но уже несколько десятилетий спустя — 60-е — 70-е годы — это технология стала очень распространенной во всём мире.

Очень давно появилась на свет первая грампластинка, которая использовалась в качестве носителя различных звуковых данных, — на неё записывали различные музыкальные мелодии, речь человека, песни.

Сама технология записи на пластинки была довольно простой. При помощи специального аппарата в специальном мягком материале, виниле, делались засечки, ямки, полоски. И из этого получалась пластинка, которую можно было прослушать при помощи специального аппарата — патефона или проигрывателя. Патефон состоял из механизма, вращающего пластинку вокруг своей оси, иглы и трубки.

Приводился в действие механизм, вращающий пластинку, и ставилась игла на пластинку. Игла плавно плыла по канавкам, прорубленным в пластинке, издавая при этом различные звуки — в зависимости от глубины канавки, её ширины, наклона и т.д., используя явление резонанса. А после труба, находившаяся около самой иголки, усиливала звук, «высекаемый» иголкой (рис. 3.7)

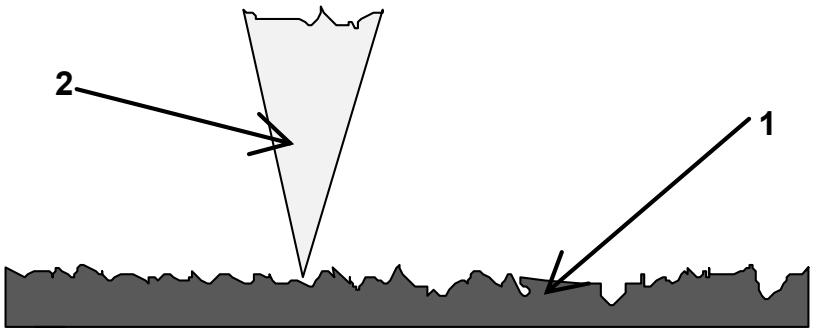


Рисунок 3.7 — Схема воспроизведения информации, записанной на виниловые носители: 1 — виниловая пластинка; 2 — игла

Почти такая же система и используется в современных (да и использовалась раньше тоже) устройствах считывания магнитной записи. Функции составных частей остались прежними, только поменялись сами составные части — вместо виниловых пластинок теперь используются ленты с напылённым на них сверху слоем магнитных частиц; а вместо иголки — специальное считывающее устройство. А трубка, усиливающая звук, исчезла совсем, и на её место пришли динамики, использующие уже более новую технологию воспроизведения и усиления звуковых колебаний. А в некоторых отраслях, в которых применяются магнитные носители (например, в компьютерах), пропала необходимость использования таких трубок.

Магнитная лента состоит из полоски плотного вещества, на которую напыляется слой ферромагнетиков. Именно на этот слой «запоминается» информация.

Процесс записи также похож на процесс записи на виниловые пластинки — при помощи магнитной индукционной головки вместо специального аппарата.

На головку подаётся ток, который приводит в действие магнит. Запись звука на плёнку происходит благодаря действию электромагнита на плёнку. Магнитное поле магнита меняется в такт со звуковыми колебаниями, и благодаря этому маленькие магнитные частички (домены) начинают менять своё местоположение на поверхности плёнки в определённом порядке, в зависимости от воздействия на них магнитного поля, создаваемого электромагнитом.

А при воспроизведении записи наблюдается процесс обратный записи: намагниченная лента возбуждает в магнитной головке электрические сигналы, которые после усиления поступают дальше в динамик (рис. 3.8)

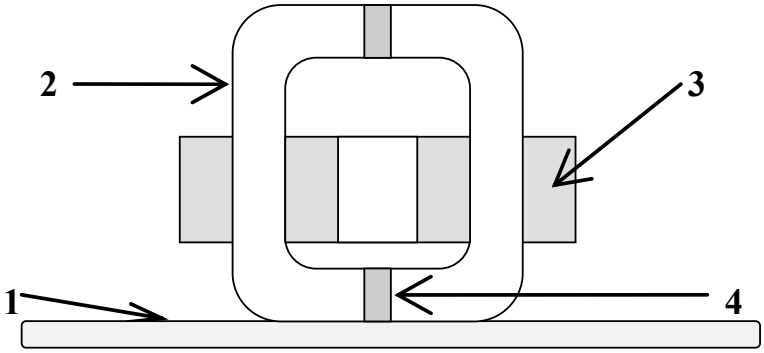


Рисунок 3.8 — Схема магнитной индукционной головки:
1 — магнитная лента; 2 — сердечник электромагнита; 3 — обмотка электромагнита; 4 — рабочий зазор

Данные, используемые в компьютерной технике, записываются на магнитные носители таким же образом, с той разницей, что для данных нужно меньше места на плёнке, чем для звука. Просто вся информация, записываемая на магнитный носитель в компьютерах, записывается в двоичной системе — если при чтении с носителя головка «чувствует» нахождение под собой домена, то это означает, что значение данной частички данных равно «1», если не «чувствует», то значение — «0». А дальше уже система компьютера преобразует данные, записанные в двоичной системе, в более понятную для человека систему.

Сейчас в мире присутствует множество различных типов магнитных носителей: дискеты для компьютеров, аудио- и видеокассеты, бобинные ленты, жёсткие диски внутри компьютеров и т.д.

Гибкие диски

В приводе флоппи-диска (гибкого диска, или просто дискеты) имеются два двигателя: один обеспечивает стабильную скорость вращения вставленной в накопитель дискеты, а второй перемещает головки записи-чтения. Скорость вращения первого двигателя зависит от типа дискеты и составляет от 300 до 360 об/мин. Двигатель для перемещения головок в этих приводах всегда шаговый. С его помощью головки перемещаются по радиусу от края диска к его центру дискретными интервалами. В отличие от привода винчестера головки в данном устройстве не «парят» над поверхностью флоппи-диска, а касаются ее.

Для подключения разных типов дисководов предназначены обычно комбинированные кабели с четырьмя разъемами, включенными попарно.

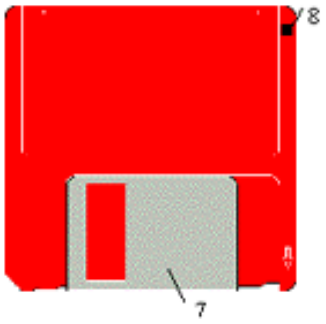
Некоторые BIOS компьютеров позволяют программно изменять назначение физического адреса: «первый» (A:) и «второй» (B:) привод. В отличие от винчестеров для флоппи-дисководов порядок накопителя (A: или B:) определяется именно положением устройства на кабеле.

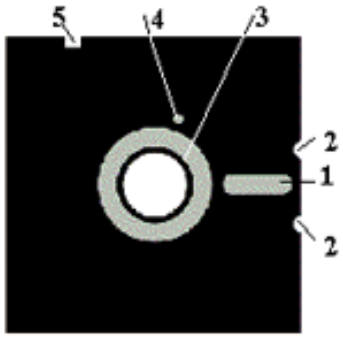
Для каждого из типоразмеров дискет (5,25 или 3,5 дюйма) существуют свои специальные приводы соответствующего форм-фактора.

Дискеты каждого типоразмера (5,25 и 3,5 дюйма) бывают обычно двусторонними (Double Sided, DS), односторонние давно стали анахронизмом. Плотность записи может быть различной: одинарной (Single Density, SD), двойной (Double Density, DD) и высокой (High Density, HD). Поскольку об одинарной плотности уже мало кто вспоминает, такую классификацию обычно упрощают, говоря только о двусторонних дискетах двойной плотности (DS/DD, емкость 360 или 720 Кбайт) и двусторонних дискетах высокой плотности (DS/HD, емкость 1,2, 1,44 или 2,88 Мбайта). Плотность записи определяется величиной зазора между диском и магнитной головкой, а от стабильности зазора зависит качество записи (считывания). Для повышения плотности записи необходимо уменьшить зазор, однако при этом значительно повышаются требования к рабочей поверхности дисков.

В качестве материала для изготовления магнитных дисков обычно применяют алюминиевый сплав Д16МП (МП — магнитная память). Этот сплав немагнитный, мягкий, достаточно прочный, хорошо обрабатывается.

Дискета представляет собой слой магнитно-мягкого материала, нанесенный на специальную подложку, выполненную из полимерного немагнитного пластического материала, степень жесткости которого может быть различна в зависимости от реализации. Носитель помещается в бумажный, пластмассовый или другой кожух-корпус. В настоящее время используются только двусторонние носители, следовательно, покрытие нанесено с обеих сторон дискеты и чтение/запись производится с обеих сторон. Дискеты различного диаметра, как правило, имеют разные оформления корпуса. Так, гибкие диски диаметром 5.25 дюйма помещаются в бумажный кожух, а 3.14 — в пластмассовый. Дискета в кожухе свободно вращается приводом устройства — дисковода через окно центрального захвата, что обеспечивает прохождение площади дорожки под устройством чтения/записи, называемым головкой чтения/записи.





На кожухе дискеты имеются соответственно отверстия: центрального захвата (3), позиционирования головки (1), физической защиты от записи (5, 8), направляющие отверстия и пазы (2), автоопределения типа магнитного покрытия (9), определения полного оборота носителя (4). Отверстие

для позиционирования магнитных головок чтения/записи у 3.14-дюймовых носителей закрыто металлической задвижкой (7), а отверстие для центрального захвата и вращения на шпинделе привода вращения диска, в отличие от носителя диаметром 5.25 дюймов, находится только с нижней стороны дискеты. Каждый сменный дисковый магнитный носитель перед использованием в какой-либо операционной системе необходимо подготовить к приему данных. Такая операция называется форматированием. Форматирование дискет производится при помощи специального программного обеспечения — программ форматирования дисков и, как правило, специфично для каждой операционной системы.

В зависимости от типа носителя, в соответствии с качеством магнитного покрытия, возможностями операционной системы и устройств дискеты можно форматировать для записи на них информации различного максимального объема, что достигается заданием таких параметров форматирования, как число дорожек и секторов. Как правило, производителями дискет указывается параметр, называемый числом точек на дюйм носителя — Track per inch (TPI). Данный параметр показывает, какую максимальную плотность размещения областей независимой намагниченности может иметь носитель. В соответствии с производственными характеристиками диска необходимо форматировать носитель только в рамках его физических возможностей, иначе риск потери данных после операции записи неограниченно возрастает.

Дисковод представляет собой устройство чтения/записи с/на носитель — дискету. Каждый тип носителя (дискет), как правило, требует собственного устройства — для чтения 5.25- и 3.14-дюй-мовых дискет, хотя выпускаются и смешанные дисководы, соединяющие в себе устройства для чтения 3.14- и 5.25-дюймовых дискет. Дисководы, как правило, располагаются внутри системного блока, однако выпускаются и внешние варианты. Снаружи системного блока находится передняя панель дисковода, на которой располагаются управляющие элементы — ручка или кнопка фиксации/извлечения дискеты внутри дисковода, отверстие для помещения/извлечения дискеты, индикатор обращения к устройству, светящийся во время операций обращения к дисководу. Внутри дисковод состоит из двигателя, системы управления вращением носителя, двигателя, системы управления позиционированием головок чтения/записи, схем формирования и преобразования сигналов и др. электронных устройств. Дисководы подключаются к другим схемам компьютера посредством интерфейсного кабеля — шлейфа. На концах и/или по длине шлейфа находятся разъемы, один из которых служит для соединения шлейфа с дисководом или дисководами, другой — с интерфейсом дискового устройства, находящимся на плате контроллера (интерфейсной карте, плате адаптера) дисковых устройств или на материнской плате. Дисковод также нуждается в подключении питающего напряжения при помощи кабеля питания.

В настоящий момент технологии хранения и чтения/записи информации на обычную дискету дают невысокие скорости обмена и позволяют добиться плотности записи для объема информации до 2-х мегабайт. Такой объем и быстродействие считаются малыми, и поэтому дискеты используют лишь как средство транспортировки и архивного хранения небольших объемов информации. Надежность дискет также оставляет желать лучшего. Они подвержены вредным воздействиям температурных, гидрометрических, магнитных, механических и др. факторов. Поэтому с дискетами следует обращаться аккуратно.

Во избежание потери данных или повреждения носителя недопустимо хранение дискет в местах, подверженных воздействию магнитных полей, влаги, сильных механических воздействий, обильного количества пыли, резких температурных перепадов. Необходимо осторожно вставлять и извлекать дискету из дисковода только после того, как индикатор обращения к диску погаснет. В зависимости от интенсивности использования дискеты ее необходимо проверять на предмет целостности и правильности логической и физической структуры при помощи специального программного обеспечения с различной частотой, но не реже одного раза в два месяца. Также необходимо производить чистку головок чтения/записи дисковода при помощи специальной чистящей дискеты и очистителя. Срок службы носителя зависит не только от способа его эксплуатации, но и от его исходного качества. Дискеты высокого качества известных крупных производителей способны форматироваться на максимальные объемы и выдерживают при эксплуатации до 70 млн проходов головки чтения/записи по дорожке, что, практически, означает срок интенсивной эксплуатации до 20 лет. Дискеты безымянных производителей и просто плохого качества, как правило, подвержены таким вредным процессам, как высыпание частичек магнитного покрытия и размагничиваемость. Не следует экономить на носителях информации, если она вам дорога. На практике, нужно стараться использовать только высококачественные дискеты известных производителей.

3.1.5 Магнитно-оптический носитель

Магнитооптика — раздел физики, в котором изучаются изменения оптических свойств сред под действием магнитного поля и обусловленные этим изменения взаимодействия оптического излучения с помещенным в поле веществом.

Очевидно, что если в названии упоминаются магниты и оптика, то оба этих явления используются в устройстве.

Так и есть. МО-диск представляет собой поликарбонатную подложку (часто его также называют слоем) толщиной 1,2 мм, на которую нанесено несколько тонкопленочных слоев, где заключается магнитная часть технологии, а оптическая представлена считывающим лазером.

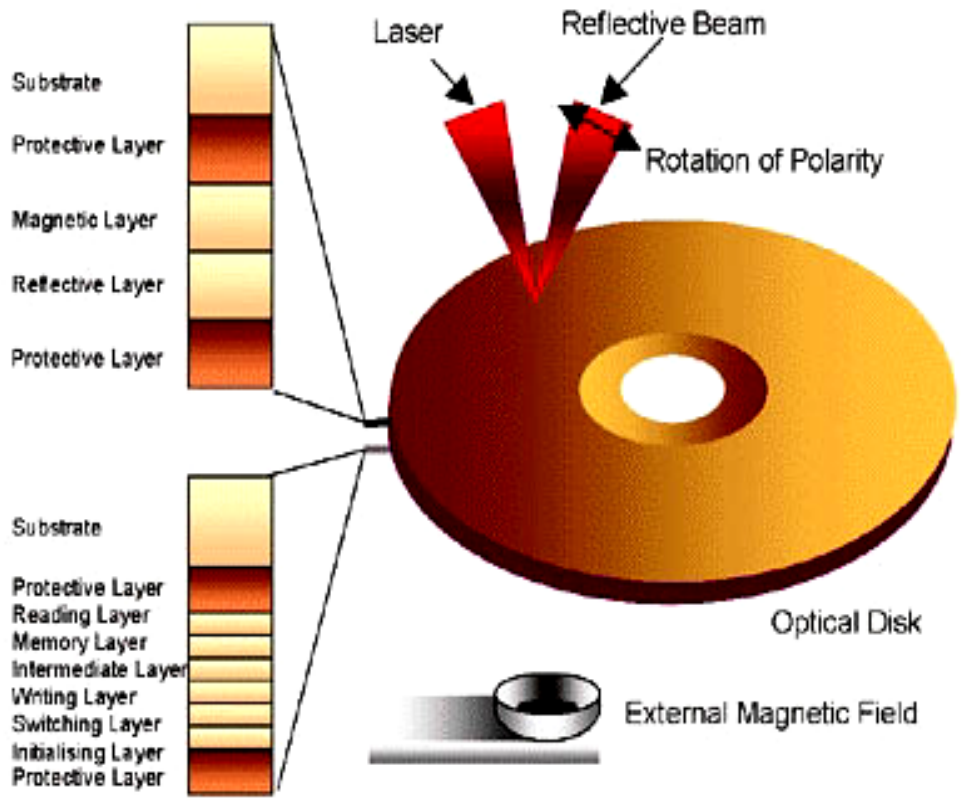


Рисунок 3.9 — Структура магнитно-оптического носителя

Теперь вернемся к устройству диска — самые главные слои это защитный слой, предохраняющий поверхность диска от повреждений, отражающий слой, необходимый для корректной работы лазера; диэлектрические слои, выполняющие две функции — теплоизолируют магнитный слой (чтобы эффективнее использовать энергию лазера при записи) и увеличивают эффект поляризации при чтении; магнитный слой — собственно хранитель информации. Сам MO-диск помещается в пластиковую коробку со «шторкой» и окошечком защиты от записи, размер составляет 1,8 дискеты.

Видовое разнообразие

Размер 5,25

Максимальная емкость носителя — 9,1 Гб, за исключением HDD более емких устройств найти невозможно (устройства Orb представляют собой чисто теоретический интерес). По скоростным параметрам записывающие устройства на базе DVD уступают магнитооптике не

только по быстродействию, но и по надежности хранения данных. Магнитооптические носители выдерживают огромное количество циклов перезаписи, не чувствительны к внешним магнитным полям и радиации, гарантируют сохранность записанной информации в течение полусотни лет. Уникальный набор характеристик обеспечил магнитооптической технологии применение в High-End-устройствах записи с повышенными требованиями к объемам и надежности хранения данных (неслучайно библиотека конгресса США оборудована магнитооптическими библиотеками).

Производят магнитооптические дисководы фирмы Sony и Maxoptix, а магнитооптические библиотеки — Hewlett-Packard, Maxoptix, Plasmon и UNIES. В библиотеках Hewlett-Packard используются магнитооптические дисководы Sony емкостью 5,2 или 9,1 Гб, в библиотеках Maxoptix — дисководы Maxoptix (5,2 Гб) или Sony (9,1 Гб).

Судя по всему, выпущенный фирмой Sony в конце 2000 года дисковод SMO-F561, поддерживающий 5,25-дюймовые диски емкостью 9,1 Гб, является последним устройством, основанным на традиционной магнитооптической технологии (MO). Дальнейшее развитие фирма Sony связывает с новой технологией оптической записи UDO (*Ultra Density Optical*). Конкурентом технологии является продукция Fujitsu под названием GigaMO.

Технология UDO базируется на новом коротковолновом лазере с длиной волны 405 нм, применение которого позволяет существенно увеличить плотность размещения дорожек записи и плотность записи в дорожке. Процесс записи основан не на магнитооптической технологии, а на технологии изменения фазы. Формат UDO предполагает начальную емкость 5,25 диска в 40 Гб (по 20 Гб на сторону). В дальнейшем емкость диска может быть доведена до 60 и даже до 120 Гб.

Фирма Sony была не первой, предложившей перейти к форматам со сверхвысокими плотностями записи. Почти двумя годами ранее это сделала фирма Maxoptix, разработавшая формат OSD (*Optical Super Density*), основанный на магнитооптической технологии. Этот формат также предусматривает начальную емкость диска 40 Гб с последующем увеличением. Для повышения плотности записи и быстродействия используется комбинация из нескольких приемов. В отличие от традиционной магнитооптической технологии рабочий слой размещается практически на поверхности диска (OCIR — *OverCoat Incident Recording*). При этом защитный слой сохраняется, но становится намного тоньше, так что головки могут приблизиться к рабочему слою почти вплотную. Считывающая оптическая головка имеет усовершенствованную линзу (*Recessed Objective Lense*) и располагается очень близко от поверхности, благодаря чему достигается минимальный размер светового пятна. Запись производится с помощью двух головок. Оптическая осуществляет нагрев, а магнитная изменяет направление магнитного поля (*Magnetic Field Modulation*). Обе стороны диска записываются одновременно (*Surface*

Array Recording), благодаря чему скорость записи и чтения данных удваивается.

Фирма Maxoptix уже неоднократно проводила технологическую демонстрацию OSD-технологии и практически готова к выпуску новой продукции. Однако будущее этой технологии находится под вопросом, так как даже двух альтернативных технологий для относительно узкого сектора рынка слишком много.

Размер 3,5

Магнитооптика формата 3,5, в отличие от магнитооптики формата 5,25, с самого начала была ориентирована на массовый рынок. Благодаря компактности, высокому быстродействию и надежности позиции 3,5 дисководов довольно прочны.

Хотя 3,5 магнитооптические дисководы предлагаются под разными торговыми марками, производятся они в настоящее время единственной фирмой — Fujitsu, которая была родоначальницей этого формата и внесла наибольший вклад в его развитие. Последнее технологическое достижение — формат высокоплотной записи GigaMO — является совместной разработкой Fujitsu и Sony. В GigaMO емкость носителей составляет 1,3 Гб и 2,3 Гб. Оба формата GigaMO предусматривают полную обратную совместимость устройств с носителями предыдущих поколений (128—640 Мб).

В настоящее время Fujitsu производит дисководы как для 1,3 Гб носителей, так и для 640 Мб носителей.

Модельный ряд включает несколько модификаций по производительности (разные значения скорости вращения диска и среднего времени доступа), используются все распространенные типы интерфейсов (ATAPI, SCSI, LPT, USB 1.1 & 2.0, IEEE 1394).

Последней новинкой являются накопители DynaMO 2300U2 и DynaMO 1300U2 (более полную информацию можно узнать здесь), использующие сменные носители емкостью 2,3 и 1,3 GB соответственно. Учитывая, что ориентировочная стоимость 2,3-гигабайтных MO-дисков составляет $30 — это достаточно экономичный и надежный способ решения таких задач, как архивация или резервное копирование больших объемов информации. Устройства U2 сохраняют полную совместимость с ранними версиями MO дисков Fujitsu емкостью 128, 230, 540 и 640 MB, а использование технологии Hi-Speed USB 2.0 позволяет эксплуатировать устройство с максимальной быстротой.

Несмотря на лидерство и некоторый монополизм Fujitsu, последнее время активно действует на рынке не менее известный гигант Olympus. Новая линейка продуктов вполне конкурирует с основным производителем

TURBO MO с интерфейсами Ultra SCSI и USB 2,0 на 1,3 Гб MO133S1S и 640 Мб модели MO646S1S.

Таблица 3.1 — Характеристика моделей MO

Характе-ристики	MCB3064AP	MCE3064SS	MCD1330AP	MCD2300 AP	DynaMO2300U2	DynaMO1300U2
Емкость (Мб)	640	640	1,3 и 640	2,3	2,3	1,3
Время поиска (мс)	28	23	28	24	20	19
Скорость вращения (об/с)	4300	4558	3500 и 4500	4000	5400	4500
Скорость передачи (Мб/с)	4,33	4,96	5,92	6	8	6
Скорость чтения (Мб/с)	3,67	3,88	4,5	5	5,5	5
Скорость записи (Мб/с)	1,23	1,29	1,5	2	2,5	2,5
Буфер (Мб)	0,512	2	2	2	4	2

Первая модель работает с MO-дисками емкостью 1,3 Гб и меньше, вторая — с дисками емкостью 640 Мб и меньше. Скорость вращения шпинделя у обеих моделей — до 6000 об/мин (при работе с 1,3 Гб дисками — только 3670 об/мин), размер буфера — 2 Мб.

3.1.6 Оптические носители

DVD

DVD-стандарт был реализован с учетом накопленного опыта по производству и распространению компакт-дисков и CD-устройств, требований и рекомендаций производителей компьютерной и киноиндустрии, а также предварительных разработок различных компаний. Новый стандарт базируется на следующих основных принципах:

- большая емкость и возможность ее дальнейшего наращивания;
- обратная совместимость с существующими CD;
- совместимость с будущими записываемыми DVD-дисками;

– единая файловая система для всех приложений;

– единый интерактивный стандарт для компьютера и телевидения;

– надежность хранения данных и их последующего считывания;

– высокая производительность при записи и считывании данных как для последовательного, так и для произвольного доступа к данным;

– отсутствие вспомогательных конструкций типа картриджей и кэдди;

– доступная цена.

Внешне конструкция DVD аналогична устройству традиционного компакт-диска — с теми же геометрическими размерами (диаметр — 120 мм, толщина — 1,2 мм), но содержательно она значительно сложнее. Для увеличения объема данных при сохранении тех же геометрических размеров диска, что и CD, были предприняты следующие шаги:

– уменьшение размеров углублений (питов) на DVD до 0,4 мкм;

– уменьшение расстояния между соседними дорожками (треками) до 0,74 мкм;

– размещение несущих информацию слоев в несколько этажей (до 8 пар, и это еще не предел).

DVD может быть как односторонним, так и двухсторонним. Конструктивно двухсторонний диск представляет собой два склеенных нерабочими поверхностями диска толщиной 0,6 мм каждый (модель, предложенная компанией Toshiba). Спецификации DVD-стандарта предусматривают четыре конструктивно различных типа дисков с разной информационной емкостью:

– односторонний однослойный диск (4,7 Гбайт, видеоресурс — 133 мин);

– односторонний двухслойный диск (8,5 Гбайт, видеоресурс — 240 мин);

– двухсторонний однослойный диск (9,4 Гбайт, видеоресурс — 266 мин);

– двухсторонний двухслойный диск (17 Гбайт, видеоресурс — 481 мин).

Таким образом, емкость одностороннего однослойного диска в семь раз, а двухстороннего двухслойного — в двадцать шесть раз превышает емкость стандартного компакт-диска. Предполагается, что первый тип дисков найдет широкое распространение для большинства компьютерных приложений, где емкости 4,7 Гбайт вполне достаточно, а более емкие диски, видимо, будут востребованы киноиндустрией.

Увеличение плотности данных стало возможным благодаря созданию более совершенных источников лазерного излучения и системы обнаружения и коррекции ошибок. Для считывания DVD используется луч красного спектра с возможностью двойного фокусирования с длиной волны 650 нм или 635 нм, в зависимости от толщины считываемого диска. Привод DVD сам определяет, какой тип диска используется, и

автоматически поворачивает линзу в положение нужной фокусировки луча.

При такой плотности записи любая внутренняя неоднородность может сделать диск непригодным к использованию. Поэтому с помощью технологии компании Sony была модернизирована и стандартизирована схема цифровой модуляции и коррекции ошибок RS-PC (Reed Solomon Product Code), которая уменьшила вероятность их появления на порядок по сравнению с компакт-диском. Кроме того, DVD, как и компакт-диск, стоек и малочувствителен к пыли, царапинам и прикосновениям пальцев.

Система самоуничтожения для DVD-дисков

Не совсем понятную новость только что сообщило новостное агентство «Reuters». Flexplay разработал интересную систему борьбы с недобросовестными обладателями DVD-дисков. Как известно, по законодательству, прокат DVD-дисков (hiring), и не только их, запрещен. Технология уже будет внедрена в жизнь в августе этого года компанией Disney.

Диски прекратят функционировать, когда процесс, названный Ez-d, сделает их неработоспособными. Как только диск вытаскивается из упаковки, он может быть использован только в течение приблизительно 48 часов. Взаимодействие поверхности диска с кислородом через данный промежуток времени создает особый слой на поверхности диска, из-за которого процесс чтения становится невозможным. Однако при наличии известного ПО можно просто скопировать содержимое диска на HDD за время, пока он работает как обычный носитель. С другой стороны — куда деваться честным покупателям?

Divx

Компания Digital Video Express разработала новый формат Divx-диска для однократной записи кинофильмов. Разработка этого формата связана с организацией системы временного видеопроката, когда купив диск, не придется возвращать его назад. Его можно будет воспроизводить только на Divx-проигрывателях в течение двух суток с момента его первого воспроизведения. О своей поддержке этого формата заявили такие крупные голливудские компании, как Disney, Dream-Works, Paramount, Universal и другие.

Этот диск не совместим с домашними DVD-проигрывате-лями, подключаемыми к телевизору. Divx — это название системы, установленной непосредственно в проигрывателе, которая позволяет потребителям в течение двух дней пользоваться правом на прокат видеофильма независимо от даты покупки диска. Идея Divx состоит в том, что она обеспечивает нарушение записи на диске. Право проката видеофильма на новый срок можно приобрести через модемную линию связи, подключенную к проигрывателю для обмена информацией с сервисным центром Divx Central и отслеживания счетчика. Внедрение

данного формата в нашей стране не представляется возможным ввиду того, что для просмотра Divx-дисков требуется дорогостоящее оборудование, постоянная телефонная связь с центром, да и цена диска предположительно составит около 6 долл.

FMD ROM — накопители третьего тысячелетия

По каким же параметрам FMD ROM превосходит DVD?

Первый параметр — соотношение размер/емкость. Тут «fluorescent multilayer disk» вне конкуренции. Разработчики заявляют, что уже сейчас первые прототипы способны вмещать при размере диска 12 см в диаметре, то есть на стандартном 5 дюймовом носителе, до 140Гб. Это при десяти слоях. А в ближайших планах компании C3D есть желание как минимум удесятерить число слоев. При этом становится вполне реальной возможность создания сменных носителей информации емкостью в десятки терабайт. Та емкость, которую на сегодняшний день можно получить лишь при использовании громадных дисковых массивов, занимающих подчас целые шкафы и даже комнаты, будет обеспечиваться компактным диском, который с легкостью умещается в кармане!

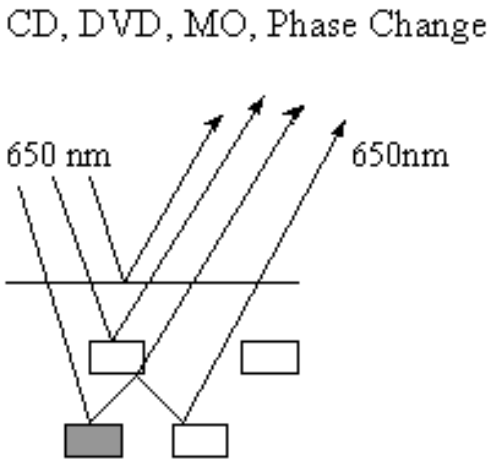


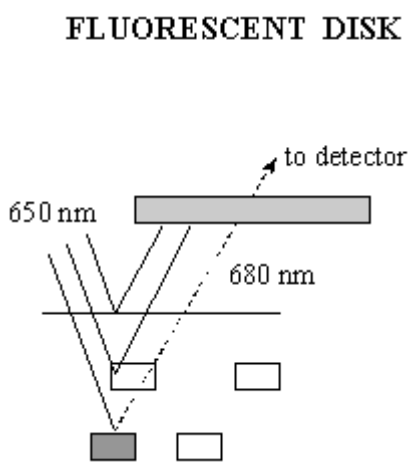
Рисунок 3.10 — Вид FMD ROM

Насчет скорости доступа еще очень мало данных. Разработчики обещают, что этот параметр будет намного выше, нежели у DVD. Хотелось бы верить, ведь иначе, с существующими скоростями, при работе с терабайтными массивами информации даже простые операции, например перечитка диска, могут затянуться на несколько часов. Новые гигантские объемы требуют и соответствующих скоростей доступа. Что же касается соотношения емкость/стоимость носителя, то и тут FMD ROM не имеет себе равных. Ведь он представляет собой практически кусок пластмассы, вернее полимерную матрицу с фотохромным веществом, но по стоимости это просто пластиковый диск. И никаких затрат по созданию дорогостоящих полупрозрачных слоев, как в DVD. Собственно, и никаких слоев в привычном смысле этого слова нет, но об этом в следующей главе.

О принципах функционирования FMD ROM

Внешний вид FMD ROM. Как вы видите, диск совершенно прозрачный, хотя и имеет формат обычного CD- или DVD-диска. В отличие от обычного CD-ROM, в котором отражающий алюминиевый слой нанесен на выдавленную подложку из полимера, из-за чего он, собственно, и непрозрачен, диск FMD ROM монолитен и при этом разделен по вертикали на некоторые условные области, названные разработчиками «слоями» (layer). Эти «слои» не являются слоями в привычном смысле, это скорее параметр форматирования диска, ближайший аналог — это сектора и дорожки для магнитных носителей. Толщина этих слоев строго фиксирована, и это не случайно. Чтобы понять, почему разработчики выбрали именно эту толщину каждого из слоев, надо рассмотреть принципы записи/считывания информации на FMD ROM.

В оптических носителях (CD, DVD, магнитооптика) во время чтения луч полупроводникового лазера отражается от слоя с записанной информацией.



Отраженный луч затем фиксируется детектором — приемником. Грубо говоря, считывание идет по принципу: попал или не попал луч в приемник. Максимальная удельная емкость диска определяется размером светового пятна от лазера, которое, в свою очередь, зависит от длины волны (у красных лазеров — 650 нм). Можно использовать два слоя, причем сделать один из слоев прозрачным для излучения с определенной длиной волны, как это реализовано в DVD. Но два слоя — это предел, больше сделать очень сложно, так как нужны очень точные фокусирующие системы, которые будут работать только в лабораторных условиях. Разумеется, массовое производство таких систем является невероятно дорогим и нерентабельным. Да и вообще, технология отражающих слоев подошла к своему пределу развития.

Но вот создатели технологии многослойных дисков (компания C3D) нашли способ и обошли проблему множественной интерференции между слоями и потери самого луча в многослойных дисках. И технологически это выглядит очень красиво и остроумно.

Разработчиками FMD было предложено следующее решение: материал, содержащий записанную информацию, не отражает, как подложка в DVD или CD, а излучает! Использовано явление флуоресценции, то есть при освещении активирующим излучением (в данном случае полупроводниковым лазером с определенной длиной волны) вещество начинает излучать, сдвигая спектр падающего на него излучения в сторону красного цвета на определенную величину. Причем

величина сдвига зависит от толщины слоя. Таким образом, выбрав такую толщину слоя, чтобы спектр отраженного света получался смещенным относительно длины волны излучающего лазера на строго определенную величину, например на 30 или 50 нм, можно с высокой достоверностью записывать информацию вглубь диска и впоследствии считывать ее без потери данных.

Для FMD ROM разработчиками так же предложено название «трехмерный диск», и в данном случае это вполне оправдано.

Таким образом, плотность записи будет зависеть и от чувствительности регистрирующего детектора. Чем меньше то дополнительное излучение флюоресцирующего вещества, добавляющееся к частоте рабочего лазера, который удастся зафиксировать, тем большее число слоев можно вместить в один диск.

Излученный свет от флуоресцентного слоя некогерентен и хорошо контрастирует с отраженным светом лазера, что является дополнительной гарантией надежности считывания, ведь без отражений все равно не обойтись, они будут происходить от поверхности диска и других записанных слоев. Качественное ухудшение сигнала в обычных (отражающих) многослойных дисках нарастает с увеличением числа слоев, но вот в случае с флуоресцентными дисками это ухудшение происходит гораздо медленнее. По заявлению разработчиков FMD ROM, даже при количестве слоев больше сотни не будет происходить сильного искажения полезного сигнала. Используя синий лазер (480 нм), можно увеличить плотность записи до десятков терабайт на один FM-диск. Вполне возможно создание диска с 1000 слоями — это уже субмолекулярные размеры. Теоретически возможно создание пятна размером в несколько молекул, проблема лишь в том, как зафиксировать столь малое флуоресцентное излучение.

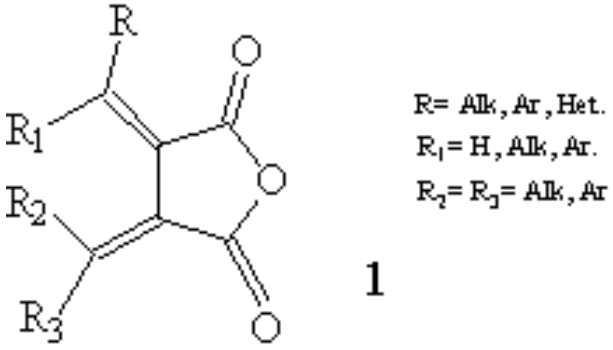
Одна из главных особенностей этой разработки — возможность параллельного чтения слоев (т.е. последовательность бит будет записана не по «дорожкам», а по слоям) — скорость выборки данных в этом случае должна быть очень высокой. Вот уж действительно «3-мерный диск».

Принцип записи на FMD ROM основан на явлении фотохромизма. Фотохромизм — это свойство некоторых веществ под действием активирующего излучения обратимо переходить из одного состояния в другое, при этом изменяя свои физические свойства (например, такие, как цвет, появление/исчезновение флюоресценции и т.д.). Материал, из которого состоит FMD ROM, содержит специальную фотохромную субстанцию, которая циклизуется под воздействием лазерного луча определенной длины волны, превращаясь в необходимый устойчивый флуоресцент. Обратная реакция рециклизации, приводящая к исчезновению флуоресцентных свойств (операция стирания), происходит под действием лазера с другой длиной волны. Стирающая частота лазера выбирается с таким расчетом, чтобы она не встречалась в повседневной жизни во избежание потери данных. Ну, и естественно, читающий лазер ни

в коем случае не должен вносить изменения в данные, хранящиеся на диске.

Наиболее ценными фотохромными свойствами обладают соединения под названием «фульгиды», поэтому можно предположить, что используемый в FMD ROM фотохром принадлежит именно к этому классу.

Общая формула фульгидов:



1

Вообще идея использования фотохромов в качестве носителей информации не нова. Ей примерно тридцать лет. И лишь теперь эта идея была реализована на практике.

Технология Blu-Ray — преемник DVD

Так уж повелось, что эволюция в области компьютерных технологий происходит быстрее, чем в остальных технических отраслях. И с течением времени период, за который мощность компьютера удваивается, становится все меньше и меньше. До 1 ГГц процессоры шли 22 года, а до 2 ГГц — всего лишь полтора. Объем винчестера растет как на дрожжах, 160—180 гигабайт — это уже обыденность, а ведь совсем недавно для достижения таких объемов конструировались целые RAID-массивы из десятков жестких дисков. Миниатюризация, увеличение быстродействия, скорости передачи данных, увеличение плотности записи — вот что мы слышим каждый день, вот что появляется каждый день на новостных страницах Интернета. Объемы, скорости, частоты — все это удваивается, учетверяется, удесятеряется, и все это никого уже не удивляет. Но есть область, в которой на фоне успехов на других фронтах до последнего времени царил, казалось бы, полный застой. Это область сменных носителей информации.

Действительно, в этой области, как и двадцать лет назад, продолжает лидировать старичок CD-ROM. Правда, за эти годы он подрос с 650 Мб до 700 Мб, а благодаря стараниям TDK — местами даже и до 800 Мб, но, увы, в наш насыщенный информацией век такие объемы становятся явно недостаточными. Столь долгой жизнью CD-ROM обязан не в последнюю очередь форматам сжатия звука (MP3) и видео (MPEG4, DivX), благодаря которым в столь мизерный объем стало возможным запихивать огромные массивы музыки и целые фильмы. Конечно, качество страдает, но народ у

нас невзыскательный, и в конечном итоге все равно выходит качественнее и долговечнее, чем переписанные аудио- и видеокассеты.

В последнее время более взыскательная публика открыла для себя DVD (digital versatile disc). Именно в последнее время, хотя этому формату насчитывается уже 8 лет. Причин столь медленного продвижения множество. Вначале на рынке царил дикий разнобой форматов DVD. Если в случае с CD компания-изобретатель SONY четко дала спецификации данного устройства, которых придерживались все производители, то в случае с DVD вышло с точностью наоборот — каждый из производителей предлагал свою версию DVD-привода, с разным максимальным объемом носителя, с разной механикой и даже с разной длиной волны читающего лазера, что в начале частенько приводило к ситуации, когда на DVD-приводе из 10 дисков читались только один-два. На следующем этапе распространения DVD сдерживающим фактором стала, прямо скажем, нелепая попытка обуздать пиратство с помощью введения зон. За несколько лет экспансии DVD-диски из разных зон настолько перемешались между собой, что стало проблематичным найти диски именно под свой дисковод. К тому же мгновенно в сети Интернет появились взломанные мультизональные прошивки DVD-приводов и специальные программы, обнуляющие счетчики проигранных дисков. Опытные пользователи решили проблемы с зонами в обход производителей, ну а неопытных вся эта канитель только отпугнула от приобретения DVD-привода. Не в последнюю очередь в медленном продвижении DVD сыграли высокая стоимость дисков и их относительная редкость. Например, в России в свободной продаже DVD-диски массово появились совсем недавно.

Но постепенно все стало налаживаться. Производители самостоятельно стали убирать ограничения на зоны из своих приводов. Так, например, программа PowerDVD сообщила об одном недавно вышедшем DVD-приводе NEC, что «данное устройство имеет номера зон 1, 2, 3, 4, 5». Приводы стали совместимы, появилось что-то похожее на единый стандарт. DVD-фильмы появились в широкой продаже, их стали даже дублировать на русском языке. Цена на диски упала. Появились пиратские диски, которые составили ценовую конкуренцию (увы, только ценовую, о качестве говорить не приходится) лицензионным дискам. Пользователи, обнаружив, что DVD-приводы не намного дороже CD-ROM, стали постепенно затовариваться данными устройствами. Появились пишущие DVD-приводы, правда, за бешенные деньги и сразу в лице трех несовместимых между собой форматов: DVD-RAM (Panasonic), DVD-RW (Pioneer) и DVD+RW (Philips). Наконец-то для DVD замаячило светлое будущее. Но это будущее было стремительно перечеркнуто в феврале 2002 года синим лучом — технологией Blu-Ray Disc.

Япония, Токио, 19 февраля 2002. Представители девяти лидирующих высокотехнологических компаний Sony, Matsushita (Panasonic), Samsung, LG, Philips, Thomson, Hitachi, Sharp и Pioneer на совместной пресс-

конференции объявили о создании и продвижении нового формата оптических дисков большой емкости под названием Blu-Ray Disc, этим самым, возможно, подписав смертный приговор DVD. Согласно объявленной спецификации Blu-Ray Disc — перезаписываемый диск следующего поколения со стандартным CD/DVD размером 12 см с максимальной емкостью записи на один слой и одну сторону до 27 Гб.

Собственно, назвать Blu-Ray принципиально новым форматом нельзя — это скорее эволюция формата DVD. Как следует из названия, в Blu-Ray для записи и воспроизведения диска вместо красного лазера, который используется в DVD и CD-ROM, применен синий лазер (blue-violet laser). У синего лазера длина волны составляет 405 нанометров, что значительно меньше длины волны красного лазера (650 нм). Меньшая длина волны — соответственно меньшая интерференция отраженного луча, соответственно можно сделать толщину дорожки данных тоньше, что приводит к значительному увеличению емкости носителя. Толщина дорожки у Blu-Ray-диска в два раза меньше, чем у DVD. Единственно, что внушает опасение, — тот факт, что энергетика синего лазера выше, чем у красного, что должно приводить к значительному разогреву поверхности диска. По-видимому, Blu-Ray-приводы потребуют мощного охлаждения.

Покрытие Blu-Ray, на которое записываются данные (optical transmittance protection layer), очень тонкое — 0.1 мм. Из этого факта можно сделать 3 вывода. Первое — чем тоньше слой, тем меньше рассеяние отраженного луча и больше данных можно вместить на квадратный дюйм, то есть тонкий слой — это необходимость для достижения большой емкости диска. Второе — настолько тонкий слой позволит без проблем сделать диск многослойным (по крайней мере, двухслойным, как DVD), так как уменьшается рефракция луча, отраженного от более глубокого слоя. Третье — настолько тонкий слой легко повредить, следовательно, Blu-Ray Disc потребует защиты, то есть будет упакован в пластиковую оболочку, наподобие MiniDisk от Sony. Последний факт, к сожалению, говорит о том, что цены на Blu-Ray-приводы, возможно, будут существенно выше, чем на DVD, так как если бы Blu-Ray Disc оставался бы диском без упаковки, то производители смогли бы использовать корпуса и механику от DVD-приводов без переделки, сменив лишь лазер и декодирующую микросхему, а так придется начинать практически с нуля. Возможен компромиссный вариант, когда односторонние диски относительно малой емкости (23—27 ГБ) будут производиться без упаковки и иметь соответствующие приводы, мало отличающиеся от DVD-приводов по внешнему виду и по цене; такие объемы для домашних мультимедийных компьютеров на первое время более чем достаточны, по крайней мере, объем Blu-Ray-диска в разы превосходит DVD, а для пользователей весьма важна цена. Потребители голосуют рублем, неважно зеленый он или нет, соответственно чем меньше будет начальная стоимость Blu-Ray для домашнего и мультимедийного сектора, тем быстрее он наберет популярность. Также

диски этого формата будут использоваться для цифровых пишущих видеоплееров нового поколения, так как на один Blu-Ray Disc умещается до 13 часов видеоинформации качества VHS (MPEG-2 с bitrate 3.8Mbps) или же 2 часа видео в модном сейчас в Японии формате HDTV (телевидение высокого разрешения до 1600x1200x32bit, MPEG-2 с bitrate от 8Mbps и выше).

Для hi-tech учреждений, предприятий, систем управления, образовательных заведений и других, где требуются большие объемы информации, понадобятся более емкие — двусторонние, двухслойные (или многослойные) Blu-Ray-диски с емкостью от 100 ГГб. Такие диски будут заключены в прозрачный картридж и использовать специальные Blu-Ray-приводы, оснащенные лазерами с разной длиной волны (в пределах синей части спектра) для чтения разных слоев. Первые прототипы 100 ГГб дисков уже созданы. Такие, кажущиеся сейчас огромными, объемы информации могут уже в ближайшем будущем стать нормой, так же как в свое время быстро привыкли к огромному скачку между 3,5′ дискетой (1.44 Мб) и CD-ROM (650 Мб). Через некоторое время и домашний сектор станет одним из потребителей многослойных Blu-ray-дисков, когда упадут первоначально высокие цены на приводы и носители информации этого формата.

Технологии Blu-Ray создавались в первую очередь для записи, хранения и воспроизведения видео- и аудиоинформации, то есть налицо сильная ориентация в сторону мультимедиа, хотя, разумеется, на Blu-Ray Disc можно записать и просто данные. Основными форматами хранения видео, как и в DVD, является MPEG2, форматы звука соответственно — AC3, MPEG1, MPEG Layer2. Для цифровых видеоплееров формата Blu-Ray декодирование будет осуществляться аппаратно, для компьютерных приводов — программно.

Нельзя не упомянуть о высокой скорости пересылки данных, которая будет осуществлена в Blu-Ray-устройствах. Так, согласно спецификации, максимальная скорость пересылки данных между Blu-Ray-приводом и целевым устройством (MPEG-2 декодер или компьютер) будет достигать 36Mbps, что при огромных объемах носителя весьма актуально. Такой скорости пересылки данных должна в полной мере соответствовать скорость считывания. К сожалению, не указывается, каким путем будет достигнута столь высокая скорость, так как если этот способ — повышение скорости вращения диска, то боюсь, что взорвавшиеся Blu-Ray-диски и сгоревшие приводы уже не за горами, разве что в игру вступит какой-нибудь неизвестный фактор, например новый состав материала, из которого будут делаться диски. Но тогда возникает вопрос совместимости с предыдущими поколениями носителей. Конечно, можно добавить логические схемы, которые будут определять тип носителя CD/DVD/Blu-Ray и соответственно менять максимальную скорость вращения для каждого типа, но это приведет к удорожанию привода. Путь же увеличения числа считывающих лазеров, как мы видим на примере

технологии True-X, ведет к взрывообразному увеличению стоимости привода.

Для обратной совместимости с предыдущими носителями информации, а это обязательное условие грядущей популярности Blu-Ray, привод должен иметь, по крайне мере, два лазера — основной синий и дополнительный красный. Сомневаюсь, что диски, для чтения которых требуется красный лазер, будут читаться синим. Много факторов мешают: меньшая толщина синего луча, иные отражательные свойства поверхности, более грубая структура самого диска и т.д. В результате опять вилка — выберешь совместимость со старыми форматами — проиграешь в цене, зато приобретешь благосклонное внимание консервативных слоев общества, откажешься от совместимости — удешевишь конструкцию, но отпугнешь покупателей, кроме наиболее радикальных hi-tech экстремалов. Это вилка для компьютерных приводов, так как там можно еще поставить несколько приводов, хотя для пользователя это означает и то, что придется платить за каждый из приводов, да и пятидюймовых слотов ограниченное количество. Для видеоплееров никаких вилок нет — совместимость с предыдущими форматами нужна в любом случае, библиотека DVD- и Video CD-фильмов уже очень велика, и никто не захочет отказываться от нее из-за призрачных перспектив, обещаемых Blu-Ray.

К сожалению, те грабли, на которые наступил в свое время DVD, ничему новый формат не научили — в Blu-Ray включена защита от нелегального копирования. К счастью, это будут не зоны, как ранее, а некий индивидуальный номер, который будет проставляться на всех записанных видеодисках. Не совсем ясно, для чего это делается, но в пресс-релизе гордо сказано, что «эта метка будет осуществлять реальную высококачественную защиту авторских прав». По всей видимости, для того устройства, на котором был записан диск, число воспроизведений будет не ограничено, а для других — какое-то число раз, то же самое будет с легально приобретенными фирменными дисками, на которых наверняка будет стоять защита от записи и перезаписи.

Характеристики Blu-Ray Disc

Емкость носилеля	23.3 Гб / 25 Гб / 27 Гб / 50 Гб / 100 Гб
Длина волны лазера	405 nm (blue-violet laser)
Шаг линзы	0.85 NA (numerical aperture)
Скорость пересылки данных	36 Mbps
Диаметр диска	120 mm
Толщина диска	1.2 mm (толщина оптически активного слоя — 0.1 mm)
Толщина трека	0.32 um
Минимальная длина точки	0.160/0.149/0.138 um
Плотность записи	16.8/18.0/19.5 Gbit/inch2

Формат записи видео	MPEG2 video (для видеоплеера), для компьютера — любые
Формат записи аудио	AC3, MPEG1, Layer2 (для видеоплеера), для компьютера — любые
Размер картриджа	129×131×7 mm

3.1.7 Голографические ЗУ

Устройства хранения информации на основе лазерной оптики уже на протяжении двух десятков лет, со времени появления CD в начале 1980-х, входят наряду с магнитными накопителями в базисный набор современных технологий внешней памяти. Но как для магнитных носителей, так и для CD, и для более продвинутой на этой основе технологии DVD всегда присутствуют принципиальные ограничения на емкость хранения, вызванные особенностями методов записи, по сути своей «плоских», т.е. размещающих информацию лишь на поверхности запоминающей среды. Как известно, эти неудобства обычно преодолеваются «в лоб» — двусторонней записью, установкой пакета параллельных пластин в жестком магнитном диске или нанесением нескольких полупрозрачных слоев в DVD (пока двух, в будущем до десятка). Но одновременно уже многие годы ученые и конструкторы работают над созданием принципиально иной технологии хранения информации, получившей название голографическая память и задействующей не поверхность, а весь объем запоминающей среды. Благодаря голографическому методу записи/считывания появляется возможность не только существенно увеличить емкость хранения информации, но и скорость ее обработки.

Вполне очевидно, что на сегодняшний день в такого рода высокопроизводительных системах хранения имеется ощутимая потребность сразу в нескольких областях: в вещательной индустрии, переходящей на цифровые форматы; в видеозаписи высокой четкости, требующей примерно в 10 раз более высоких объемов памяти, нежели занимают нынешние фильмы формата DVD; в разнообразных формах мультимедиа-контента, усложняющегося вместе с ростом пропускной способности широкополосных сетей; наконец, нельзя не упомянуть и саму по себе бездонную область видеоигр в условиях виртуальной реальности.

Голографическая память (или, кратко, холопамять) в принципе дает возможность размещения 1 терабайта, т.е. триллиона байт данных, в кристалле размером с кубик сахара. Говоря другими словами, 1 терабайт — это суммарная емкость более чем 1000 компакт-дисков. Причем, что любопытно, сама идея голографического устройства хранения была впервые выдвинута ученым-исследователем фирмы Polaroid Питером ван Хеерденом (Pieter J. van Heerden) еще в начале 1960-х годов, т.е. задолго до появления технологии CD. Примерно десятилетие спустя, в начале 1970-х,

разработчики из исследовательского центра RCA Laboratories сразу на двух примерах продемонстрировали практическую работоспособность концепции, записав 500 голограмм в кристалл ниобата лития, легированный атомами железа, и еще 550 голограмм с картинками высокого разрешения в особый светочувствительный полимерный материал. Однако дальнейшее развитие технологии на много лет оказалось заторможенным, с одной стороны, из-за отсутствия дешевых решений, пригодных к массовому производству, с другой же — вследствие стремительного развития недорогих технологий магнитной, полупроводниковой, а затем и лазерно-оптической памяти.

Прогресс последней, правда, невольно способствовал и возрождению интереса к голографическим методам хранения информации. В течение 1990-х годов сразу несколько серьезных ведомств и компаний, в первую очередь американское военное Агентство передовых исследовательских проектов (DARPA), корпорация IBM и исследовательский центр Bell Labs фирмы Lucent существенно продвинулись в создании практичной технологии холопамяти. Прототипы голографических устройств хранения, или HDSS (от holographic data storage system), созданные за последнее десятилетие, слегка различаются в деталях, однако все построены на основе единой концепции и сходных базовых компонентов.

Для всех современных устройств HDSS характерно использование следующих ключевых элементов: лазер, расщепитель луча для разделения лазерного пучка, зеркала для направления лазерных лучей, жидкокристаллическая панель в роли пространственного модулятора света, линзы для фокусировки лазерных лучей, кристалл ниобата лития или фотополимер в роли запоминающей среды, фотодетектор для считывания информации (камера с ПЗС1- или КМОП2-матрицей).

При записи данных, когда лазер генерирует луч, расщепитель (полупрозрачное зеркало) создает два когерентных, т.е. согласованных по длине волны, лазерных пучка. Один, именуемый объектным или предметным лучом, идет в пространственный светомодулятор (ПСМ), преобразующий подлежащие записи биты данных в массив темных и светлых пикселов. Сейчас в качестве ПСМ обычно выступает жидкокристаллический дисплей, размещающий данные в виде страницы размером порядка 1 миллиона бит. Вся эта информация целой страницей переносится в предметном луче далее, к светочувствительной среде (кристалл ниобата лития или полимер). Второй луч, именуемый опорным лучом, проходит по собственной «обходной» траектории и также падает на записывающий кристалл. В месте схождения двух лучей создается интерференционная картина, которая и записывается в устройство хранения как голограмма страницы данных. Небольшими изменениями угла наклона опорного луча или изменением длины волны лазера в тот же кристалл можно записывать множество других страниц данных.

Важным преимуществом системы голографической памяти является то, что и считывание целого массива информации (страницы) происходит

очень быстро, одним разом. Для того чтобы извлечь и восстановить данные, хранимые в виде голограммы, опорный луч освещает кристалл точно под тем же углом, под которым он находился при записи страницы. Лишь в этих условиях луч будет преломлен кристаллом именно так, что воссоздастся записанное изображение нужной страницы. Восстановленная страница проецируется в фотодетектор (сейчас это обычно ПЗС-матрица камеры, которая интерпретирует и передает считанную двоичную информацию в компьютер).

Ключевым моментом для любой системы голографического хранения данных является мультиплексирование, т.е. наложение множества страниц в одну запоминающую среду, обеспечивающее огромную емкость холопамяти. По сути дела, это напоминает проецирование множества картинок на один и тот же кадр фотопленки, однако здесь манипуляции опорным лучом позволяют выделять каждый снимок-голограмму индивидуально.

Ясно, что описанная технология радикально отличается от ныне используемых сразу в нескольких аспектах. Во-первых, наложение сотен голограмм друг на друга означает хранение существенно больших объемов информации в носителе малого размера. Во-вторых, данные записываются и считываются в параллели большими порциями, что существенно увеличивает скорость обработки. И что самое главное, в отличие от традиционных запоминающих устройств, все это обеспечивается не увеличением скорости вращения механических деталей, ростом частоты сканирования или уменьшением размеров физического носителя бита информации, а просто принципиально иным методом хранения данных на основе волновой интерференции. Кроме того, голографическая природа записываемых образов обеспечивает избыточность информации (т.е. повышенную стойкость к локальным повреждениям носителя), поскольку каждый бит данных хранится не в конкретной ячейке, а распределенным по всей интерференционной картине.

3.2 Надежность памяти ЭВМ

Запоминающие устройства составляют значительную часть объема аппаратных средств ЭВМ и систем обработки данных, поэтому на них приходится существенная доля отказов ЭВМ.

ЗУ являются сложными изделиями вычислительной техники. Для их построения в значительной степени используются элементы той же технологии и аналогичного конструктивного исполнения, что и для других устройств ЭВМ, поэтому для них характерны те же причины и виды отказов, что и для других устройств ЭВМ. Однако специфика выполняемых функций, особенности запоминающей среды и организации ЗУ обуславливают специфические причины и виды отказов [4].

Под отказом будем понимать устойчивое нарушение работоспособности ЗУ или его элемента. Сбой — нарушение

работоспособности, которое устраняется само или оператором без проведения ремонта. Ошибка — это искажение хранимого или считанного из ЗУ слова под воздействием отказа или сбоя [4].

Особенностью выполняемых ЗУ функций является то, что они предназначены для хранения информации, в то время как другие устройства ЭВМ информацию обрабатывают (преобразуют) и осуществляют ее ввод и вывод. Интервал времени между записью и воспроизведением информации, в течение которого к данной области ЗУ обращения отсутствуют, может быть значительным, и при отказах или сбоях бывает невозможно восстановить исходные данные. Кроме этого, хранящиеся данные являются исходным материалом для обработки другими устройствами ЭВМ и их искажение вызывает неправильную работу ЭВМ в целом. Что касается устройств обработки и ввода-вывода данных, то выполняемые ими операции в большинстве случаев могут быть повторены, а ошибка обнаружена и, если она случайная, исправлена. Поэтому значимость отказов или сбоев ЗУ может быть более высокой, чем для других устройств ЭВМ [4].

Особенности отказов ЗУ по сравнению с отказами других устройств ЭВМ обусловлены спецификой запоминающей среды и организацией ЗУ. Характерные для ЗУ причины сбоев и их влияние на хранимую в ЗУ информацию показаны в таблице [4].

Таблица 3.2 — Причины и последствия сбоев ЗУ

Тип ЗУ	Причины сбоев	Последствия сбоев
ОЗУ динамические	Воздействие альфа-частиц, стекание заряда	Ошибки в хранимых данных
ОЗУ статические биполярные	Влияние невыбранных ЗЭ на выбранные	Изменение времени выборки
ППЗУ с нихромовыми перемычками	Восстановление перемычек	Ошибки в хранимых данных
РПЗУ на МНОП и МОП ПЗ	Стекание заряда	То же
ЗУ на ФС	Влияние невыбранных ЗЭ на выбранные	Ошибки в хранимых данных
ЦМД ЗУ	Взаимовлияние доменов	То же
Электромеханические ЗУ на магнитных носителях	Взаимовлияние считанных сигналов	Ошибки в считанных данных

*Данные искажаются за счет перезаписи неверно считанной информации. ОЗУ — оперативные ЗУ; ППЗУ — перепрограммируемые ЗУ; РПЗУ — репрограммируемые ЗУ; МНОП — структуры металл-нитрид-окисел-полупроводник; ЗЭ — запоминающий элемент; МОП — структуры металл-окисел-полупроводник; МОП ПЗ — МОП-структуры с плавающим затвором; ЗУ на ФС — ЗУ на ферритовых сердечниках

Изменение (увеличение) времени выборки биполярных ОЗУ и искажения вследствие взаимовлияния считанных сигналов в ЗУ на магнитных носителях могут быть устранены повторным считыванием данных. Остальные искажения устраняются перезаписью исправленной информации. Зависимость положения границ области устойчивой работы (ОУР) запоминающего устройства от характера хранимой информации и адресной последовательности при обращении к ЗУ является причиной сбоев при «тяжелых» последовательностях адресов и данных [4].

Рассмотрим влияние организации ЗУ на ошибки в считанной информации. Все ЗУ в зависимости от организации можно условно разбить на две большие группы: ЗУ с произвольной выборкой и ЗУ с последовательной и произвольно-последователь-ной выборкой. В ЗУ с произвольной выборкой схемы дешифрации обеспечивают выбор любого слова данных. К таким ЗУ относятся полупроводниковые интегральные ЗУ и ЗУ на ФС. В них возможны как однобитовые ошибки из-за отказов ЗЭ, так и пакетные ошибки, когда из-за отказов схем выборки и управления искажается информация в группе соседних разрядов считанного слова [4].

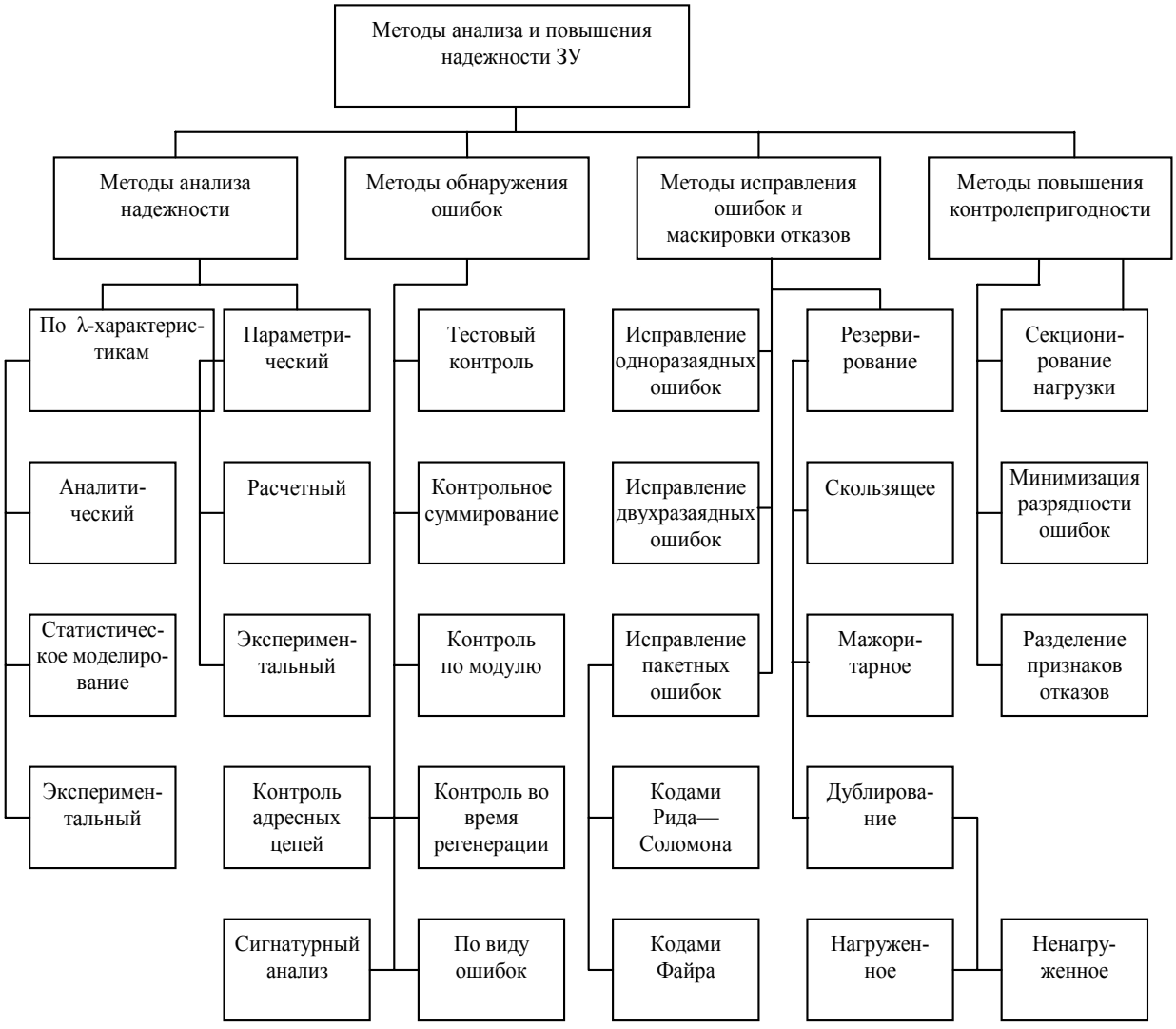


Рисунок 3.11 — Методы повышения надежности ЗУ

В зависимости от типа и емкости ЗУ, используемой схемотехники и конструкции изменяется разрядность пакетных ошибок, количество искаженных адресов и соотношения между интенсивностями ошибок различных видов. Значительную долю составляют однобитовые ошибки, вызванные отказами запоминающих элементов (ЗЭ), на которые приходится большая часть оборудования и интенсивности отказов ЗУ [4].

В ЗУ с последовательной и произвольно-последовательной выборкой (ЗУ на магнитных носителях, ЦМД ЗУ) для хранения данных используются участки магнитной поверхности. Вследствие высокой плотности размещения хранимых данных дефекты запоминающей среды или взаимовлияние элементов приводят к ошибкам нескольких соседних бит, т.е. пакетным ошибкам. Для таких ЗУ характерны пакетные ошибки в считанных данных [4].

Указанные особенности обуславливают специфику используемых методов повышения надежности ЗУ. На рисунке классифицированы методы анализа и повышения надежности ЗУ. Приведенные методы используются как отдельно, так и совместно для одного ЗУ [4].

3.2.1 Коды Хемминга

Коды Хемминга — простейшие линейные коды с минимальным расстоянием 3, то есть способные исправить одну ошибку.

Основные характеристики корректирующих кодов

В настоящее время наибольшее внимание с точки зрения технических приложений уделяется двоичным блочным корректирующим кодам. При использовании блочных кодов цифровая информация передаётся в виде отдельных кодовых комбинаций (блоков) равной длины. Кодирование и декодирование каждого блока осуществляется независимо друг от друга.

Почти все блочные коды относятся к разделимым кодам, кодовые комбинации которых состоят из двух частей: информационной и проверочной. При общем числе n символов в блоке число информационных символов равно k, а число проверочных символов

$$r = n - k. \qquad (3.1)$$

К основным характеристикам корректирующих кодов относятся:

- число разрешённых и запрещённых кодовых комбинаций;
- избыточность кода;
- минимальное кодовое расстояние;
- число обнаруживаемых или исправляемых ошибок;
- корректирующие возможности кодов.

Число разрешённых и запрещённых кодовых комбинаций.

Для блочных двоичных кодов, с числом символов в блоках, равным n, общее число возможных кодовых комбинаций определяется значением

$$N_0 = 2^n. \quad (3.2)$$

Число разрешённых кодовых комбинаций при наличии k информационных разрядов в первичном коде равно

$$N_k = 2^k. \quad (3.3)$$

Очевидно, что число запрещённых комбинаций равно:

$$N_3 = N_0 - N_k = 2^n - 2^k, \quad (3.4)$$

а с учётом (3.2) отношение будет:

$$N_0 / N_k = 2^n / 2^k = 2^{n-k} = 2^r, \quad (3.5)$$

где r — число избыточных (проверочных) разрядов в блочном коде.

Избыточность корректирующего кода

Избыточностью корректирующего кода называют величину

$$\chi = \frac{r}{n} = \frac{n-k}{n} = 1 - \frac{k}{n}, \quad (3.6)$$

откуда следует

$$B_k = \frac{k}{n} = 1 - \chi. \quad (3.7)$$

Эта величина показывает, какую часть общего числа символов кодовой комбинации составляют информационные символы. В теории кодирования величину B_k называют относительной скоростью кода. Если производительность источника информации равна H символов в секунду, то скорость передачи после кодирования этой информации окажется равной

$$B = H \frac{k}{n}, \quad (3.8)$$

поскольку в закодированной последовательности из каждых n символов только k символов являются информационными.

Если число ошибок, которые нужно обнаружить или исправить, значительно, то необходимо иметь код с большим числом проверочных символов. Чтобы при этом скорость передачи оставалась достаточно высокой, необходимо в каждом кодовом блоке одновременно увеличивать как общее число символов, так и число информационных символов. При этом длительность кодовых блоков будет существенно возрастать, что приведёт к задержке информации при передаче и приёме. Чем сложнее кодирование, тем длительнее временная задержка информации.

Минимальное кодовое расстояние — d_{min}

Для того чтобы можно было обнаружить и исправлять ошибки, разрешённая комбинация должна как можно больше отличаться от запрещённой. Если ошибки в канале связи действуют независимо, то

вероятность преобразования одной кодовой комбинации в другую будет тем меньше, чем большим числом символов они различаются.

Если интерпретировать кодовые комбинации как точки в пространстве, то отличие выражается в близости этих точек, т.е. в расстоянии между ними.

Количество разрядов (символов), которыми отличаются две кодовые комбинации, можно принять за кодовое расстояние между ними. Для определения этого расстояния нужно сложить две кодовые комбинации по модулю 2 и подсчитать число единиц в полученной сумме. Например, две кодовые комбинации

$x_i = 01011$ и $x_j = 10010$ имеют расстояние $d(x_i,x_j)$, равное 3, так как

$$\begin{aligned} x_i &= 01011 \rightarrow W = 3 \\ &\oplus \\ x_j &= 10010 \rightarrow W = 2 \\ x_i \oplus x_j &= 11001 \rightarrow d(x_i,x_j) = 3. \end{aligned} \qquad (3.9)$$

(Здесь под операцией «⊕» понимается сложение по mod2, логическое «или не».)

Заметим, что кодовое расстояние $d(x_i, x_0)$ между комбинацией x_i и нулевой $x_0 = 00..0$ называют весом W комбинации x_i, т.е. вес x_i равен числу «1» в ней.

Расстояние между различными комбинациями некоторого конкретного кода могут существенно отличаться. Так, в частности, в безызбыточном первичном натуральном коде ($n = k$) это расстояние для различных комбинаций может изменяться от единицы до величины n, равной значности кода. Особую важность для характеристики корректирующих свойств кода имеет минимальное кодовое расстояние d_{min}, определяемое при попарном сравнении всех кодовых комбинаций, которое называют расстоянием Хемминга.

В безызбыточном коде все комбинации являются разрешёнными, и, следовательно, его минимальное кодовое расстояние равно единице — $d_{min} = 1$. Поэтому достаточно исказиться одному символу, чтобы вместо переданной комбинации была принята другая разрешённая комбинация. Чтобы код обладал корректирующими свойствами, необходимо ввести в него некоторую избыточность, которая обеспечивала бы минимальное расстояние между любыми двумя разрешёнными комбинациями не менее двух — $d_{min} \geq 2$.

Минимальное кодовое расстояние является важнейшей характеристикой помехоустойчивых кодов, указывающей на гарантируемое число обнаруживаемых или исправляемых заданным кодом ошибок.

Число обнаруживаемых или исправляемых ошибок

При применении двоичных кодов учитывают только дискретные искажения, при которых единица переходит в нуль (1 → 0) или нуль переходит в единицу (0 → 1). Переход 1 → 0 или 0 → 1 только в одном

элементе кодовой комбинации называют единичной ошибкой (единичным искажением). В общем случае под кратностью ошибки подразумевают число позиций кодовой комбинации, на которых под действием помехи одни символы оказались заменёнными на другие. Возможны двукратные ($g = 2$) и многократные ($g > 2$) искажения элементов в кодовой комбинации в пределах $0 \le g \le n$.

Минимальное кодовое расстояние является основным параметром, характеризующим корректирующие способности данного кода. Если код используется только для обнаружения ошибок кратностью g_0, то необходимо и достаточно, чтобы минимальное кодовое расстояние было равно

$$d_{min} \ge g_0 + 1. \qquad (3.10)$$

В этом случае никакая комбинация из g_0 ошибок не может перевести одну разрешённую кодовую комбинацию в другую разрешённую. Таким образом, условие обнаружения всех ошибок кратностью g_0 можно записать в виде:

$$g_0 \le d_{min} - 1. \qquad (3.11)$$

Чтобы можно было исправить все ошибки кратностью g_u и менее, необходимо иметь минимальное расстояние, удовлетворяющее условию:

$$d_{min} \ge 2\, g_u + 1. \qquad (3.12)$$

В этом случае любая кодовая комбинация с числом ошибок g_u отличается от каждой разрешённой комбинации не менее чем в $g_u + 1$ позициях. Если условие (3.12) не выполнено, возможен случай, когда ошибки кратности g исказят переданную комбинацию так, что она станет ближе к одной из разрешённых комбинаций, чем к переданной, или даже перейдёт в другую разрешённую комбинацию. В соответствии с этим, условие исправления всех ошибок кратностью не более g_u можно записать в виде:

$$g_u \le (d_{min} - 1) / 2. \qquad (3.13)$$

Из (3.10) и (3.12) следует, что если код исправляет все ошибки кратностью g_u, то число ошибок, которые он может обнаружить, равно $g_0 = 2g_u$. Следует отметить, что соотношения (3.10) и (3.12) устанавливают лишь гарантированное минимальное число обнаруживаемых или исправляемых ошибок при заданном d_{min} и не ограничивают возможность обнаружения ошибок большей кратности. Например, простейший код с проверкой на чётность с $d_{min} = 2$ позволяет обнаруживать не только одиночные ошибки, но и любое нечётное число ошибок в пределах $g_0 < n$.

Корректирующие возможности кодов

Вопрос о минимально необходимой избыточности, при которой код обладает нужными корректирующими свойствами, является одним из важнейших в теории кодирования. Этот вопрос до сих пор не получил полного решения. В настоящее время получен лишь ряд верхних и нижних оценок (границ), которые устанавливают связь между максимально

возможным минимальным расстоянием корректирующего кода и его избыточностью.

Так, граница Плоткина даёт верхнюю границу кодового расстояния d_{min} при заданном числе разрядов n в кодовой комбинации и числе информационных разрядов k, и для двоичных кодов:

$$d_{\min} \le \frac{n2^{k-1}}{2^k - 1} \tag{3.14}$$

или $$r \ge 2(d_{min} - 1) - \log_2 d_{min} \tag{3.15}$$

при $n \ge 2d_{min} - 1$.

Верхняя граница Хемминга устанавливает максимально возможное число разрешённых кодовых комбинаций (2^k) любого помехоустойчивого кода при заданных значениях n и d_{min}:

$$2^k \le 2^n / \sum_{i=0}^{\frac{d_{\min}-1}{2}} C_n^i, \tag{3.16}$$

где C_n^i — число сочетаний из n элементов по i элементам.

Отсюда можно получить выражение для оценки числа проверочных символов:

$$r \ge \log_2 \left(\sum_{i=0}^{\frac{d_{\min}-1}{2}} C_n^i \right). \tag{3.17}$$

Для значений $(d_{min}/n) \le 0.3$ разница между границей Хемминга и границей Плоткина сравнительно невелика.

Граница Варшамова—Гильберта для больших значений n определяет нижнюю границу для числа проверочных разрядов, необходимого для обеспечения заданного кодового расстояния:

$$r \ge \log_2 \left(\sum_{i=0}^{\frac{d-2}{2}} C_{n-1}^i \right). \tag{3.18}$$

Отметим, что для некоторых частных случаев Хемминг получил простые соотношения, позволяющие определить необходимое число проверочных символов:

$r \ge \log_2(n + 1)$ для $d_{min} = 3$,

$r \ge \log_2(2n)$ для $d_{min} = 4$.

Блочные коды с $d_{min} = 3$ и 4 в литературе обычно называют кодами Хемминга.

Все приведённые выше оценки дают представление о верхней границе числа d_{min} при фиксированных значениях n и k или оценку снизу числа проверочных символов r при заданных k и d_{min}.

Существующие методы построения избыточных кодов решают в основном задачу нахождения такого алгоритма кодирования и декодирования, который позволял бы наиболее просто построить и реализовать код с заданным значением d_{min}. Поэтому различные корректирующие коды при одинаковых dmin сравниваются по сложности кодирующего и декодирующего устройств. Этот критерий является в ряде случаев определяющим при выборе того или иного кода.

Корректирующие коды Хемминга

Построение кодов Хемминга базируется на принципе проверки на чётность веса W (числа единичных символов) в информационной группе кодового блока. Поясним идею проверки на чётность на примере простейшего корректирующего кода, который так и называется кодом с проверкой на чётность или кодом с проверкой по паритету (равенству).

В таком коде к кодовым комбинациям безызбыточного первичного двоичного k-разрядного кода добавляется один дополнительный разряд (символ проверки на чётность, называемый проверочным, или контрольным). Если число символов «1» исходной кодовой комбинации чётное, то в дополнительном разряде формируют контрольный символ 0, а если число символов «1» нечётное, то в дополнительном разряде формируют символ 1. В результате общее число символов «1» в любой передаваемой кодовой комбинации всегда будет чётным.

Таким образом, правило формирования проверочного символа сводится к следующему:

$$r_1 = i_1 \oplus i_2 \oplus \ldots \oplus i_k,$$

где i — соответствующий информационный символ (0 или 1), k — общее их число, а под операцией «⊕» здесь и далее понимается сложение по mod 2. Очевидно, что добавление дополнительного разряда увеличивает общее число возможных комбинаций вдвое по сравнению с числом комбинаций исходного первичного кода, а условие чётности разделяет все комбинации на разрешённые и неразрешённые. Код с проверкой на чётность позволяет обнаруживать одиночную ошибку при приёме кодовой комбинации, так как такая ошибка нарушает условие чётности, переводя разрешённую комбинацию в запрещённую.

Критерием правильности принятой комбинации является равенство нулю результата S суммирования по mod 2 всех n символов кода, включая проверочный символ r_1. При наличии одиночной ошибки S принимает значение 1:

$$S = \underbrace{r_1 \oplus i_1 \oplus i_2 \oplus \ldots}_{n} \oplus i_k = \begin{cases} 0 \text{ — ошибки нет,} \\ 1 \text{ — однократная ошибка.} \end{cases}$$

Существует классический код Хемминга с маркировкой

$$(n,k) = (2^r - 1,\ 2^r - 1 - r), \qquad (3.19)$$

т.е. — (7,4), (15,11), (31,26) …

При других значениях числа информационных символов k получаются так называемые усечённые (укороченные) коды Хемминга. Так, для международного телеграфного кода МТК-2, имеющего 5 информационных символов, потребуется использование корректирующего кода (9,5), являющегося усечённым от классического кода Хемминга (15,11), так как число символов в этом коде уменьшается (укорачивается) на 6. Для примера рассмотрим классический код Хемминга (7,4). В простейшем варианте при заданных четырёх ($k = 4$) информационных символах (i_1, i_2, i_3, i_4) будем полагать, что они сгруппированы в начале кодового слова, хотя это и не обязательно. Дополним эти информационные символы тремя проверочными символами ($r = 3$), задавая их следующими равенствами проверки на чётность, которые определяются соответствующими алгоритмами:

$$r_1 = i_1 \oplus i_2 \oplus i_3;$$
$$r_2 = i_2 \oplus i_3 \oplus i_4;$$
$$r_3 = i_1 \oplus i_2 \oplus i_4.$$

Рассмотрим декодирование для (7,4) — кода Хемминга слова

$$V = (i_1', i_2', i_3', i_4', r_1', r_2', r_3').$$

Апостроф означает, что любой символ слова может быть искажён помехой в канале передачи.

Для исправления ошибок необходимо построить последовательность:

$$s_1 = r_1' \oplus i_1' \oplus i_2' \oplus i_3';$$
$$s_2 = r_2' \oplus i_2' \oplus i_3' \oplus i_4';$$
$$s_3 = r_3' \oplus i_1' \oplus i_2' \oplus i_4'.$$

Трёхсимвольная последовательность (s_1, s_2, s_3) называется синдромом. Термин «синдром» используется и в медицине, где он обозначает сочетание признаков, характерных для определённого заболевания. В данном случае синдром $S = (s_1, s_2, s_3)$ представляет собой сочетание результатов проверки на чётность соответствующих символов кодовой группы и характеризует определённую конфигурацию ошибок (шумовой вектор).

Число возможных синдромов определяется выражением

$$S = 2^r. \qquad (3.20)$$

При числе проверочных символов $r = 3$ имеется восемь возможных синдромов ($2^3 = 8$). Нулевой синдром (000) указывает на то, что ошибки при приёме отсутствуют или не обнаружены. Всякому ненулевому синдрому соответствует определённая конфигурация ошибок, которая и исправляется. Классические коды Хемминга (3.19) имеют число синдромов, точно равное их необходимому числу, позволяют исправить все однократные ошибки в любом информативном и проверочном символах и включают один нулевой синдром. Такие коды называются плотноупакованными.

Для рассматриваемого кода (7,4) в таблице 3.1 представлены ненулевые синдромы и соответствующие конфигурации ошибок.

Таблица 3.1 — Синдромы и ошибки классического кода Хемминга

Синдром	001	010	011	100	101	110	111
Конфигурация ошибок	0000001	0000010	0001000	0000100	1000000	0010000	0100000
Ошибка в символе	r_3	r_2	i_4	r_1	i_1	i_3	i_2

Таким образом, код (7,4) позволяет исправить все одиночные ошибки. Простая проверка показывает, что каждая из ошибок имеет свой единственный синдром. При этом возможно создание такого цифрового корректора ошибок (дешифратора синдрома), который по соответствующему синдрому исправляет соответствующий символ в принятой кодовой группе.

Две или более ошибки превышают возможности корректирующего кода Хемминга.

Идея построения подобного корректирующего кода, естественно, не меняется при перестановке позиций символов в кодовых словах. Все такие варианты также называются (7,4) — кодами Хемминга.

Исправление двухразрядных ошибок в запоминающих устройствах с кодами Хемминга методом двойного инвертирования

В ЗУ возможны как отказы, так и сбои. При отказе ЗЭ всегда находится в одном состоянии — либо 1, либо 0. Таким образом, в считанном из ЗУ слове находится постоянная ошибка. Ошибки из-за сбоев являются случайными и исправляются при последующем считывании или записи информации [2].

Двухразрядные постоянные или комбинированные ошибки (одна постоянная ошибка и одна случайная) могут быть исправлены кодом Хемминга при использовании метода двойного инвертирования. Для этого считанное из ЗУ слово, содержащее двухразрядную ошибку, инвертируется вместе с контрольными разрядами и вновь записывается в ЗУ по прежнему адресу. Постоянные ошибки при этом исправляются. Затем следует считывание инверсного слова, его инвертирование и исправление случайной ошибки, если таковая имелась.

Если в разряде слова имеются две постоянные ошибки, то данные разряды имеют инверсное значение по отношению к разрядам ранее записанного слова. Поэтому записанное по данному адресу инверсное слово этих ошибок не содержит. Если инверсное слово считать и проинвертировать второй раз, то получим слово, не содержащее ошибок.

Исправление комбинированных ошибок проиллюстрировано рис. 3.9. После обнаружения двухразрядной ошибки в считанном ЗУ слове производится его инвертирование и последующая запись в ЗУ. Все разряды слова, кроме ошибочного, будут инверсными по отношению к первоначально записанной информации. При последующем считывании и инвертировании считанного слова в нем будет один неисправный разряд. Рассчитанный синдром позволит исправить ошибку.

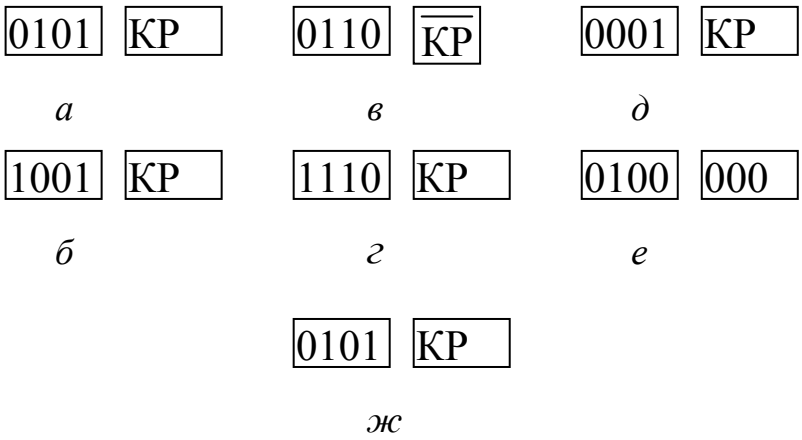


Рисунок 3.9 — Иллюстрация метода исправления комбинированных ошибок методом двойного инвертирования:
а — записываемое слово; *б* — считанное слово; *в* — записываемое инвертированное слово; *г* — считанное инвертированное слово; *д* — слово после второго инвертирования, постоянная ошибка исправлена; *е* — выходы дешифратора ошибок; *ж* — исправленное слово

Данный способ обеспечивает исправление двухразрядных постоянных или комбинированных ошибок кодом Хемминга без введения дополнительных контрольных разрядов. Требуется лишь дополнительный цикл записи и считывания инверсного слова по прежнему адресу и повторный цикл формирования синдрома и записи слова в прямом коде при возникновении двухразрядной ошибки.

3.2.2 Коды БЧХ

Полиномиальное определение циклических кодов и операции с ними

Циклические коды являются частным случаем систематических, линейных [n, k] кодов. Название ЦК получили из-за своего основного свойства: циклическая перестановка символов разрешённой кодовой комбинации даёт также разрешённую кодовую комбинацию. При циклической перестановке символы кодового слова перемещаются слева направо на одну позицию, причем крайний справа символ переносится на место крайнего левого.

Если, например, A_1 — 101100, то разрешённой кодовой комбинацией будет и A_2 — 010110, полученная циклической перестановкой. Отметим, что перестановка производится вместе с проверочными символами, и по правилам линейных кодов сумма по модулю 2 двух разрешённых кодовых комбинаций даёт также очередную разрешённую кодовую комбинацию.

Описание ЦК связано с представлением кодовых комбинаций в виде полиномов (многочленов) фиктивной переменной «*X*». Для примера переведём кодовое слово $A_1 = 101100$ в полиномиальный вид

i	6	5	4	3	2	1
код	1	0	1	1	0	0

При этом

$$A_1(X) = 1 \cdot X^5 + 0 \cdot X^4 + 1 \cdot X^3 + 1 \cdot X^2 + 0 \cdot X^1 + 0 \cdot X^0 = X^5 + X^3 + X^2.$$

Действия с кодовыми векторами, представленными в виде полиномов, производятся по правилам арифметики по модулю 2, в которой вычитание равносильно сложению. Так, из равенства $X_{n-1} = 0$ получаем $X_n = 1$. Прибавив к левой и правой частям по единице, имеем $X_n + 1 = 1 \oplus 1 = 0$. Таким образом, вместо двучлена X_{n-1} можно ввести бином X_{n+1} или $1 + X_n$, из чего следует, что $X_k \oplus X_k = X_k\ (1 \oplus 1) = 0$ и при последующих операциях с полиномами необходимо вычёркивать пары фиктивных переменных X с одинаковыми степенями.

Приведём далее порядок суммирования (вычитания), умножения и деления полиномов с учётом того, что операция суммирования осуществляется по модулю 2. В примерах используем вышеприведённые кодовые комбинации

$$A_1(X) = X^5 + X^3 + X^2 \quad \text{и} \quad A_2(X) = X^4 + X^2 + X.$$

Суммирование (вычитание):

$$A_1 + A_2 = A_1 - A_2 = X^5 + X^4 + X^3 + X^2 + X^2 + X = X^5 + X^4 + X^3 + X$$

или

$$\begin{array}{ll} A_1 & 101100 \\ & \oplus \\ A_2 & \underline{010110} \\ & 111010 = X^5 + X^4 + X^3 + X \end{array}$$

Умножение:

$$A_1 \times A_2 = (X^5 + X^3 + X^2) \times (X^4 + X^2 + X) = X^9 + X^7 + X^6 + X^7 + X^5 + X^4 + X^6 + X^4 + X^3 = X^9 + X^5 + X^3 = 1000101000.$$

Деление: $\dfrac{A_1}{A_2}$

$$\begin{array}{l|l} X^5 + X^3 + X^2 & X^4 + X^2 + X \\ \hline X^5 + X^3 + X^2 & X \end{array}$$

0 0 0 — остаток при делении R(X) = 0

Из последнего примера следует, что при циклическом сдвиге вправо на один разряд необходимо исходную кодовую комбинацию поделить на X, а умножение на X эквивалентно сдвигу влево на один символ.

3.2.2.1 Порождающие полиномы циклических кодов (ЦК)

Формирование разрешённых кодовых комбинаций ЦК B_i (X) основано на предварительном выборе так называемого порождающего (образующего) полинома $G(X)$, который обладает важным отличительным признаком: все комбинации B_i (X) делятся на порождающий полином $G(X)$ без остатка, т.е.

$$\frac{B_i(X)}{G(X)} = A_i(X) \text{ (при остатке } R(X) = 0), \quad (3.20)$$

где $A_i(X)$ — информативный полином (кодовая комбинация первичного кода, преобразуемого в корректирующий ЦК).

Поскольку, как отмечалось выше, ЦК относятся к классу блочных разделимых кодов, у которых при общем числе символов n число информационных символов в $A_i(X)$ равно k, то степень порождающего полинома определяет число проверочных символов $r = n – k$.

Из этого свойства следует сравнительно простой способ формирования разрешённых кодовых комбинаций ЦК — умножение комбинаций информационного кода $A_i(X)$ на порождающий полином $G(X)$:

$$B_i(\mathrm{X}) = A_i(X)G(X). \quad (3.21)$$

В теории циклических кодов доказывается, что порождающими могут быть только такие полиномы, которые являются делителями двучлена (бинома) X_{n+1}:

$$\frac{X^n + 1}{G(X)} = H(X) \text{ (при остатке } R(X) = 0). \quad (3.22)$$

Некоторые из порождающих полиномов приведены в табл. 3.3.

Таблица 3.3 — Порождающие полиномы

r-степень полинома G(X)	Порождающий полином G(X)	Запись полинома по mod 2	Запись полинома по mod 8	n	k	Примечание
1	X+1	11	3	3	2	Код с проверкой на чётность (КПЧ)
2	X^2+X+1	111	7	3	1	Код с повторением
3	X^3+X^2+1 X^3+X+1	1101 1011	13 15	7 7	4 4	Классический код Хемминга
4	X^4+X^3+1 X^4+ X+1 X^4+X^2+X+1 X^4+X^3+X^2+1	11001 10011 10111 11101	31 23 27 35	15 15 7 7	11 11 3 3	Классический код Хемминга Коды Файра—Абрамсона
5	X^5+X^2+1 X^5+X^3+1 ...	100101 101001	45 51	31 31	26 26	Классический код Хемминга
6	... X^6+X^5+X^4+ +X^3+X^2+X+1	1111111	177	7	1	Код с повторением

Коды Хемминга также принадлежат к классу ЦК, однако при этом группа проверочных символов кода получается сразу «в целом» при делении информативной кодовой группы на порождающий полином, а не «поэлементно», когда последовательное суммирование по модулю 2 соответствующих информативных символов давало очередной символ проверочной группы. Отметим, что два варианта порождающих полиномов кода Хемминга (7,4), с записью по модулю 2 в виде 1101 и 1011, представляют собой так называемые двойственные многочлены (полиномы).

Двойственные многочлены определяются следующим образом: если задан полином в виде

$$h(X) = h_0 + h_1X + h_2X^2 + \ldots + h_rX^r,$$

то двойственным к нему полиномом является

$$\tilde{h}(X) = h_0X^r + h_1X^{r-1} + \ldots + h_r, \qquad (3.23)$$

т.е. весовые коэффициенты исходного полинома, зачитываемые слева направо, становятся весовыми коэффициентами двойственного полинома

при считывании их справа налево. Говоря образно, набор весовых коэффициентов «вывёртывается наизнанку».

Обратим внимание на то, что в полных таблицах порождающих ЦК полиномов двойственные полиномы, как правило, не приводятся.

Наряду с тем, что порождающие полиномы кода Хемминга (7,4) являются двойственными друг другу, они также являются неприводимыми.

Неприводимые полиномы не делятся ни на какой другой полином степени меньше *r,* поэтому их называют ещё неразложимыми, простыми и примитивными.

В соответствии со свойством (3.22) порождающих полиномов G(X) бином X^7+1 раскладывается на три неприводимых полинома

$$X^7+1 = (X+1)\cdot(X^3+X^2+1)\cdot(X^3+X+1) = G_1(X)\times G_2(X)\times G_3(X), \quad (3.24)$$

каждый из которых является порождающим для следующих кодов:

$G_1(X) = X+ 1$ — код с проверкой на чётность, КПЧ (7, 6);

$G_2(X) = X^3 + X^2+ 1$ — первый вариант кода Хемминга (7,4);

$G_3(X) = X^3 + X+ 1$ — двойственный $\overline{G}_2(X)$, второй вариант кода Хемминга.

Кроме того, различные вариации произведений $G_{1,2,3}(X)$ дают возможность получить остальные порождающие полиномы:

$G_4(X) = G_1(X)\cdot G_2(X) = (X + 1)\cdot(X_3 + X_2+ 1) = X_4 + X_2+ X +1$ — код Абрамсона (7,3);

$G_5(X) = G_1(X)\cdot G_3(X) = (X + 1)\cdot(X^3 + X+ 1) = X^4 + X^3+ X^2 +1$ — двойственный G_4(X);

$G_6(X) = G_2(X)\cdot G_3(X) = (X^3 + X^2+ 1)\cdot(X^3 + X+ 1) =$

$= X^6 + X^5 +X^4 + X^3 + X^2+ X +1$ — код с повторением (7,1).

Таким образом, для постоянного заданного значения n все возможные порождающие полиномы ЦК размещаются между кодами с проверкой на чётность (n, n –1) (r = 1) и кодами с повторением (n, 1) ($r = n$ – 1), которые правомерно и называют «кодами антиподами».

При выборе применяемых в системах связи корректирующих кодов необходимо позаботиться о том, чтобы, во-первых, избыточность кода была минимальной, т.е. относительная скорость кода или эффективность кода была максимальной, а, во-вторых, техника кодирования и декодирования была по возможности проста.

3.2.2.2 Принципы формирования и обработки разрешённых кодовых комбинаций циклических кодов

Циклические коды (ЦК) составляют множество многочленов $B_i(X)$ степени n – 1 и менее (до $r = n – k$, где r — число проверочных символов), кратных порождающему (образующему) полиному G(X) степени r, который, в свою очередь, должен быть делителем бинома X_{n+1}, т.е. остаток

после деления бинома на G(X) должен равняться нулю. Учитывая, что ЦК принадлежат к классу линейных, групповых кодов, сформулируем ряд их основных свойств.

1. Сумма разрешённых кодовых комбинаций ЦК образует разрешённую кодовую комбинацию

$$B_i(X) \oplus B_j(X) = B_k(X). \qquad (3.25)$$

2. Поскольку к числу разрешённых кодовых комбинаций ЦК относится нулевая комбинация 000...00, то минимальное кодовое расстояние d_{min} для ЦК определяется минимальным весом разрешённой кодовой комбинации:

$$D_{min} = W_{min}. \qquad (3.26)$$

3. Циклический код не обнаруживает только такие искажённые помехами кодовые комбинации, которые приводят к появлению на стороне приёма других разрешённых комбинаций этого кода из набора N_0.

4. Значения проверочных элементов $r = n - k$ для ЦК могут определяться путём суммирования по модулю 2 ряда определённых информационных символов кодовой комбинации $A_i(X)$. Например, для кода Хемминга (7,4) с порождающим полиномом $G(X) = X^3+X+1$ алгоритм получения проверочных символов будет следующим

$$r_1 = i_1 \oplus i_2 \oplus i_3;$$
$$r_2 = i_2 \oplus i_3 \oplus i_4;$$
$$r_3 = i_1 \oplus i_2 \oplus i_4.$$

Эта процедура свидетельствует о возможности «поэлементного» получения проверочной группы для каждой кодовой комбинации $A_i(X)$.

5. Как было показано ранее, умножение полинома на X приводит к сдвигу членов полинома на один разряд влево, а при умножении на X^r — соответственно на r разрядов влево, с заменой r младших разрядов полинома «нулями». Умножение полинома на X свидетельствует о том, что при этой процедуре X является «оператором сдвига». Деление полинома на X приводит к соответствующему сдвигу членов полинома вправо с уменьшением показателей членов на 1. Процедура сдвига позволяет к исходной кодовой комбинации $A_i(X)$, после домножения её на X^r, дописать справа r проверочных символов.

6. Поскольку разрешённые кодовые слова ЦК $B_i(X)$ без остатка делятся на порождающий полином G(X) с получением итога в виде информационной комбинации $A_i(X)$ (3.20), то имеется возможность формировать $B_i(X)$ на стороне передачи (кодирующее устройство) простым методом умножения (3.21).

Два последних свойства ЦК позволяют осуществить построение кодеров ЦК двумя методами: методом умножения и методом деления полиномов.

Метод деления полиномов позволяет представить разрешённые к передаче кодовые комбинации в виде разделённых информационных $A_i(X)$ и проверочных $R_i(X)$ символов, т.е. получить блочный код.

Поскольку число проверочных символов равно *r,* то для компактной их записи в последние младшие разряды кодового слова надо предварительно к $A_i(X)$ справа приписать *r* «нулей», что эквивалентно умножению $A_i(X)$ на оператор сдвига X^r (см. свойство 5 ЦК).

На практике предпочитают использование метода деления полиномов при построении кодеков, поскольку при этом имеется возможность представить кодовую комбинацию в виде разделённых информационных и проверочных символов:

$$B_i(X) = A_i(X) \cdot X^r + R_i(X), \qquad (3.27)$$

где $R_i(X)$ — остаток от деления $A_i(X) \cdot X^r / G(X)$.

В алгоритме (3.27) можно выделить три этапа формирования разрешённых кодовых комбинаций в кодирующем устройстве:

1) к комбинации первичного кода $A_i(X)$ дописывается справа r нулей, что эквивалентно умножению $A_i(X)$ на X^r;

2) произведение $A_i(X) \cdot X^r$ делится на соответствующий порождающий полином $G(X)$ и определяется остаток $R_i(X)$, степень которого не превышает $r - 1$, этот остаток и даёт группу проверочных символов;

3) вычисленный остаток присоединяется справа к $A_i(X) \cdot X^r$.

Пример 1. Рассмотрим процедуру кодирования по алгоритму (4.15): для кодовой комбинации $A = 1001$ сформировать кодовую комбинацию циклического кода (7,4).

В заданном ЦК $n = 7$, $k = 4$, $r = 3$, и из табл. 1 выберем порождающий полином $G(X) = X^3 + X + 1$ (код Хемминга). Выполним три необходимые операции для получения кодовой комбинации ЦК согласно алгоритму (3.27):

$A_i(X) = 1001 \sim X^3 + 1$ (знак «~» — тильда — означает соответствие).

1. $A_i(X) \cdot X^r = (X^3 + 1) \cdot X^3 = X^6 + X^3 \sim 1001000$ ($n = 7$).

2.

$$A_i(X) \cdot X^r / G(X) = X^6 + X^3 \;\Big|\; \underline{X^3 + X + 1}$$

$$\underline{X^6 + X^4 + X^3} \qquad X^3 + X$$

$$X^4$$

$$\underline{X^4 + X^2 + X}$$

$X^2 + X$ — остаток при делении

$R(X) = X^2 + X \sim 110.$

3. $B_i(X) = A_i(X) \cdot X^r + R_i(X) = 1001110$ — итоговая комбинация ЦК.

Синдромный метод декодирования (СМД) предполагает в ДУ принятую кодовую комбинацию поделить на порождающий полином. Если принятая комбинация является разрешённой, т.е. не искажена

помехами в канале связи, то остаток от деления будет нулевым. Ненулевой остаток свидетельствует о наличии в принятой кодовой комбинации ошибок, остаток от деления и называется синдромом. Для исправления ошибки на стороне приёма необходимо знать не только факт её существования, но и её местонахождение, которое определяется по установленному виду вектора ошибки z(*X*).

После передачи по каналу с помехами принимается кодовое слово

$$B_i'(X) = B_i(X) + z(X), \qquad (3.28)$$

где $B_i(X)$ — передаваемая кодовая комбинация; z(*X*) — полином (вектор) ошибки, имеющий степень от 1 до $n - 1$.

При декодировании принятое кодовое слово делится на

$$\frac{B_i'(X)}{G(X)} = U_i(X) + S_i(X), \qquad (3.29)$$

где остаток от деления $S_i(X)$ и является синдромом.

Если при делении получается нулевой остаток $S_i(X) = 0$, то выносится решение об отсутствии ошибки z(*X*) = 0. Если остаток (синдром) ненулевой $S_i(X) \neq 0$, то выносится решение о наличии ошибки и определяется шумовой вектор (полином) z(*X*), а затем — передаваемое кодовое слово, поскольку из (3.2.28) следует

$$B_i(X) = B_i'(X) + z(X). \qquad (3.30)$$

Всякому ненулевому синдрому соответствует определённое расположение (конфигурация) ошибок. Взаимосвязь между видом синдрома и местоположением ошибочного символа находится довольно просто. Достаточно в любую разрешённую кодовую комбинацию ввести ошибку и выполнить деление на G(*X*). Полученный остаток (3.29) — синдром и будет указывать на ошибку в этом символе.

В качестве примера для ЦК Хемминга (7,4), позволяющего исправлять однократную ошибку при $d_{min} = 3$ (см. табл. 3.2), взаимосвязь между синдромом и ошибочным символом для двух возможных порождающих полиномов кода (7,4) приведена в табл. 3.4. Пользуясь этой таблицей, можно найти местоположение ошибки и исправить её.

Для параметров рассмотренного ранее примера 1, где была показана процедура кодирования кодовой комбинации $A_i = 1001$ при использовании порождающего полинома $G(X) = X_3 + X + 1$ для кода Хемминга (7,4), исправляющего однократную ошибку, приведём в примере 2 процедуру декодирования принятой с помехой кодовой комбинации.

Пример 2. Принятая кодовая комбинация ЦК (7,4) имеет вид $B_i'(X) = 1011110$. Определить и исправить ошибку в $B_i'(X)$, если она имеется.

Выполним три необходимые операции, проводимые при декодировании:

Таблица 3.4 — Синдромы и ошибки

№ символа комбинации со старшего разряда	Ошибочный символ полинома комбинации	Синдром для порождающего полинома $G(X) = X^3+X+1$	Синдром для порождающего полинома $G(X) = X^3+X^2+1$	Шумовой вектор $z(X)$
7	X^6	101	110	1000000
6	X^5	111	011	0100000
5	X^4	110	111	0010000
4	X^3	011 $= H_k(7,4)$	101 $= H_k(7,4)$	0001000
3	X^2	100	100	0000100
2	X^1	010	010	0000010
1	X^0	001	001	0000001
	Нет ошибки	000	000	0000000

1) в соответствии с алгоритмом (3.29) производим деление

$$B_i'(X)/G(X) = X^6 + X^4 + X^3 + X^2 + X \;\Big|\; \underline{X^3 + X + 1}$$
$$\underline{+X^6 + X^4 + X^3} \qquad\qquad X^3$$

$X^2 + X$ — остаток при делении

$R(X) = X^2 + X \sim 110.$

Отметим, что совпадение остатков в примере 1 и 2 — чисто случайное, в примере 1 остаток являлся проверочной группой кода, а в примере 2 — синдромом;

2) по полученному синдрому 110 в соответствующем опознавателе синдрома (дешифраторе синдрома, локаторе ошибки) определяем вид шумового вектора $z(X)$ 0010000 (см. табл. 3.3);

3) воспользовавшись алгоритмом (3.2.30), исправляем принятую кодовую комбинацию $B_i'(X)$ и получаем переданную комбинацию $B_i(X)$:

$$B(X) = B_i'(X)+z(X) = \begin{array}{l} 1011110 \\ \underline{0010000} \\ 1001110 \end{array}$$

— исправленная комбинация на выходе.

3.2.2.3 Построение порождающих и проверочных матриц циклических кодов

Наряду с полиномиальным способом задания кода, структуру построения кода можно определить с помощью матричного представления. При этом в ряде случаев проще реализуется построение кодирующих и декодирующих устройств ЦК.

Рассмотрим варианты формирования и обработки ЦК, заданных в виде порождающих и проверочных матриц, на конкретном примере ЦК Хемминга (7, 4), воспользовавшись выражением (3.24), в котором определены двойственные (дуальные) порождающие полиномы кода:

$X^7+1 = (X+1)(X^3+X+1)(X^3+X^2+1)$, что соответствует кодам (7, 6); (7, 4); (7, 4).

Пример 3. Задан ЦК (7,4) дуальными порождающими полиномами $G(7,4) = X^3+X+1$ и $G\sim(7,4)= X^3 + X^2 +1$.

Составить порождающие матрицы для формирования разрешённых кодовых комбинаций и проверочные матрицы для получения синдромов.

Первой строкой в матрице записывается порождающий полином (в двоичном представлении) с домножением его на оператор сдвига X^r для резервирования места под запись $r = 3$ проверочных символов. Следующие $k - 1$ строк матриц получаются путём последовательного циклического сдвига базового кодового слова матрицы G и $G\tilde{}$ на одну позицию вправо, поскольку при этом по определению ЦК также получаются разрешённые к передаче кодовые комбинации:

$$G(7,4) = \left|\begin{matrix}1011000\\0101100\\0010110\\0001011\end{matrix}\right|\begin{matrix}1\\2\\3\\4\end{matrix}; \quad \tilde{G}(7,4) = \left|\begin{matrix}1101000\\0110100\\0011010\\0001101\end{matrix}\right|\begin{matrix}1\\2\\3\\4\end{matrix}. \qquad (3.31)$$

Однако в таком виде эти порождающие матрицы размерностью $k \times n$ — (n столбцов, k строк) могут образовать только неразделимый ЦК, т.е. код, у которого не определены жёстко места информационных и проверочных элементов. Для построения порождающей матрицы, формирующей разделимый блочный код, необходимо матрицу преобразовать к каноническому виду путём простых линейных операций над строками, промаркированными № 1—4.

С учётом свойства ЦК (3.2.25), каноническую форму матрицы можно получить путём сложения ряда разрешённых кодовых комбинаций. Каноническая матрица должна в левой части порождающей ЦК матрицы содержать единичную диагональную квадратную подматрицу порядка «k» для получения в итоге блочного ЦК. С этой целью для получения первой строки канонической матрицы $G_k(7,4)$ необходимо сложить по модулю 2 строки с номерами 1, 3 и 4 матрицы G(7, 4), а для матрицы $\tilde{G}_k$ (7,4) —

строки с номерами 1, 2 и 3 матрицы $\overline{G}$ (7,4). В этом случае в матрицах (3.31) в первых строках остаются «1» только на первых позициях, а остальные «k –1» символов заменяются «0», что и соответствует первым строкам единичных подматриц порядка «k». Нормирование последующих трёх строк канонических матриц производится путём соответствующего суммирования строк матриц (3.31).

В итоге имеем следующий вид дуальных канонических матриц:

$$G_k(7,4) = \begin{vmatrix} 1000|101 & 1 = 1 \oplus 3 \oplus 4 \\ 0100|111 & 2 = 2 \oplus 4 \\ 0010|110 & 3 = 3 \\ 0001|011 & 4 = 4 \end{vmatrix} \qquad \overline{G}_k(7,4) = \begin{vmatrix} 1000|110 & 1 = 1 \oplus 2 \oplus 3 \\ 0100|011 & 2 = 2 \oplus 3 \oplus 4 \\ 0010|111 & 3 = 3 \oplus 4 \\ 0001|101 & 4 = 4 \end{vmatrix} \tag{3.32}$$

Процесс кодирования первичных кодов на стороне источника сообщений сводится к умножению информационных посылок, представленных в виде векторов $\bar{A}_i(X)$, на соответствующую порождающую каноническую матрицу:

$$\overline{B}_i(X) = \bar{A}_i(X) \cdot G_k. \tag{3.33}$$

Эта процедура позволяет получить блочные коды Хемминга «в целом», т.е. получить проверочную группу символов r_1, r_2, r_3 сразу после выполнения операции (3.33).

При матричном варианте обработки принятых кодов на стороне получателя сообщений для получения синдрома необходимо принятую, возможно искажённую в канале, кодовую комбинацию $\overline{B_i'}(X)$ умножить на проверочную матрицу $H(X)$:

$$\overline{S} = \overline{B_i'}(X) \cdot H(X). \tag{3.34}$$

Заметим, что матрица H с размерностью n x r может быть получена из порождающей матрицы канонического вида (3.2.32) путём дополнения проверочной подматрицы единичной матрицей размерности r x r, что даёт следующий вид дуальных проверочных матриц:

$$H_k(7,4) = \begin{vmatrix} 101 \\ 111 \\ 110 \\ 011 \\ 100 \\ 010 \\ 001 \end{vmatrix}; \qquad \overline{H}_k(7,4) = \begin{vmatrix} 110 \\ 011 \\ 111 \\ 101 \\ 100 \\ 010 \\ 001 \end{vmatrix}. \tag{3.35}$$

По определённому с помощью полученного синдрома (3.34) соответствующему шумовому вектору исправляются ошибки (3.30).

Интересно отметить, что в табл. 3.3, в которой рассмотрена связь между синдромом и шумовым вектором для кода (7,4), колонки с синдромами дуальных порождающих полиномов полностью совпадают с (3.35).

В ЦК Хемминга (n, k) все проверочные $r = n - k$ разряды размещаются в конце кодовой комбинации и, как отмечалось, формируются «в целом». При поэлементном получении проверочных символов (4.27) целесообразно, чтобы каждый синдром представлял собой двоичное число, указывающее на номер разряда, в котором произошла ошибка. Коды, в которых синдромы (опознаватели) соответствуют указанному принципу, и предложил впервые Хемминг. В этом случае для кода (7, 4) проверочные символы r_1, r_2, r_3 (табл. 2) размещаются на первой, второй и четвертой позициях кодовой комбинации, отсчитываемых справа налево.

Такое построение кодов упрощает декодирующее устройство на стороне получателя сообщений.

3.2.2.4 Циклические коды Боуза—Чоудхури—Хоквингема

Коды Боуза—Чоудхури—Хоквингема (БЧХ) представляют собой обширный класс кодов, способных исправлять несколько ошибок и занимающих заметное место в теории и практике кодирования. Интерес к кодам БЧХ определяется по меньшей мере тремя следующими обстоятельствами:

1) среди кодов БЧХ при небольших длинах существуют хорошие (но, как правило, не лучшие из известных) коды;

2) известны относительно простые и конструктивные методы их кодирования и декодирования;

3) полное понимание построения кодов БЧХ является наилучшей отправной точкой для изучения многих других классов кодов.

Коды БЧХ независимо открыли Хоквингем (1959) и Боуз и Рой—Чоудхури (1960), которые доказали ряд теорем, устанавливающих существование таких ЦК, у которых минимизируется число проверочных символов, а также указывающих пути нахождения порождающих полиномов для этих кодов.

Корректирующие свойства ЦК могут быть определены на основании следующей теоремы: для любых значений m и g_u существует ЦК длиной $n = 2^m - 1$, исправляющий все ошибки кратности g_u и менее ($g_u < m$) и содержащий не более $n - k \le mg_u$ проверочных символов. Так, например, при $n = 15$, $m = 4$ и различных g_u число проверочных символов будет равно: $g_u = 1$, $n - k = m \cdot g_u = = 4 \cdot 1 = 4$; $g_u = 2$, $m \cdot g_u = 4 \cdot 2 = 8$; $g_u = 3$, $m \cdot g_u = 4 \cdot 3 = 12$. Соответствующие коды ($n$, k) будут (15,11), (15,7), (15,3). Напомним, что минимальное кодовое расстояние $d_{min} = 2 \cdot g_u + 1$ и применительно к ЦК оно чаще называется конструктивным расстоянием. Если величины g_u или d выбрать слишком большими, то получившийся в результате код будет

тривиальным — в нём будет лишь один либо (при $r > n$) ни одного информационного символа.

В табл. 3.5 даны параметры и порождающие полиномы некоторых кодов БЧХ. Полиномы приведены в восьмеричной форме записи, старшая степень расположена слева.

Например, коду (15, 7) соответствуют двоичное представление 111010001 и многочлен $G(X) = X^8 + X^7 + X^6 + X^4 + 1$. Подробные таблицы порождающих полиномов циклических кодов БЧХ приведены в [6].

Таблица 3.5 — Параметры кодов БЧХ

m	n	k	r	g_u	$G(X)$-mod 8
3	7	4	3	1	13
4	15	11	4	1	23
		7	8	2	721
5	31	26	5	1	45
		21	10	2	3551
		16	15	3	107657
		11	20	5	5423325
6	63	57	6	1	103
		51	12	2	12471
		45	18	3	1701317
		39	24	4	166623567
		36	27	5	1033500423
7	127	120	7	1	211
		113	14	2	41567
		106	21	3	11554743
		99	28	4	3447023271
		92	35	5	624730022327
8	255	247	8	1	435
		239	16	2	267543
		231	24	3	156720665
		223	32	4	75626641375
		215	40	5	23157564726421

Коды БЧХ с длиной $2^m - 1$ называют примитивными кодами БЧХ. К ним, в частности, относятся классические коды Хемминга, исправляющие однократные ошибки. К кодам БЧХ относятся также коды, длина n которых является делителем $2^m - 1$. Например, код Голея (23, 12, 7) также принадлежит классу кодов БЧХ, поскольку при $m = 11$ примитивный код БЧХ имеет длину $n = \quad = 2^{11} - 1 = 2047$, причём это значение без остатка делится на длину кода Голея $n = 23$ (2047: 23 = 89), который относится к непримитивным БЧХ-кодам [5, 6].

Все примитивные коды БЧХ обладают конструктивным расстоянием $d_{min} \geq 2g_u + 1$. Расстояние можно увеличить до $2g_u + 2$. Для этого нужно основной порождающий полином БЧХ-кода домножить на бином $X + 1$, т.е. $G_1(X) = (X + 1) \times G_{БЧХ}(X)$, что повлечёт за собой прибавление к коду одного проверочного символа, обеспечивающего проверку на чётность всех символов БЧХ-кода. Таким образом получается расширенный БЧХ-код.

Коды Рида—Соломона (РС) являются важным и широко используемым подмножеством кодов БЧХ. Двоичный код Рида—Соломона получится, если взять основание кода $q = 2^s$. Это означает, что каждый символ кода заменяется s-значной двоичной последовательностью. Если исходный код с основанием q исправляет ошибки кратности $< g_u$, то

полученный из него двоичный код имеет $2g_u \cdot s$ проверочных символов (по $2g_u$ на каждый блок из символов) из общего числа $n = s \cdot (2^s - 1)$. Код может исправлять серийные ошибки (пакеты ошибок) длиной $\leq b = s \cdot (g_u - 1) + 1$.

Коды РС, наряду с кодами Файра, являются наиболее подходящими для исправления серийных ошибок, а также в каскадных системах кодирования в качестве внешних кодов.

3.3 RAID-системы

Исследования надежности различных вычислительных архитектур на основе микропроцессоров показали, что до 80 % отказов в таких системах происходит по причине сбоев в памяти системы и по причине выхода из строя накопителя на жестких дисках. Данная проблема стимулировала появление интереса разработчиков к построению мощных отказоустойчивых накопителей на жестких магнитных дисках с высокой скоростью доступа — raid-систем.

В декабре 1987 года коллектив авторов в составе Дэвида Паттерсона, Гарта Гибсона и Рэнди Катца из Калифорнийского университета Беркли опубликовали статью «A Case for Redundant Arrays of Inexpensive Discs (RAID)» СНОСКА. В статье описывались основные конфигурации (уровни) избыточных массивов независимых (или недорогих) дисков — RAID.

Выделялись три основных признака RAID-массивов:

– набор физических дисков с точки зрения пользователя представляет собой единый виртуальный диск большой емкости;

– данные распределены по набору дисков;

– в наборе имеется избыточная емкость, обеспечивающая возможность восстановления данных при отказе одного или нескольких дисков.

Авторы определяли пять разных способов организации, или уровней, от RAID 1 до RAID 5. Каждый уровень RAID характеризуется своим способом распределения данных и способом использования избыточной емкости.

Для стандартизации RAID в 1992 году был создан промышленный консорциум — Комиссия советников по RAID (RAID Advisory Board — RAB, http://www.raid-advisory.com).

В настоящее время комиссией стандартизировано 8 уровней объединения дисков в массивы: от RAID-0 до RAID-7. Номера уровней определены в порядке, в котором были предложены различные варианты, и не связаны с характеристиками RAID. Применяются также комбинированные уровни, например уровень 0+1 означает RAID уровня 0, но в этот RAID объединены не одиночные диски, а несколько RAID уровня 1 (несколько зеркальных дисков).

3.3.1 Общие понятия и принципы функционирования

Прежде чем переходить к описанию каждого из уровней RAID в отдельности, определим основные используемые понятия.

Чередование (striping) — данные записываются на диски полосами, каждая полоса состоит из блоков, каждый блок помещается на отдельный диск в массиве. Число дисков в массиве называется шириной полосы (stripe width). Это позволяет ускорить запись/чтение файлов за счет параллельного проведения операций ввода/вывода.

Зеркалирование (mirroring) — данные с одного диска копируются на другой диск. В случае отказа одного диска информация по-прежнему будет доступна с другого.

Дуплексирование (duplexing) — дублирование различных аппаратных компонентов, например контроллеров и дисков; повышает общую отказоустойчивость системы.

Контроль четности (parity) — данные записываются на диски полосами, каждая полоса состоит из блоков, каждый блок помещается на отдельный диск в массиве. Число дисков в массиве называется шириной полосы (stripe width). Суммирование блоков с помощью операции «исключающего или» (xor) дает блок четности: parity=block1 xor block2 xor... xor blockN. Эта операция обратима и позволяет восстановить данные при потере одного из блоков, позволяет восстановить информацию в случае отказа одного из дисков.

Основные задачи, которые позволяют решить RAID, — это обеспечение отказоустойчивости дисковой системы и повышение ее производительности.

Отказоустойчивость достигается тем, что вводится избыточность. В RAID объединяется больше дисков, чем это необходимо для получения требуемой емкости.

Производительность дисковой системы повышается за счет того, что современные интерфейсы (в частности, SCSI) позволяют осуществлять операции записи и считывания фактически одновременно с несколькими дисками. Поэтому в первом приближении можно рассчитывать, что скорость записи или чтения в случае применения RAID увеличивается пропорционально количеству дисков, объединяемых в RAID.

Возмзжность одновременной работы с несколькими дисками можно реализовать двумя способами: с использованием параллельного доступа (parallel-access array) и с использованием независимого доступа (independent-access array).

Для организации параллельного доступа рабочее пространство дисков размечается на зоны определенного размера (блоки) для размещения данных и избыточной информации. Информация, подлежащая записи на диски (запрос на обслуживание), разбивается на такие же по величине блоки, и каждый блок записывается на отдельный диск. При

поступлении запроса на чтение необходимая информация собирается из нескольких блоков.

Понятно, что в этом случае скорость записи (равно как и скорость чтения) увеличивается пропорционально количеству дисков, объединенных в RAID.

Для организации независимого доступа рабочее пространство дисков также размечается на зоны определенного размера (блоки). Однако, в отличие от предыдущего случая, каждый запрос на запись или чтение обслуживается только одним диском.

Естественно, в этом случае скорость записи будет не выше, чем при работе с одним диском. Однако массив с независимым доступом в каждый момент времени может обслуживать одновременно несколько запросов, каждый диск обслуживает свой запрос.

Таким образом, оба архитектурных решения способствуют повышению производительности, но механизм повышения производительности у этих решений различен. Соответственно, свойства RAID существенно зависят от того, какой из этих двух механизмов в нем используется. Именно поэтому при сравнении RAID различного уровня в первую очередь необходимо сравнивать размер логических блоков. Точнее говоря, не собственно размер, а соотношение размера блока и величины запроса на обслуживание (объем информации, подлежащей записи или считыванию).

Другим фактором, влияющим на производительность, является способ размещения избыточной информации. Избыточная информация может храниться на специально выделенном для этого диске и может распределяться по всем дискам.

И наконец, в RAID различного уровня применяются различные способы вычисления избыточной информации. Это также влияет на характеристики RAID (надежность, в первую очередь, производительность и стоимость). Основные способы: полное дублирование информации, применение кодов с коррекцией ошибок (применяется код с коррекцией одиночных ошибок и обнаружением двойных ошибок ECC — код Хемминга) и вычисление четности (Parity).

3.3.2 Уровни RAID

Рассмотрим особенности уровней RAID. Будем считать, что RAID-массив представляет собой набор физических дисков, связанных с компьютером, и не содержит никакой дополнительной логики, кроме минимально необходимой для обеспечения обмена данными.

RAID-0

Эта конфигурация была предложена Digital Equipment Corporation и реализована в одной из подсистем ОС VMS на машине VAX 11. RAID-0 не

имеет избыточной емкости. Выход из строя любого диска приводит к невозможности доступа к данным и может повлечь их потерю.

Распределение данных по физическим дискам осуществляется следующим образом. Совокупное пространство памяти разбивается на блоки равной длины. Пронумеруем эти блоки 0, 1, 2 и т.д. Блок 0 располагается на первом физическом диске, Блок 1 — на втором диске в том же самом месте, что и Блок 0 на первом, Блок 2 — на третьем, а Блок 3 — на четвертом. Пятый блок размещается опять на первом диске вслед за нулевым блоком, шестой — на втором вслед за блоком 1 и т.д. до тех пор, пока все блоки не будут распределены по всем физическим дискам.

Диск 1 Диск 2 Диск 3 Диск 4

Блок 0 Блок 1 Блок 2 Блок 3

Блок 4 Блок 5.....

.... Блок N Блок N+1 Блок N+2

Эффективность распределения данных будет зависеть от соотношения размера блока и длины читаемого/записываемого файла. Допустим, размер блока больше по отношению к размеру файла. При записи или чтении обращение происходит только к этой части блока, где расположен файл. Все запросы с таким соотношением размера файла к размеру блока будут относиться к единственному блоку. При чтении данных виртуальный адрес отображается в физический адрес блока на диске, содержащем данные. Данные читаются из этого блока. Аналогичные действия выполняются и при записи. Поскольку в этом случае почти каждому запросу будет соответствовать единственный блок, то разные диски в массиве смогут обслуживать разные запросы параллельно. Этим обеспечивается хорошая производительность для запросов небольших и средних размеров.

Рассмотрим другой вариант. Размеры запросов велики, а размер блока мал по отношению к размерам запроса. В этом случае запрос распределен по всем дискам, и для его обслуживания диски будут передавать информацию одновременно. Скорость при передаче запросов большого объема оказывается больше, но число запросов, обслуживаемых в единицу времени, ограничено, так как в каждый момент времени может быть обслужен только один запрос. При таком типе распределения RAID-0 обеспечивает хорошую производительность для последовательного потока больших запросов.

Из-за отсутствия необходимости выделять место для записи контрольной информации реальная емкость массива равна суммарной емкости составляющих его дисков. Однако при отказе одного из дисков вся информация будет потеряна.

RAID 1

Технология, на которой основана конфигурация RAID 1, называется «зеркальные диски». Как и ранее, для пользователя массив представляет собой виртуальное дисковое устройство с необходимым объемом памяти, разбитым на блоки. Данные размещаются точно так же, как на традиционных дисках, то есть Блок 0, Блок 1, Блок 2,... размещены в естественном порядке на первом диске. Когда диск заполнен, данные размещаются в таком же порядке на третьем диске и так далее до исчерпания всего объема памяти. Избыточность обеспечивается с помощью второго диска, который является точной копией первого, т.е. сегменты Блок 0, Блок 1, Блок 2,... расположены на нем в тех же самых местах, что и на первом.

Диск 0 Диск 1

Блок 0 Блок 0

Блок 1 Блок 1

.....

Блок N Блок N

При поступлении в устройство запроса на чтение данные могут быть прочитаны с любого из двух дисков, так что каждая зеркальная пара дисков в среднем исполняет только половину запросов. При записи данные записываются в оба диска.

Если один из дисков выходит из строя, доступ не прерывается, не происходит и потерь данных. Все операции выполняются оставшимся работоспособным диском.

Если использовать более двух дисков для данной конфигурации, то получим вариант RAID 0+1. Отличие состоит в том, что данные распределены по дисковому массиву аналогично тому, как это делается в конфигурации RAID-0.

Диск 0 Диск 1 Диск 2 Диск 3

Блок 0 Блок 0 Блок 1 Блок 1

Блок 2 Блок 2 Блок 3 Блок 3

.....

Блок N Блок N Блок N+1 Блок N+1

Блок 0 располагается на первом и втором дисках, следующий Блок 1 — на третьем и четвертом дисках. Блок 2 располагается на первом и втором диске следом за Блоком 0. Блок 3 размещается на третьем и четвертом дисках следом за Блоком 1 и так далее.

Выберем размер блока достаточно большой по сравнению с размерами запросов пользователя. Если поступает запрос на чтение, доступ осуществляется только к тем блокам в одном из двух дисков, в которых находятся требуемые данные. Если это запрос на запись, обращение происходит также только к тем блокам, адрес которых соответствует логическому адресу в запросе. Производительность схемы заметно выше, чем в ранее описанной конфигурации. Массив из дисков «зеркальной пары» может исполнять запросы на чтение одновременно.

Если размер блока меньше, чем размер запроса, то для последовательности запросов большого размера можно получить производительность, близкую к производительности RAID 0. Рабочая нагрузка может быть сбалансирована между дисками путем подбора размера сегмента таким образом, чтобы обеспечить в среднем наилучшую производительность.

Данная конфигурация предназначена в основном для обеспечения отказоустойчивости системы. За счет полного дублирования информации обеспечивается очень высокий уровень надежности. Однако и стоимость хранения информации получается немалой.

4 ТЕХНОЛОГИЯ ПРЕОБРАЗОВАНИЯ ИНФОРМАЦИИ

Немаловажную роль в информационных технологиях играет преобразование информации. Мы сталкиваемся с этим каждый день, когда разговариваем, смотрим, ощущаем, все полученные нами сигналы преобразуются в доступный для нашего мозга вид.

В промышленности для выполнения и контроля какого-либо технологического процесса обычно используется ЭВМ. Согласно статистике около 40 % всех микроЭВМ используются в качестве управляющих машин. Как мы знаем, ЭВМ работает с электрическими сигналами, поэтому естественные (физической природы) или унифицированные сигналы необходимо преобразовывать в электрические. На рис. 4.1 приведена типичная схема замкнутой системы управления с использованием ЭВМ.

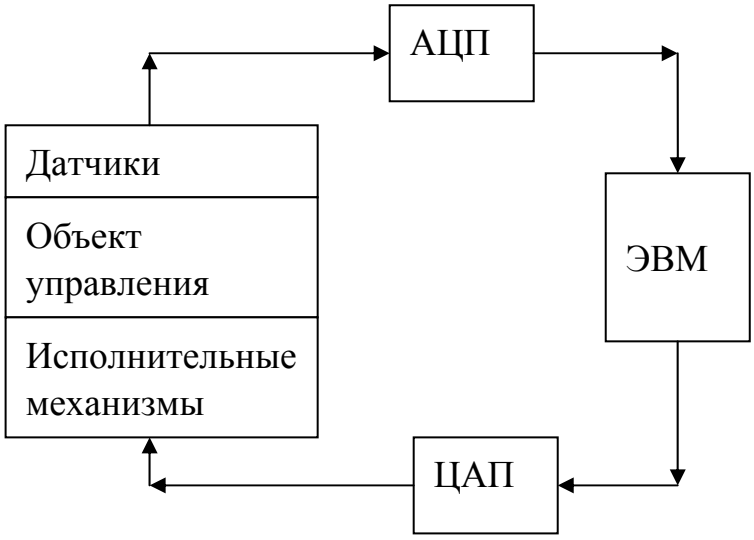


Рисунок 4.1 — Схема замкнутой системы управления с использованием ЭВМ

Параметрами, характеризующими объект управления, могут быть температура, давление, ток, напряжение и т.д. С помощью датчиков информация об этих параметрах представляется в аналоговом виде. Для дальнейшей обработки информации на ЭВМ эти сигналы нужно преобразовать в цифровую форму. Такое преобразование осуществляет аналого-цифровой преобразователь (АЦП). ЭВМ обрабатывает полученную информацию и вырабатывает сигналы управления в цифровой форме. Чтобы эти сигналы воздействовали на исполнительные механизмы, они должны быть представлены в аналоговой форме, такое преобразование осуществляет цифроаналоговый преобразователь (ЦАП).

4.1 Цифроаналоговые преобразователи

Как уже было сказано ранее, ЦАП преобразует цифровые сигналы в аналоговые. При этом каждому значению цифрового сигнала в двоичной форме соответствует его аналоговое значение в единицах напряжения.

Большинство обычно используемых структур ЦАП (отличных от простого одноразрядного ЦАП, основанного на одном коммутаторе с использованием опорного напряжения) являются двоичными взвешивающими ЦАП или многозвенными схемами лестничного типа. Данные схемы, хотя и являются несложными по структуре, требуют весьма тщательного анализа. Мы рассмотрим одну из простейших структур — делитель Кельвина, представленный на рис. 4.2.

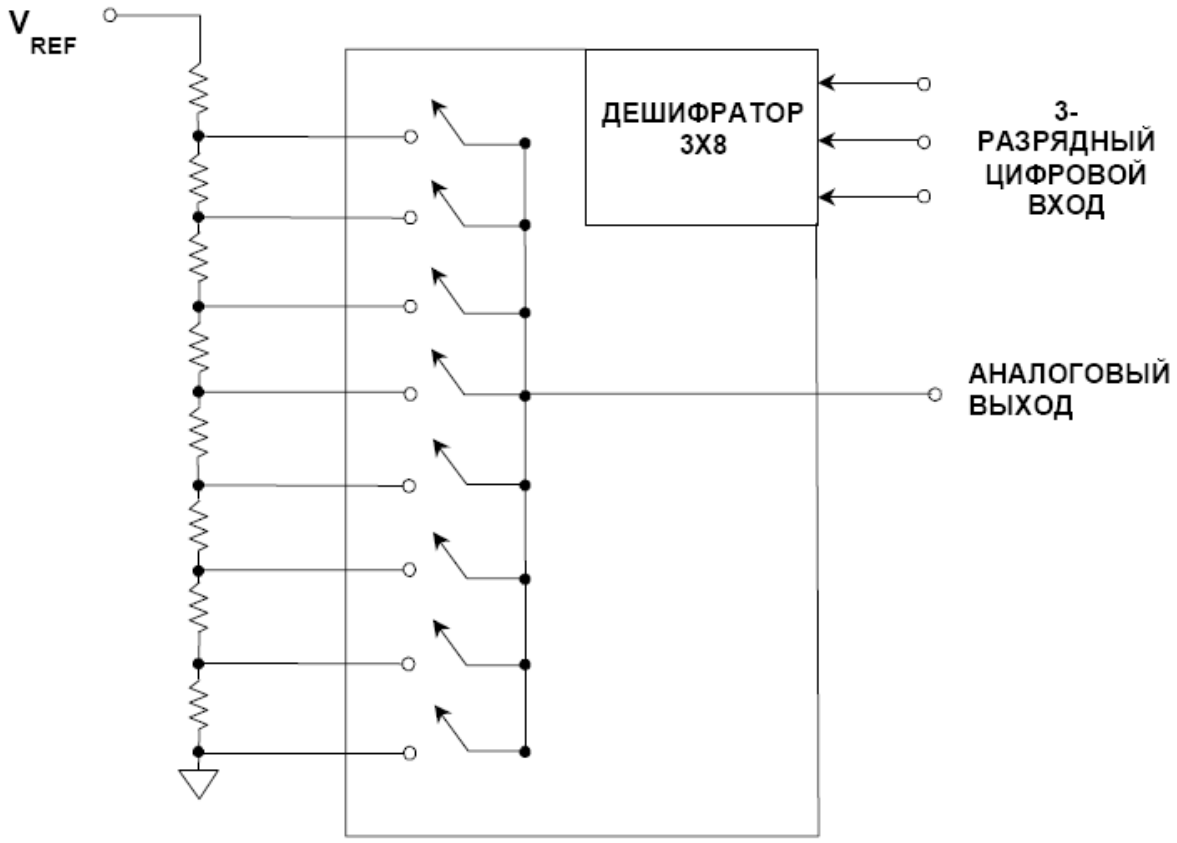


Рисунок 4.2 — Простейшая схема ЦАП с выходом напряжения

N-разрядная версия этого ЦАП просто содержит 2^N равных по величине последовательно соединенных резисторов. Выходной сигнал снимается с соответствующего отвода замыканием одного из 2^N коммутаторов после декодирования N-разрядных данных. Современные ЦАП, использующие эту архитектуру, называются строковыми ЦАП.

Эта архитектура проста, имеет выход с изменяющимся значением выходного напряжения и изначально обеспечивает монотонный сигнал. Архитектура линейна, если все резисторы равны по значению, но может быть преднамеренно сделана нелинейной, если требуется нелинейный ЦАП. Так как в момент переключения работают только два коммутатора, эта архитектура обладает малым ложным сигналом (low-glitch).

Ее главным недостатком является большое количество резисторов, требуемых для обеспечения высокой разрешающей способности, поэтому в качестве отдельного устройства она обычно не используется, но, как мы увидим позже, применяется в роли компонента более сложных структур ЦАП.

Существует аналогичный ЦАП с токовым выходом, который также состоит из 2^N резисторов, или источников тока, но подключенных теперь параллельно между входом опорного напряжения и виртуальным заземленным выходом (рис. 4.3).

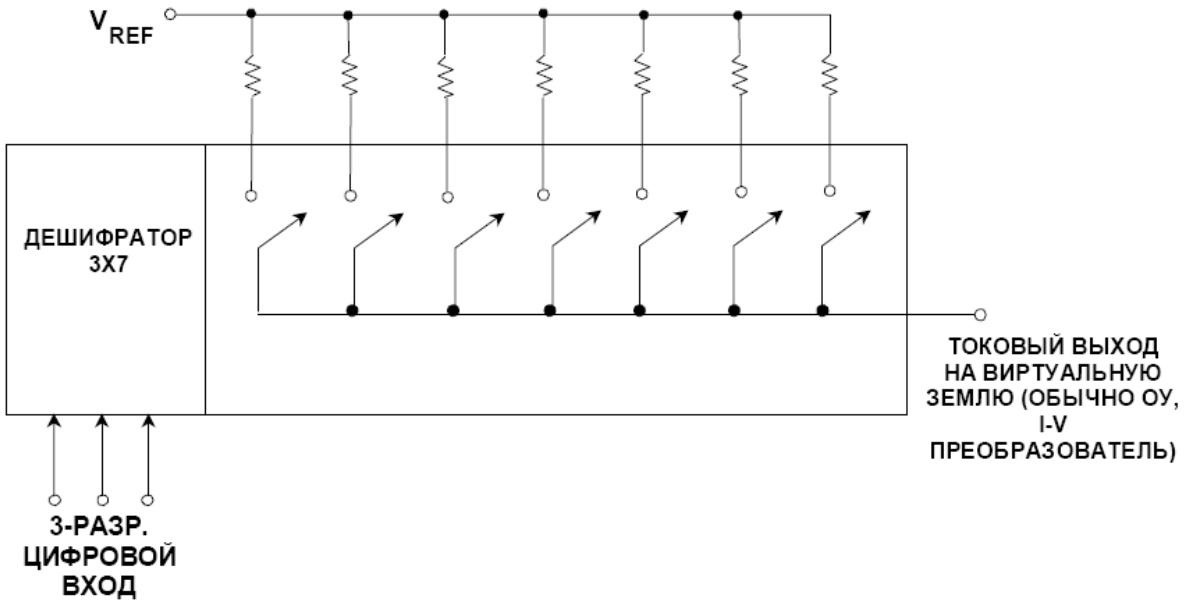


Рисунок 4.3 — Простейшая схема ЦАП с токовым выходом

В данном ЦАП, как только какой-либо резистор подключается к цепи, любые дальнейшие увеличения цифрового кода уже не могут его отключить. Таким образом, структура является изначально монотонной, независимо от погрешностей резисторов и, подобно предыдущему случаю, может быть сделана преднамеренно нелинейной там, где эта нелинейность требуется. Опять, как и в предыдущем случае, архитектура является редкостью, так как, если попытаться ее использовать для изготовления полного ЦАП, потребуется большое количество резисторов и коммутаторов. Но опять же она часто используется в качестве компонента в ЦАП более сложной структуры.

В отличие от делителя Кельвина этот тип ЦАП не имеет уникального названия, хотя оба типа упомянуты как полно-декодирующие (fully decoded) ЦАП, ЦАП типа «столбик термометра» (thermometer) или строковые (string) ЦАП.

Полнодекодирующие ЦАП часто используются как компоненты более сложных ЦАП. Наиболее популярными являются сегментные ЦАП, где часть выходного сигнала полно-декоди-рующего ЦАП в дальнейшем вновь поступает на делитель. Данная структура используется потому, что полно-декодирующий ЦАП изначально монотонен, так что если последующий делитель тоже монотонен, в целом является таковым же и результирующий ЦАП.

В сегментных ЦАП с выходом по напряжению (рис. 4.4) сигнал подается с одного из резисторов делителя Кельвина на новый делитель Кельвина (в этом случае полная структура известна как «делитель Кельвина—Варлея») или на ЦАП какой-либо другой структуры.

Во всех ЦАП выходной сигнал представляет собой результат комбинации опорного напряжения и цифрового кода. В этом смысле все

ЦАП являются перемножающими, но многие из них хорошо работают только в ограниченном диапазоне Vref.

Настоящие перемножающие ЦАП (MDAC) ориентированы на работы в широком диапазоне Vref. Строгое определение перемножающего ЦАП требует, чтобы его диапазон опорного напряжения включал 0 В, и многие схемы, особенно лестничного типа с токовым режимом и с переключателями CMOS, допускают положительное, отрицательное и переменное значение Vref. ЦАП, которые не работают при значении Vref = 0 В, тоже полезны, и их типы, допускающие изменение значения Vref в пропорции 10:1 или около того, часто относят к перемножающим ЦАП (MDAC), хотя более точно их можно было бы назвать полуперемножающими ЦАП.

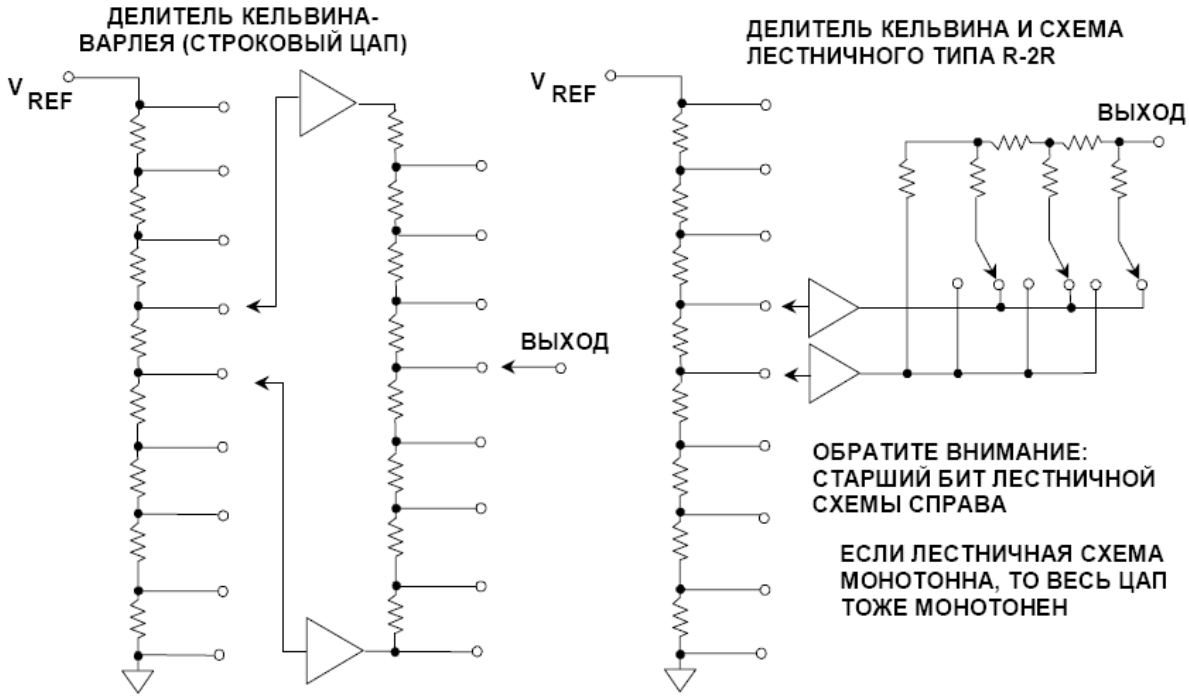


Рисунок 4.4 — Делитель Кельвина—Варлея

4.2 Аналого-цифровые преобразователи

Преобразование аналоговых сигналов в цифровые осуществляется с помощью аналого-цифровых преобразователей (АЦП) и является измерительным процессом, основанным на сравнении аналогового сигнала с эталонным напряжением, значение которого известно с высокой степенью точности.

Наиболее популярные АЦП для приложений цифровой обработки сигналов (ЦОС) базируются на пяти основных архитектурах: АЦП последовательного приближения, сигма-дельта АЦП, АЦП параллельной обработки (flash), АЦП конвейерной обработки (pipelined) и АЦП последовательного счета (Bit-Per-Stage). Мы же рассмотрим АЦП последовательного приближения и АЦП параллельной обработки.

4.2.1 АЦП последовательного приближения

АЦП последовательного приближения много лет были главным инструментом преобразования сигнала. Недавние усовершенствования разработчиков расширили диапазон частот дискретизации этих АЦП до мегагерц. Использование методов внутренних коммутируемых конденсаторов вместе с методами автокалибровки расширяет разрешающую способность этих АЦП до 16 разрядов на стандартных CMOS-процессах без необходимости в дорогой тонкопленочной лазерной подстройке.

Основные элементы АЦП последовательного приближения представлены на рис. 4.5.

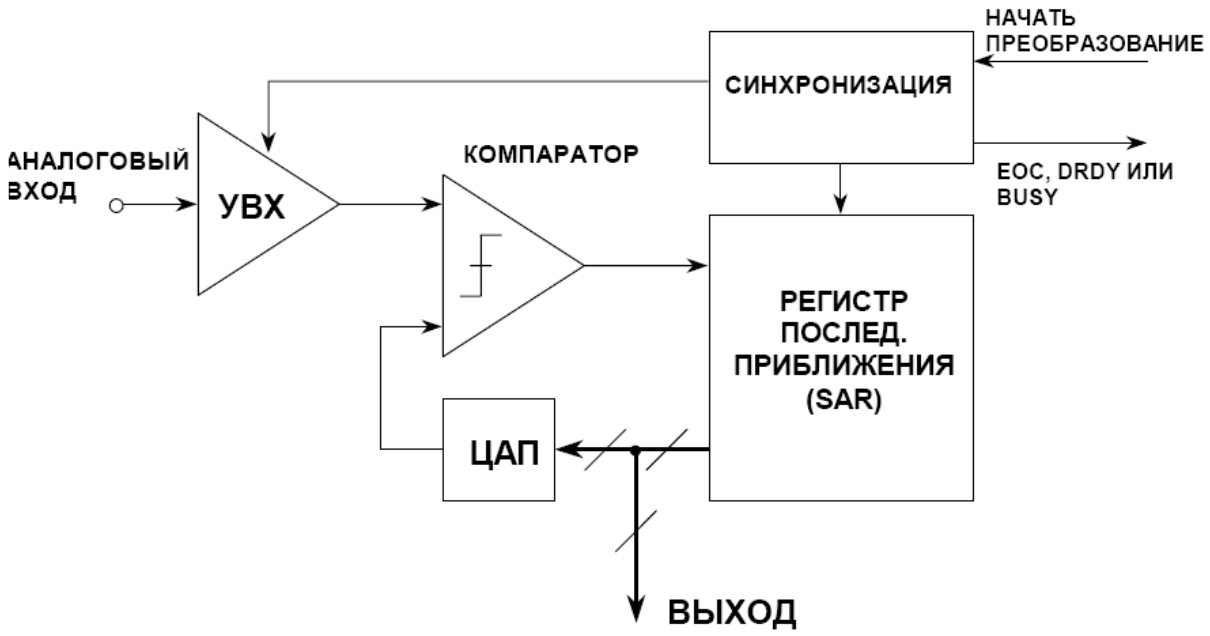


Рисунок 4.5 — АЦП последовательного приближения

Этот АЦП выполняет преобразования в командном режиме. После установки устройства выборки-хранения УВХ (SHA) в режим хранения все разряды регистра последовательного приближения РПП (SAR) сбрасываются в «0», кроме старшего значащего разряда (MSB), который устанавливается в «1». Выходной сигнал регистра последовательного приближения (РПП) подается на внутренний ЦАП. Если выходной сигнал ЦАП больше, чем аналоговый входной сигнал, старший разряд РПП сбрасывается, в противном случае он остается установленным. Затем следующий старший значащий разряд устанавливается в «1». Если сигнал на выходе ЦАП больше, чем аналоговый входной сигнал, старший разряд РПП сбрасывается, в противном случае бит остается установленным. Описанный процесс поочередно повторяется для каждого разряда. Когда все разряды в соответствии с входным сигналом будут установлены в «0»

или в «1», содержимое регистра последовательного приближения придет в соответствие со значением аналогового входного сигнала и преобразование завершится. Если рассматриваемый АЦП имеет выход в виде последовательного порта, то последовательно поступаемые биты можно непосредственно передавать на выход.

Окончание преобразования индицируется сигналами end-of-convert (EOC), data-ready (DRDY) или BUSY (фактически, отсутствие сигнала BUSY индицирует окончание преобразования). Полярности и наименование этого сигнала могут отличаться для различных АЦП последовательного приближения, но основная концепция сохраняется. В начале интервала преобразования логический уровень сигнала высокий (или низкий) и остается в этом состоянии, пока преобразование не закончено. Затем уровень сигнала становится низким (или высоким). Фронт сигнала индицирует наличие выходных данных.

N-разрядное преобразование осуществляется за N шагов. На первый взгляд может показаться, что 16-разрядному преобразователю для выполнения преобразования требуется в два раза больше времени, чем 8-разрядному преобразователю, но это не так. В 8-разрядном преобразователе перед принятием решения о значении очередного бита ЦАП должен установить на своем выходе сигнал с точностью, соответствующей 8 разрядам, в то время как ЦАП 16-разрядного преобразователя должен установить сигнал на своем выходе с точностью, соответствующей 16 разрядам, что занимает значительно больше времени. На практике 8-разрядный АЦП последовательного приближения может затрачивать на преобразование несколько сотен наносекунд, в то время как 16-разрядному АЦП требуется несколько микросекунд.

Обратите внимание, что общая точность и линейность АЦП последовательного приближения определяется, прежде всего, внутренним ЦАП. До недавнего времени в большинстве прецизионных АЦП последовательного приближения для достижения желательной точности и линейности использовалась тонкопленочная лазерная подгонка. Процесс подстройки тонкопленочного резистора увеличивает стоимость системы, а значение сопротивления тонкопленочного резистора может измениться при механическом воздействии на корпус микросхемы.

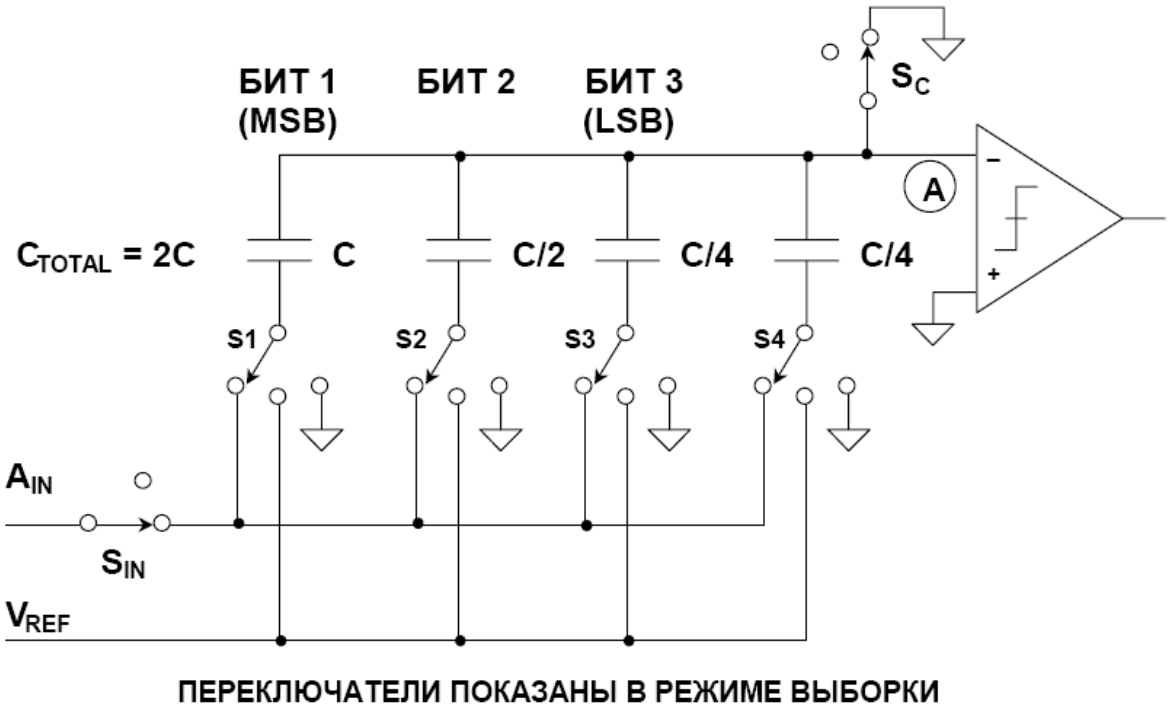


Рисунок 4.6 — Трехразрядный ЦАП с коммутируемыми конденсаторами

По этим причинам в более новых АЦП последовательного приближения стали популярными ЦАП с коммутируемыми конденсаторами (или конденсаторами с перераспределением заряда). Преимущество ЦАП с коммутируемыми конденсаторами состоит в том, что их точность и линейность определяются, прежде всего, качеством фотолитографии, которое, в свою очередь, зависит от площади конденсаторных пластин, емкости и соотношения емкостей конденсаторов. Кроме того, для достижения высокой точности и линейности конденсаторы малой емкости могут подключаться параллельно основным конденсаторам или отключаться от них в соответствии с алгоритмом автокалибровки без необходимости применения тонкопленочной лазерной подстройки.

4.2.2 Параллельные (flash) АЦП

Параллельные АЦП (Flash АЦП) являются самым быстрым типом АЦП, использующим большое количество компараторов, работающих параллельно. N-разрядный параллельный АЦП состоит из 2^N резисторов и $2^N - 1$ компараторов, размещенных как это показано на рис. 4.7.

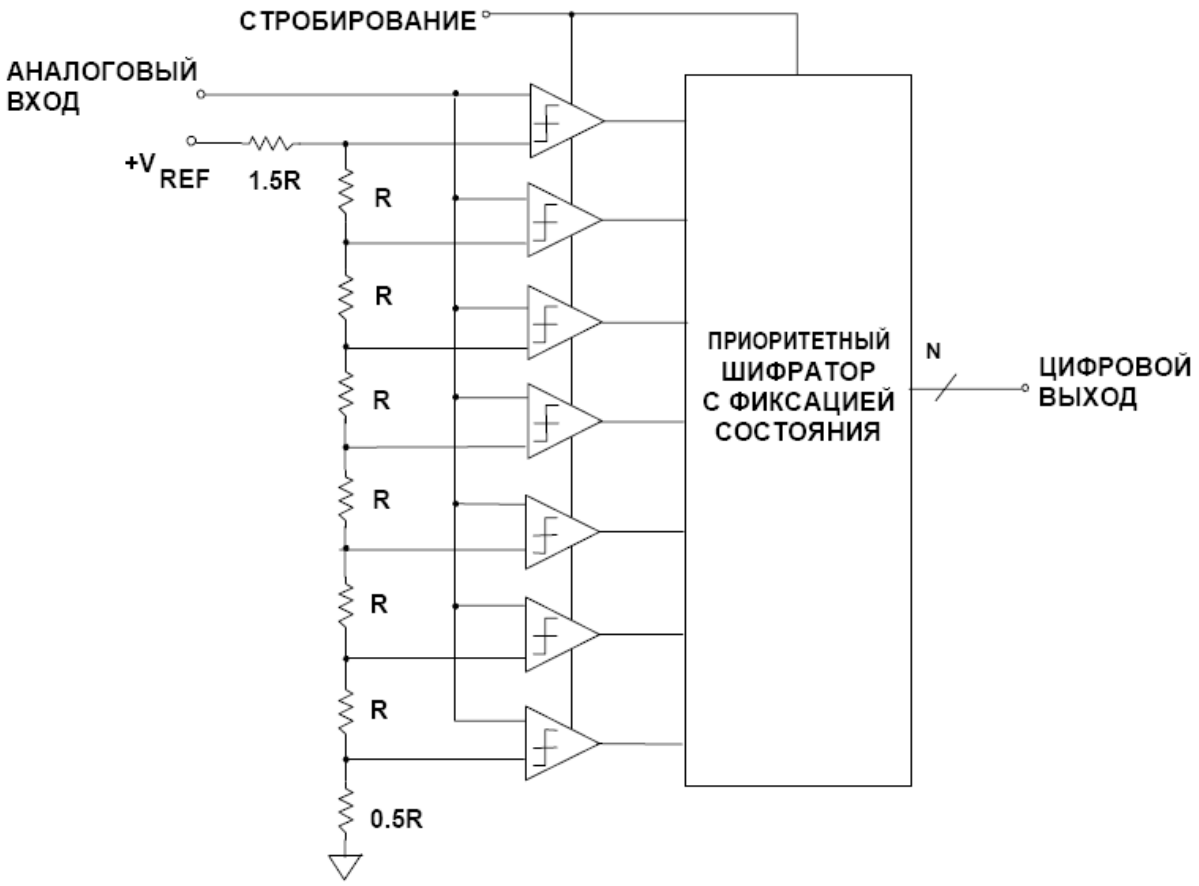


Рисунок 4.7 — Параллельный (flash) АЦП

На каждый компаратор подается опорное напряжение, значение которого для соседних точек отличается на величину, соответствующую одному младшему значащему разряду (LSB) (более старшие разряды — в верхних по схеме элементах). При фиксированном входном напряжении все компараторы, размещенные на схеме ниже некоторой точки, имеют входное напряжение выше опорного напряжения. На их логическом выходе присутствует «1». У всех же компараторов выше этой точки опорное напряжение больше входного, и их логический выход установлен в «0». Поэтому $2^N - 1$ выходов компаратора ведут себя аналогично ртутному термометру, и выходной код такого АЦП иногда называют «кодом термометра». В действительности, было бы непрактично выводить $2^N - 1$ линий данных наружу, поэтому они преобразуются шифратором в N-разрядный двоичный код.

Входной сигнал подается на все компараторы сразу, поэтому «выход термометра» имеет задержку по отношению к входному сигналу, равную задержке только одного компаратора и N-разрядного кодера. Это соответствует задержке нескольких логических элементов, так что процесс преобразования осуществляется очень быстро. Но такая архитектура предполагает использование большого числа резисторов и компараторов, имеет ограничение по максимальной разрешающей способности, и чтобы обеспечить высокое быстродействие, каждый компаратор должен иметь

довольно высокий уровень потребления энергии. Следовательно, к проблемам параллельных АЦП относятся ограниченная разрешающая способность, высокий уровень рассеивания энергии вследствие большого количества высокоскоростных компараторов (особенно на частотах дискретизации больших, чем 50 MSPS) и относительно большие размеры кристалла (и потому — высокая стоимость). Кроме того, для питания быстрых компараторов необходимым током смещения цепочка опорных резисторов должна иметь низкое сопротивление, чтобы этот источник давал весьма большие токи (> 10 мА).

На практике реализуются преобразователи до 10 разрядов, но обычно параллельные АЦП имеют разрешающую способность, соответствующую 8 разрядам. Их максимальная частота дискретизации может достигать 1 ГГц при ширине полосы пропускания по уровню полной мощности более 300 МГц.

Как упоминалось ранее, полоса пропускания по уровню полной мощности не обязательно равна полосе, соответствующей полной разрешающей способности. Идеальный компаратор параллельного преобразователя имеет хорошие характеристики и по постоянному, и по переменному току. Поскольку синхронизирующий строб подается на все компараторы одновременно, параллельный преобразователь автоматически реализует схему выборки-хранения на своем входе. На практике существуют различия в задержках компараторов и другие рассогласования по переменному току, которые вызывают уменьшение эффективного числа разрядов (ENOB) на высоких входных частотах. Это происходит потому, что скорость нарастания сигналов непосредственно на входах сопоставима со временем преобразования компаратора.

Вход параллельного АЦП непосредственно подключается к большому количеству компараторов. Каждый компаратор имеет изменяющуюся в зависимости от напряжения емкость перехода, и наличие этой емкости, зависящей от сигнала, приводит в большинстве параллельных АЦП к уменьшению эффективного числа разрядов (ENOB) и к большим искажениям на высоких входных частотах.

Добавление одного разряда к общей разрешающей способности параллельного преобразователя требует удвоения количества компараторов! Это ограничивает практическую разрешающую способность высокоскоростных параллельных преобразователей до 8 разрядов, так как при более высоких разрешающих способностях слишком велико выделение тепла.

4.3 Цифровые фильтры

Цифровая фильтрация является одним из наиболее мощных инструментальных средств ЦОС. Кроме очевидных преимуществ устранения ошибок в фильтре, связанных с флуктуациями параметров пассивных компонентов во времени и по температуре, дрейфом ОУ (в

активных фильтрах) и т.д., цифровые фильтры способны удовлетворять таким техническим требованиям по своим параметрам, которых, в лучшем случае, было бы чрезвычайно трудно или даже невозможно достичь в аналоговом исполнении. Кроме того, характеристики цифрового фильтра могут быть легко изменены программно. Поэтому они широко используются в телекоммуникациях, в приложениях адаптивной фильтрации, таких, как подавление эха в модемах, подавление шума и распознавание речи.

Процесс проектирования цифровых фильтров состоит из тех же этапов, что и процесс проектирования аналоговых фильтров. Сначала формулируются требования к желаемым характеристикам фильтра, по которым затем рассчитываются параметры фильтра. Амплитудная и фазовая характеристики формируются аналогично аналоговым фильтрам. Ключевое различие между аналоговым и цифровым фильтрами заключается в том, что вместо вычисления величин сопротивлений, емкостей и индуктивностей для аналогового фильтра, рассчитываются значения коэффициентов для цифрового фильтра. Иными словами, в цифровом фильтре числа заменяют физические сопротивления и емкости аналогового фильтра. Эти числа являются коэффициентами фильтра, они постоянно находятся в памяти и используются для обработки (фильтрации) дискретных данных, поступающих от АЦП.

Цифровой фильтр, работающий в реальном масштабе времени, оперирует с дискретными по времени данными в противоположность непрерывному сигналу, обрабатываемому аналоговым фильтром. При этом очередной отсчет, соответствующий отклику фильтра, формируется по окончании каждого периода дискретизации. Вследствие дискретной природы обрабатываемого сигнала на отсчеты данных зачастую ссылаются по их номерам, например отсчет 1, отсчет 2, отсчет 3 и т.д. На рис. 4.8 представлен низкочастотный сигнал, содержащий высокочастотный шум, который должен быть отфильтрован.

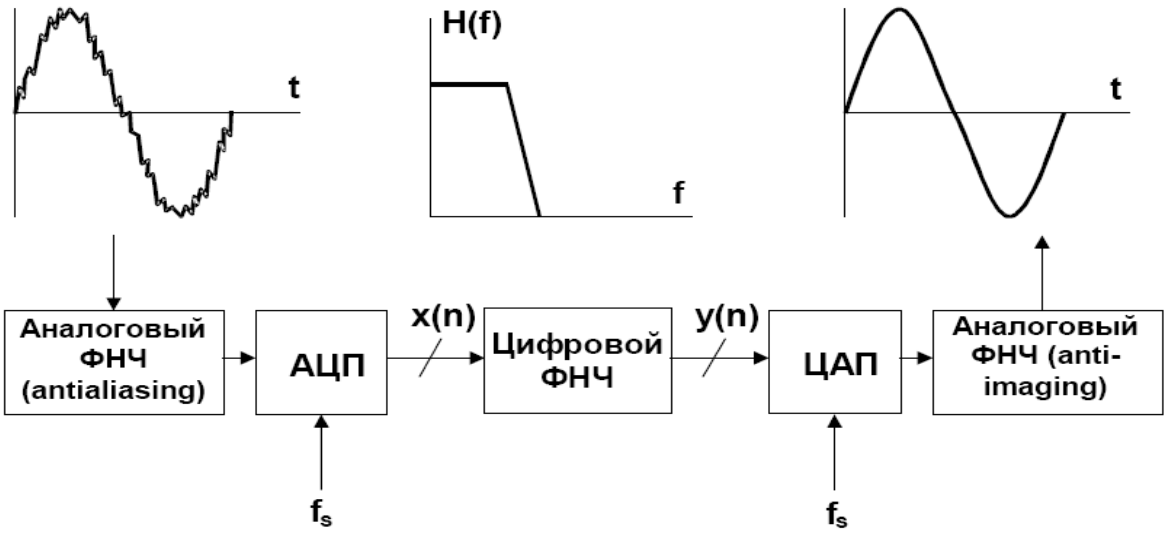


Рисунок 4.8 — Цифровая фильтрация

Вначале сигнал должен быть оцифрован с помощью АЦП для получения выборки x(n). Далее эта выборка поступает на цифровой фильтр, который в данном случае является НЧ-фильтром. Отсчеты выходных данных y(n) используются для восстановления аналогового сигнала с использованием ЦАП с низким уровнем ложного сигнала

Тем не менее цифровые фильтры не могут являться решением всех возможных задач фильтрации, возникающих при обработке сигналов. Для работы в реальном масштабе времени DSP-процессор должен быть рассчитан на выполнение всех шагов в программе фильтрации в пределах промежутка времени, соответствующего одному такту дискретизации, то есть 1/fs. Высокопроизводительный универсальный DSP-процессор с фиксированной точкой типа ADSP-2189M, обладающий быстродействием 75MIPS, способен выполнить операцию умножения с накоплением при реализации одного каскада фильтра за 13,3 нс. DSP-процессор ADSP-2189M затрачивает N+5 инструкций при реализации фильтра с количеством каскадов N. Для 100-каскадного фильтра полное время вычисления составляет приблизительно 1,4 мкс. Это соответствует максимально возможной частоте дискретизации 714 кГц, ограничивая, таким образом, ширину полосы частот обрабатываемого сигнала несколькими сотнями килогерц.

Можно заменить универсальный DSP-процессор специализированным аппаратным цифровым фильтром, способным работать на частотах дискретизации, соответствующих видеосигналу. В других случаях ограничения по быстродействию могут быть преодолены сохранением выборки данных, поступающих с большой скоростью от АЦП, в буферной памяти. Затем буферная память читается со скоростью, совместимой с быстродействием цифрового фильтра, основанного на DSP. Используя данный метод, можно осуществлять обработку сигнала в псевдореальном масштабе времени в таких системах, как радар, где обычно обрабатываются пакеты данных, накапливаемые после каждого излучаемого импульса.

Существует два основных типа цифровых фильтров: фильтры с конечной импульсной характеристикой (КИХ) и фильтры с бесконечной импульсной характеристикой (БИХ). Как следует из терминологии, эта классификация относится к импульсным характеристикам фильтров. Изменяя веса коэффициентов и число звеньев КИХ-фильтра, можно реализовать практически любую частотную характеристику. КИХ-фильтры могут иметь такие свойства, которых невозможно достичь методами аналоговой фильтрации (в частности, совершенную линейную фазовую характеристику). Но высокоэффективные КИХ-фильтры строятся с большим числом операций умножения с накоплением и поэтому требуют использования быстрых и эффективных процессоров DSP. С другой стороны, БИХ-фильтры имеют тенденцию имитировать принцип действия традиционных аналоговых фильтров с обратной связью. Поэтому их импульсная характеристика имеет бесконечную длительность. Благодаря

использованию обратной связи БИХ-фильтры могут быть реализованы с меньшим количеством коэффициентов, чем КИХ-фильтры. Другим способом реализации КИХ- или БИХ-фильтрации являются решетчатые фильтры, которые часто используются в задачах обработки речи. Цифровые фильтры применяются в приложениях адаптивной фильтрации благодаря своему быстродействию и простоте изменения характеристик воздействием на его коэффициенты.

Типы цифровых фильтров:

- фильтр скользящего среднего;
- фильтр с конечной импульсной характеристикой (КИХ);
- фильтр с бесконечной импульсной характеристикой;
- решетчатые фильтры (могут быть КИХ или БИХ);
- адаптивные фильтры.

4.4 Сигнальные процессоры

Цифровой сигнальный процессор (англ. Digital signal processor, DSP) — специализированный микропроцессор, предназначенный для цифровой обработки сигналов.

Традиционные компьютеры особенно хороши для применения в двух областях деятельности: манипуляция данными, например подготовка текстов и управление базами данных; и математические вычисления, используемые в науке, технике и цифровой обработке сигналов. Однако большинство компьютеров не могут одинаково хорошо работать в обеих сферах. В компьютерных приложениях, таких, как, например, подготовка текстов, данные запоминаются, сортируются, сравниваются, перемещаются и т.д., и время на выполнение этих операций не имеет большого значения до тех пор, пока оно удовлетворяет конечного пользователя. В приложениях, работающих с базами данных, периодически возникает необходимость реализации математических операций, но скорость их выполнения не является главным фактором. В большинстве случаев при проектировании приложений общего назначения компании-производители не концентрируют внимания на создании более эффективных программ. Прикладные программы оказываются перегруженными различными дополнительными возможностями, для каждого обновления которых требуется все больше памяти и нужны все более быстрые процессоры.

С другой стороны, для цифровой обработки сигналов важно, чтобы математические операции выполнялись быстро, и время, требуемое на выполнение команд, должно быть известно точно и заранее. Для этого и программа, и аппаратура должны быть очень эффективными. Наиболее важной математической операцией и ядром всех алгоритмов цифровой обработки сигналов является умножение с последующим суммированием. Быстрое выполнение операции умножения с последующим

суммированием очень важно для реализации быстрого преобразования Фурье, цифровых фильтров реального времени, умножения матриц, манипуляции с графическими изображениями и т.д.

Проведенное предварительное обсуждение требований, предъявляемых к цифровым сигнальным процессорам, важно для понимания различий между микроконтроллерами, микропроцессорами и цифровыми сигнальными процессорами. Хотя микроконтроллеры при использовании в промышленных устройствах управления процессами могут выполнять такие функции, как умножение, сложение, деление, они лучше подходят для приложений, где возможности процессора по реализации ввода-вывода и управления важнее, чем скорость. Микроконтроллеры, например семейства 8051, обычно содержат ЦПУ, ПЗУ, ОЗУ, последовательный и параллельный интерфейсы, счетчики и схемы прерываний. Микроконвертеры MicroConverter™ компании Analog Devices содержат не только ядро, построенное по архитектуре 8051, но также высококачественные ЦАП, АЦП и блок энергонезависимой памяти, реализованной по технологии FLASH.

Требования, предъявляемые к цифровым процессорам обработки сигналов

Наиболее важная операция в цифровой обработке сигналов — это суммирование результатов:

$$y(n) = h(0) \cdot x(n) + h(1) \cdot x(n - 1) + \ldots + h(N - 1) \cdot x(n - N).$$

Пример: цифровая фильтрация.

1. Многократное умножение значений входных отсчетов на коэффициенты фильтра (или на поворотные множители при БПФ).
2. Накопление результатов умножения в регистре-аккумуляторе.
3. Повторение этих действий N раз.

Требования, предъявляемые к DSP:

- быстрое выполнения умножения с накоплением;
- высокая точность представления результата (в аккумуляторе);
- одновременная выборка двух операндов;
- наличие циклических буферов;
- реализация циклов с автоматической проверкой условий.

При использовании ядра ADSP-21xx за один цикл возможно осуществить:

- выборку значения отсчета из памяти данных;
- выборку значения коэффициента из памяти программ;
- умножение с накоплением.

Данная операция одинаково важна для цифровых фильтров, БПФ и для множества других алгоритмов цифровой обработки сигналов. Цифровой сигнальный процессор (DSP) оптимизирован для осуществления повторяющихся математических операций, таких, как умножение с

накоплением. Пять основных требований предъявляется к DSP, чтобы оптимизировать производительность процессора: быстрое выполнение арифметических операций, повышенная точность представления операндов, возможность одновременной выборки двух операндов, поддержка циклических буферов, организация циклов с автоматической проверкой условия завершения цикла.

Быстрое выполнение арифметических действий

Быстрое выполнение арифметических действий — наиболее простое для понимания требование. Так как возможность реализации цифровой обработки сигналов в реальном масштабе времени зависит от производительности процессора, быстрота выполнения операций умножения с накоплением является главным требованием; большая скорость выполнения данной операции означает возможность обработки большей полосы частот. Но необходимо помнить, что эффективность DSP определяется не только временем выполнения операции умножения с накоплением. Этот часто забываемый факт приводит к неадекватному подходу в оценке производительности процессора, когда скорость работы процессора оценивается количеством операций, выполняемых процессором за единицу времени (в MIPS — миллионах операций в секунду). Так как большинство DSP и других процессоров, имеющих сходную архитектуру, могут выполнять за один машинный цикл команду MAC (умножение с накоплением), для большинства процессоров при оценке производительности в MIPS подразумевается производительность процессора при выполнении команды умножения с накоплением (MAC). Эта величина не учитывает другие свойства процессора, которые на практике могут повлиять на его общую производительность. Если остальные четыре критерия производительности окажутся неудовлетворительными, то высокая производительность процессора при выполнении MAC мало что даст.

В дополнение к требованиям по быстрому выполнению арифметических действий, DSP должен эффективно выполнять другие математические функции общего назначения и должен иметь соответствующее арифметико-логическое устройство (АЛУ) и возможность программировать операции сдвига для манипуляции с битами.

Повышенная точность

Кроме очевидной необходимости быстрого выполнения операции умножения со сложением (MAC), от DSP требуется высокая точность представления результата в регистре-аккуму-ляторе. Например, когда перемножаются два 16-битных слова, результат представляется 32-битным словом. Ядро процессоров компании Analog Devices семейства ADSP-21xx с фиксированной точкой имеет встроенный 40-битный аккумулятор, который обеспечивает большой запас суммирования без переполнения. Хотя использование DSP с плавающей точкой автоматически устраняет

большинство проблем, связанных с точностью и переполнением, процессоры с фиксированной точкой остаются очень популярными для многих приложений, и поэтому при их использовании нужно обращать достаточное внимание на возможное переполнение, потерю результатов (выход результата операции за пределы разрядной сетки) и масштабирование операндов.

Одновременная выборка двух операндов

Независимо от типа используемого микропроцессора ограничения в его работе в основном связаны с пропускной способностью шины. В случае микропроцессоров общего назначения или микроконтроллеров программа в основном состоит из команд, подразумевающих однократное обращение к памяти, обычно адресуемых при помощи сдвига относительно базового адреса. Это заставляет разработчиков микропроцессоров так проектировать систему команд, чтобы фиксированные данные встраивались в код, поскольку такой тип получения операндов является быстрым и эффективным с точки зрения использования памяти. С другой стороны, в DSP преобладают команды, требующие двух независимых обращений к памяти. Данное требование вытекает из самой сути операции свертки (перемножение с суммированием) $\Sigma h(i) \cdot x(i)$. Целью быстрой одновременной выборки двух операндов является необходимость непрерывной загрузки накапливающего умножителя (MAC). При описании MAC мы видели, что быстродействие DSP в основном определяется скоростью MAC. Если мы считаем, что MAC выполняется за приемлемое время, то очевидно, что для каждой операции требуется с той же скоростью подавать на MAC два операнда. Увеличение времени выборки операндов из памяти соответствующим образом отразится на скорости работы MAC. В идеале обращение происходит одновременно с выполнением операции в MAC в одном и том же машинном цикле.

Одновременная выборка двух операндов в DSP осуществляется по двум независимым шинам: шине данных памяти программ и шине данных памяти данных. Кроме того, имеются отдельные шина адреса памяти программ и шина адреса памяти данных. Таким образом, MAC может получать входные данные по каждой шине данных одновременно. Такая архитектура обычно называется гарвардской.

Циклические буферы

Если мы более внимательно исследуем самую распространенную при цифровой обработке сигналов операцию, то преимущества использования циклических буферов в DSP станут очевидными. Возьмем для примера фильтр с конечной импульсной характеристикой (КИХ). Во-первых, набор коэффициентов КИХ-фильтра по своей природе имеет периодический характер. Во-вторых, при каждом вычислении значения отсчета выходного

сигнала КИХ-фильтр использует новый отсчет входного сигнала и отбрасывает самый старый отсчет.

При последовательных вычислениях произведений коэффициентов КИХ-фильтра на отсчеты сигнала доступ к N коэффициентам фильтра осуществляется последовательно от h(0) до h(N –1). Набор отсчетов входного сигнала циркулирует в памяти следующим образом: новый отсчет входного сигнала сохраняется в памяти вместо старого отсчета всякий раз, когда вычисляется выходное значение фильтра. Для такого циркулирующего буфера может использоваться фиксированная область в ОЗУ. Самое раннее значение в памяти заменяется новым после каждого вычисления операции свертки. При этом информация об N последних отсчетах сохраняется в ОЗУ.

В виде буфера в ОЗУ DSP-процессора может быть реализована задержка, если новые значения записываются в память на место старых. Для упрощения адресации памяти старые значения считываются из памяти, начиная со значения, расположенного сразу после того, которое было только что записано. Например, в КИХ-фильтре с четырьмя коэффициентами новый отсчет x(4) записывается в ячейку памяти с адресом 0. Далее чтение данных осуществляется из ячеек с адресами 1, 2, 3 и 0 в указанном порядке. Этот способ применяется при любом числе звеньев фильтра. При такой адресации ячеек памяти генератор адреса должен выдавать лишь последовательные значения адресов, вне зависимости от того, какая операция с памятью — чтение или запись — осуществляется в настоящий момент. Буфер такого типа называется циклическим, потому что когда при записи достигается последняя ячейка, указатель памяти устанавливается на начало буфера.

Выборка коэффициентов из памяти осуществляется одновременно с выборкой данных. При рассмотренной схеме адресации самые старые отсчеты извлекаются из памяти первыми. Поэтому последний из коэффициентов должен выбираться из памяти первым. Коэффициенты могут заноситься в памяти в обратном порядке: h(N –1) — в первую ячейку, а h(0) — в последнюю, и генератор адреса в этом случае должен генерировать последовательно возрастающие адреса. И наоборот, коэффициенты могут быть записаны в памяти в нормальном порядке, но доступ к ним при этом должен осуществляться начиная с конца буфера, а генератор адреса должен генерировать последовательно убывающие адреса.

Описанные выше механизмы позволяют реализовать задержку, требуемую при реализации КИХ-фильтра, без каких-либо дополнительных затрат процессорного времени. Использование циклических буферов является специфическим для цифровой обработки сигналов, и для достижения максимальной эффективности циклические буферы должны поддерживаться аппаратно. Аппаратная реализация циклических буферов позволяет установить параметры буфера (такие, как адрес начала буфера, длина и т.д.) в программе вне тела цикла, непосредственно вычисляющего

алгоритм. Это позволяет избежать включения дополнительных команд в тело цикла. Отсутствие аппаратной реализации циклических буферов может существенным образом ухудшить возможности DSP-процессора по реализации алгоритмов цифровой обработки сигналов.

Организация циклов с автоматической проверкой условий

Необходимость поддержки циклов с автоматической проверкой условий завершения вызвана циклическим характером алгоритмов цифровой обработки сигналов (ЦОС). Функция умножения с накоплением и выборка данных повторяются N раз при каждом вычислении типового алгоритма. В традиционных микропроцессорах организация цикла предполагает наличие в заголовке цикла команд для проверки условия окончания цикла. Архитектура DSP-процессоров компании Analog Devices обеспечивает аппаратную поддержку программных циклов без необходимости программной проверки условия продолжения или завершения в теле цикла. Для типичной DSP-архитектуры различие в производительности при аппаратной поддержке цикла с автоматической проверкой условия завершения и при программной проверке условия завершения цикла может превышать 20 % времени выполнения цикла.

Подведя итог, можно сказать, что любой процессор может выполнить любой алгоритм при наличии достаточного времени. Однако DSP-процессоры оптимизированы под конкретные вычисления, связанные с обработкой реальных сигналов в реальном масштабе времени. Традиционные компьютеры больше подходят для вычислительных задач, не связанных с реальным временем.

5 ТЕХНОЛОГИЯ ПЕРЕДАЧИ ИНФОРМАЦИИ

5.1 Интерфейс «общая шина»

Интерфейс — это аппаратное и программное обеспечение (элементы соединения и вспомогательные схемы управления, их физические, электрические и логические параметры), предназначенное для сопряжения систем или частей системы (программ или устройств). Под сопряжением подразумеваются следующие функции:

- выдача и прием информации;
- управление передачей данных;
- согласование источника и приемника информации.

В связи с понятием интерфейса рассматривают также понятие «шина» (магистраль) — это среда передачи сигналов, к которой может параллельно подключаться несколько компонентов вычислительной системы и через которую осуществляется обмен данными. Очевидно, для аппаратных составляющих большинства интерфейсов применим термин «шина», поэтому зачастую эти два обозначения выступают как синонимы, хотя интерфейс — понятие более широкое.

Для интерфейсов, обеспечивающих соединение «точка-точка» (в отличие от шинных интерфейсов), возможны следующие реализации режимов обмена: дуплексный, полудуплексный и симплексный. К дуплексным относят интерфейсы, обеспечивающие возможность одновременной передачи данных между двумя устройствами в обоих направлениях. В случае когда канал связи между устройствами поддерживает двунаправленный обмен, но в каждый момент времени передача информации может производиться только в одном направлении, режим обмена называется полудуплексным. Важной характеристикой полудуплексного соединения является время реверсирования режима — то время, за которое производится переход от передачи сообщения к приему и наоборот. Если же интерфейс реализует передачу данных только в одном направлении и движение потока данных в противоположном направлении невозможно, такой интерфейс называют симплексным.

Важное значение имеют также следующие технические характеристики интерфейсов:

– вместимость (максимально возможное количество абонентов, одновременно подключаемых к контроллеру интерфейса без расширителей);

– пропускная способность или скорость передачи (длительность выполнения операций установления и разъединения связи и степень совмещения процессов передачи данных);

– максимальная длина линии связи;

– разрядность;

– топология соединения.

Архитектура системных интерфейсов

По функциональному назначению можно выделить системные интерфейсы (интерфейсы, связывающие отдельные части компьютера как микропроцессорной системы) и интерфейсы периферийных устройств.

Микро-ЭВМ с точки зрения архитектуры можно разделить на 2 основных класса:

– использующие внутренний интерфейс МП (унифицированный канал);

– использующие внешний по отношению к МП системный интерфейс.

Системный интерфейс выполняется обычно в виде стандартизированных системных шин. Однако в последнее время наметились тенденции внедрения концепций сетевого взаимодействия в архитектуру системных интерфейсов.

Различают два класса системных интерфейсов: с общей шиной (сигналы адреса и данных мультиплексируются) и с изолированной шиной (раздельные сигналы данных и адреса). Прародителями современных системных шин являются:

– Unibus фирмы DEC (интерфейс с общей шиной);
– Multibus фирмы Intel (интерфейс с изолированной шиной).

Шинная архитектура Unibus была разработана фирмой DEC для мини-ЭВМ серии PDP-11. Общая шина для периферийных устройств, памяти и процессора состоит из 56 двунаправленных линий. Unibus поддерживает пересылку одного 16-разрядного слова за 750 нс. Все пересылки инициируются ведущим устройством и подтверждаются принимающим (запоминающим) устройством, что позволяет работать с модулями различного быстродействия. Выбор устройства на роль ведущего является динамической процедурой, поэтому в ответ на запрос периферийного устройства процессор может передать ему управление шиной. Благодаря этой особенности на основе Unibus возможна разработка мультипроцессорных систем. Unibus позволяет подключать к магистрали большое число устройств, хотя необходимо учитывать снижение надежности по мере увеличения длины магистрали. Данные регистров внешних устройств могут обрабатываться теми же командами, что и данные в памяти. Следует, однако, отметить сложность технической реализации интерфейсных модулей, связанных с пересылкой адресов и данных по одним и тем же линиям.

Свое развитие архитектура Unibus получила в системном интерфейсе NuBus. Интерфейс NuBus (табл. 5.1) был разработан совместно с Western Digital в 1979 г. Затем, при участии Texas Instruments, архитектура NuBus была стандартизована (стандарт IEEE 1196-1987) и применялась фирмой Apple в компьютерах Macintosh. В NuBus также используется мультиплексирование адреса и данных. Предусмотрена автоматическая конфигурация. Возможно использование нескольких задатчиков магистрали с децентрализованным арбитражем. Имеется режим блочной передачи данных. К недостаткам NuBus можно отнести слабые возможности режима ПДП, сложный метод обработки прерываний (предусмотрен всего один сигнал запроса прерывания и программный опрос потенциальных источников прерываний).

Альтернативная шинная архитектура Multibus была разработана фирмой Intel. Шина также обеспечивает системную архитектуру с одним или несколькими ведущими узлами и с квитированием установления связи между устройствами, работающими с разной скоростью. Благодаря разделению шины адреса и шины данных возможны реализации этой архитектуры для процессоров разной разрядности. Существовали 8-разрядный и 16-раз-рядный варианты архитектуры Multibus для IBM PC. Шина адреса — 20 бит. Multibus подразумевает достаточно простую аппаратную реализацию, однако число устройств, одновременно использующих ресурсы шины, ограничено 16 абонентами. Следует отметить, что скорость обмена на шине Multibus была ниже, чем на шине Unibus.

Таблица 5.1 — Системные интерфейсы

Шина	NuBus	ISA	EISA	MCA	VLB	PCI
Год выпуска	1979	1984	1989	1987	1987	1992
Разрядность данных	32	8/16	32	32/64	32	32/64
Разрядность адреса	32	20/24	32	32	32	32
Тактовая частота, МГц	10	4/8	8	10	<33 (Fцп)	33, 66
Макс. скорость, Мбайт/с	37	8—16	33	20/40	130	132/264, 520
Макс. кол-во устройств		6	15	16	2—3	10
Кол-во сигналов	96	62/98	188	178	112	124/188

Упрощенная структура интерфейса «Общая шина» представлена на рис. 5.1, а основные сигналы приведены в табл. 5.2.

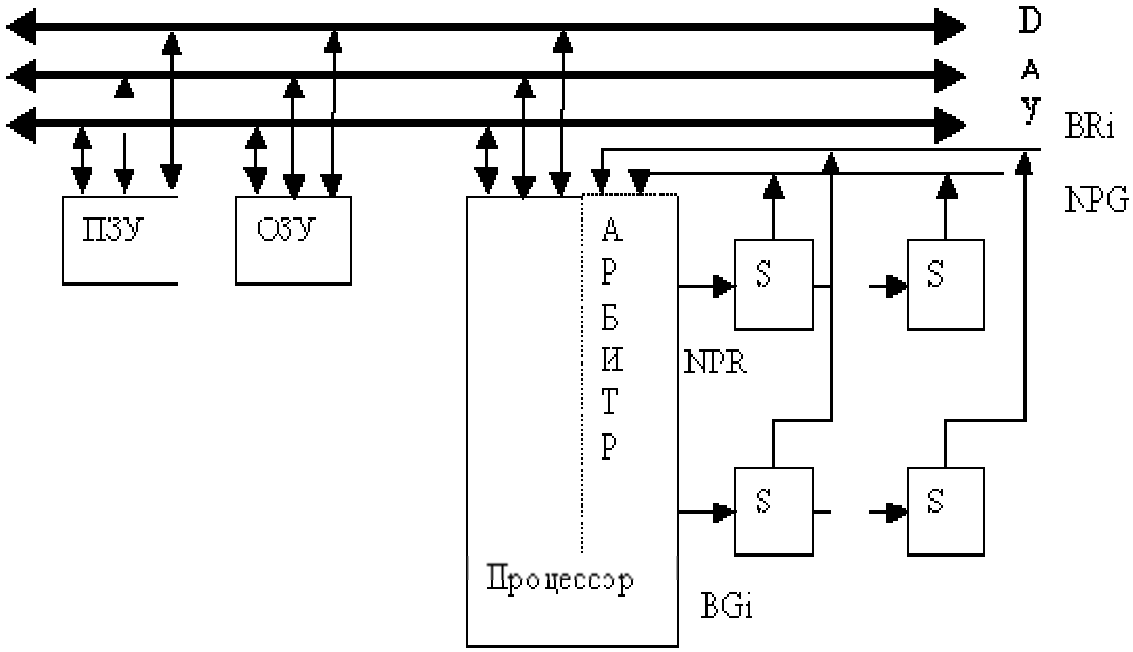


Рисунок 5.1 — Упрощенная структура интерфейса «Общая шина»

Таблица 5.2 — Назначение сигналов интерфейса «Общая шина»

Обозначение сигналов	Назначение выводов
Шина данных	
Д15—ДОО РА, РВ	Параллельная 16-разрядная шина данных Сигналы контроля передачи информации
Шина адреса	
А17—АОО	Параллельная 18-разрядная шина адреса
Шина управления	
CI,CO MSYN SSYN INIT ACLO DCLO	Шина управления передачей информации Сигнал синхронизации активного устройства Сигнал синхронизации пассивного устройства Сигнал сброса системы в начальное состояние Авария сети питания переменного напряжения Авария источника питания постоянного напряжения
Шина прерываний	
BBSY INTR BR4—BR7 (Bus request) BG4—BG7 (Bus grant) NPR (non-processor request) NRG (non-processor grant) SACK	Сигнал занятости общей шины Строб передачи вектора прерывания от внешнего устройства к процессору Сигналы запросов прерывания от 4-х групп внешних устройств Сигналы предоставления прерывания 4-м группам внешних устройств Сигнал запроса прямого доступа к памяти Сигнал предоставления прямого доступа к памяти Сигнал подтверждения выборки устройства ПДП

Интерфейс «Общая шина» имеет классическую архитектуру с тремя основными шинами — данных (D), адреса (А) и управления (У), к которым подключаются модули ОЗУ, ПЗУ, процессора и других исполнителей (рис. 5.1). В модуле процессора существует специальный узел — арбитр, который производит обработку сигналов прямого доступа к памяти (ПДП) (NPR) и прерываний (BR). Если соответствующая процедура разрешена, арбитр устанавливает сигналы разрешения ПДП (NPG) или подтверждения прерывания (BG).

Данные в процессе обмена передаются по 16-разрядной двунаправленной шине данных. При этом к данным могут быть добавлены сигналы РА, РВ для проведения в процессе работы контроля/проверки по четности. Кодировка этих сигналов выглядит следующим образом:

00	нет ошибки
01	ошибка при операции чтения
10, 11	зарезервировано

В процессе обмена данными участвуют только два устройства «задатчик» и «исполнитель». Последний определяется в цикле обмена информацией с помощью сигналов, устанавливаемых задатчиком на адресную шину. Направление обмена (циклы записи или чтения) задается с помощью сигналов С1 и С2, которые определяют приказ интерфейса (табл. 5.3).

Таблица 5.3 — Кодировка приказов в интерфейсе «Общая шина»

С1	С2	Тип приказа
1	1	Запись байта
0	1	Запись слова
0	0	Чтение слова
1	0	Чтение слова с паузой

Циклы обмена в интерфейсе осуществляются с помощью классических операций «чтение и запись». При этом в циклах чтения происходит передача всего слова, тогда как в циклах записи возможны операции со словами и байтами. В интерфейсе предусмотрен элементарный цикл обмена — чтение слова с паузой. Данная операция подразумевает чтение слова из ячейки памяти с последующей записью в ту же ячейку памяти информации после обработки в процессоре. Так как данная операция считается элементарной, то между циклами записи и чтения исключена процедура арбитража, что несколько повышает производительность МПС. На рис. 5.2 представлен протокол обмена информацией для цикла чтения в интерфейсе «Общая шина».

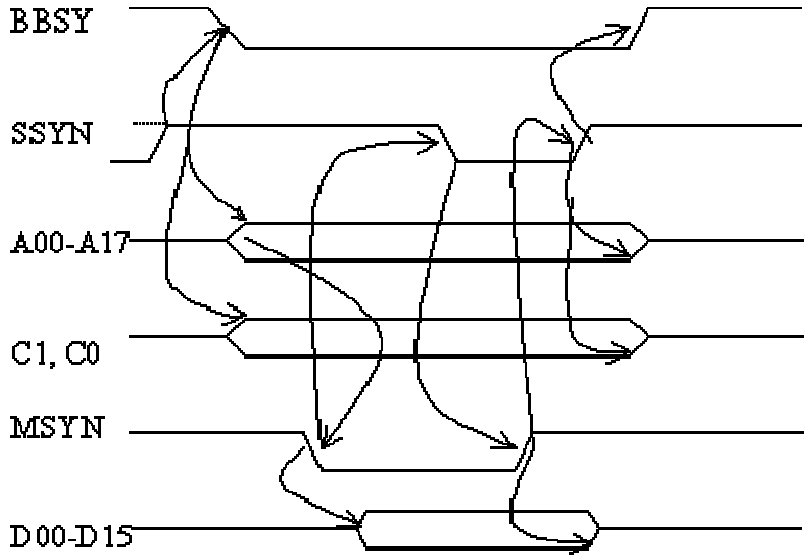


Рисунок 5.2 — Цикл чтения в интерфейсе «Общая шина»

Процедура обмена в цикле чтения описывается следующим образом.

1. Задатчик выставляет сигнал занятости магистрали BBSY.

2. На шине адреса задатчик выставляет адрес внешнего устройства A и код приказа C0, C1.

3. С задержкой 150 нс и при пассивном сигнале квитирования SSYN задатчик устанавливает строб данных MSYN.

4. Определенный по адресной информации исполнитель по приходу сигнала MSYN устанавливает информацию на шину данных.

5. Исполнитель выставляет сигнал квитирования SSYN.

6. Задатчик с задержкой не менее 75 нс принимает данные от исполнителя.

7. Задатчик снимает строб данных MSYN.

8. Исполнитель освобождает магистраль, снимая сигналы D и SSYN.

9. Задатчик с задержкой не менее 75 нс после снятия MSYN освобождает магистраль (A, C, BBSY).

Аналогично реализуются и другие элементарные циклы обмена (приказы).

Процедуры смены задатчика и обработка прерываний осуществляются с помощью арбитра. В данной системе реализован радиальный арбитраж с возможностью подключения еще нескольких устройств по каждой линии запроса, для которых реализован последовательный арбитраж в неоднородной структуре.

5.2 Контроллеры

Активные (ПДП)

Обычно в микропроцессорной системе именно процессор управляет всей ее работой. Он выдает на шину адреса рабочий адрес, он выдает на шину управления сигналы RD, WR, MREQ и IORQ. Режим прямого доступа к памяти — это тот случай, когда это не так. Он применяется в тех случаях, когда нужно перенести большой блок информации из одной области памяти в другую. Либо байт за байтом выдать блок информации на один из портов ввода/вывода. Либо наоборот, получить блок информации байт за байтом из порта ввода/вывода и записать его в какую-нибудь область памяти. Под блоком информации естественно подразумевают некую последовательность байтов, хранящуюся в памяти в смежных ячейках (например, с адреса ADDR1 по адрес ADDR2).

Хороший пример такой задачи — вывод текста на печать. Текстовая информация хранится в памяти компьютера следующим образом: каждая буква, знак препинания, арифметические значки и другие символы закодированы некими числами. Существует множество таблиц кодировки. В настоящее время наиболее распространен американский стандарт ASCII. В памяти компьютера располагаются последовательно коды букв текста, как они пишутся на экране. Пробел между словами и предложениями — это тоже символ. Он также имеет свой код. Свои коды имеют и специальные символы. Например, символ перевода строки и символ перехода на новую

страницу. Правда, такой способ представления текста используется только в простейших случаях. Например, в формате TXT (файлы программы «Блокнот»). Современные текстовые процессоры, такие, как «Microsoft Word», кроме кодов символов, хранят огромное количество служебной информации. Это информация о размере и типе шрифта, таблицы, форматирование и еще множество свойств, которые введены для получения высококачественного текста. Мы не будем рассматривать этот случай. Считайте, что это простейший текстовый редактор. А принтер — простейший матричный. Матричные принтеры исторически устроены так, что имеют в своей памяти фиксированный набор шрифтов. Для того чтобы такой принтер напечатал текст, нужно просто последовательно выдавать на него коды символов с компьютера. Получив код очередного символа, процессор, встроенный в принтер, самостоятельно печатает этот символ на бумаге. Получив код перевода строки, принтер прокрутит бумагу на одну строку. Получив код перевода страницы, принтер прокрутит бумагу до конца и произведет загрузку нового листа. Естественно, с такой задачей прекрасно может справиться и главный процессор. Нужно только написать соответствующую программу. Так раньше и делали. Но с появлением скоростных принтеров такой способ перестал удовлетворять. Если блок информации передается программным путем, нельзя достигнуть максимальной производительности, так как для вывода одного байта процессор должен выполнить 5—6 команд программы. Нужно считать из памяти очередной код, выдать его в порт, увеличить счетчик адреса на единицу, проверить, не достигнут ли конец блока. Затем повторить все для всех остальных байтов, составляющих этот блок информации.

Для ускорения таких операций был придуман режим прямого доступа к памяти. Для осуществления этого режима были разработаны специальные микросхемы — контроллеры прямого доступа к памяти. На рисунке приведена схема подключения контроллера прямого доступа к памяти (ПДП) к микропроцессору и микропроцессорной системе.

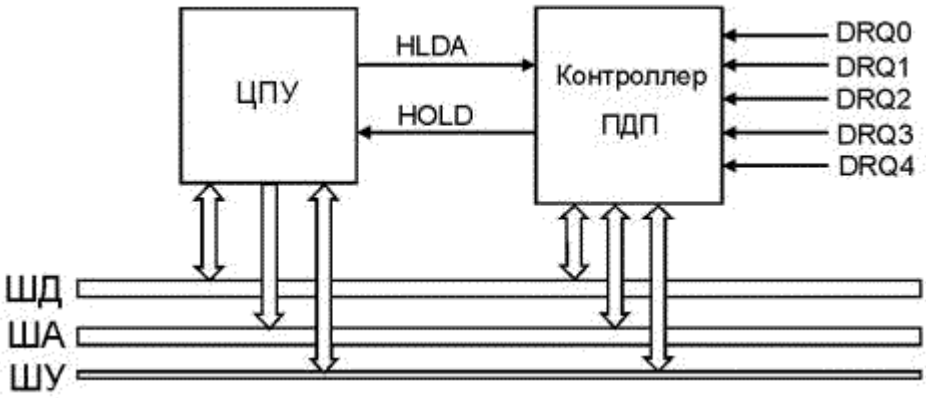


Рисунок 5.3 — Схема подключения контроллера прямого доступа к памяти

Микросхема контроллера прямого доступа к памяти выполнена по технологии программируемых микросхем. В данном случае понятие «программируемая» означает то, что микросхема имеет внутри специальные регистры, которые подключаются к системной шине как порты ввода/вывода. Процессор может записывать в эти порты различные числа и тем самым менять режимы работы микросхемы. Так, процессор записывает в контроллер прямого доступа к памяти адреса начала и конца блока памяти, подлежащего передаче. Кроме того, таким же образом в микросхему ПДП записывается режим работы (выдать блок в порт / считать блок из порта / переместить блок в памяти). Обычно команды программирования микросхем помещают в начале управляющей программы микропроцессора. Выполнив эти команды, процессор переходит к выполнению своей основной программы. А контроллер прямого доступа в это время находится в режиме ожидания. Он ждет специального сигнала — запроса на ПДП от одного из внешних устройств. Такие запросы должны поступать на один из входов DRQ0…DRQ4 контроллера ПДП. В нашем случае сигнал запроса ПДП поступает от схемы параллельного интерфейса, который вырабатывает его при получении сигнала готовности от принтера. При получении сигнала запроса на ПДП микросхема контроллера формирует сигнал запроса на захват (HOLD), который поступает на центральный процессор. Получив этот сигнал, процессор сначала заканчивает текущую операцию, затем переходит в специальный режим прямого доступа к памяти и сообщает об этом контроллеру ПДП при помощи сигнала подтверждения захвата (HLDA). В режиме прямого доступа процессор отключается от системной шины, и управление полностью берет на себя контроллер ПДП. Он сам вырабатывает сигналы адреса, сигналы управления (RD, WR, MREQ, IORQ). Таким образом, данные с максимальной скоростью передаются на принтер. По окончании процесса контроллер ПДП отключается от системной шины и снимает с процессора сигнал HOLD. Процессор возобновляет свою работу.

В любом персональном компьютере имеется несколько устройств, которые пользуются системой прямого доступа к памяти. В IBM-совместимых компьютерах, к которым относится славное семейство «пентиумов», существует четыре канала прямого доступа к памяти. Такой механизм работы с памятью, кроме принтеров, использует сетевые и звуковые карты, а так же накопители на гибких и жестких дисках.

Пассивные

Использование последовательных линий связи для обмена данными с внешними устройствами возлагает на контроллеры ВУ дополнительные, по сравнению с контроллерами для параллельного обмена, функции. Во-первых, возникает необходимость преобразования формата данных: из параллельного формата, в котором они поступают в контроллер ВУ из системного интерфейса микроЭВМ, в последовательный при передаче в

ВУ и из последовательного в параллельный при приеме данных из ВУ. Во-вторых, требуется реализовать соответствующий режиму работы внешнего устройства способ обмена данными: синхронный или асинхронный.

Простой контроллер для синхронной передачи данных в ВУ по последовательной линии связи (последовательный интерфейс) представлен на рис. 5.4.

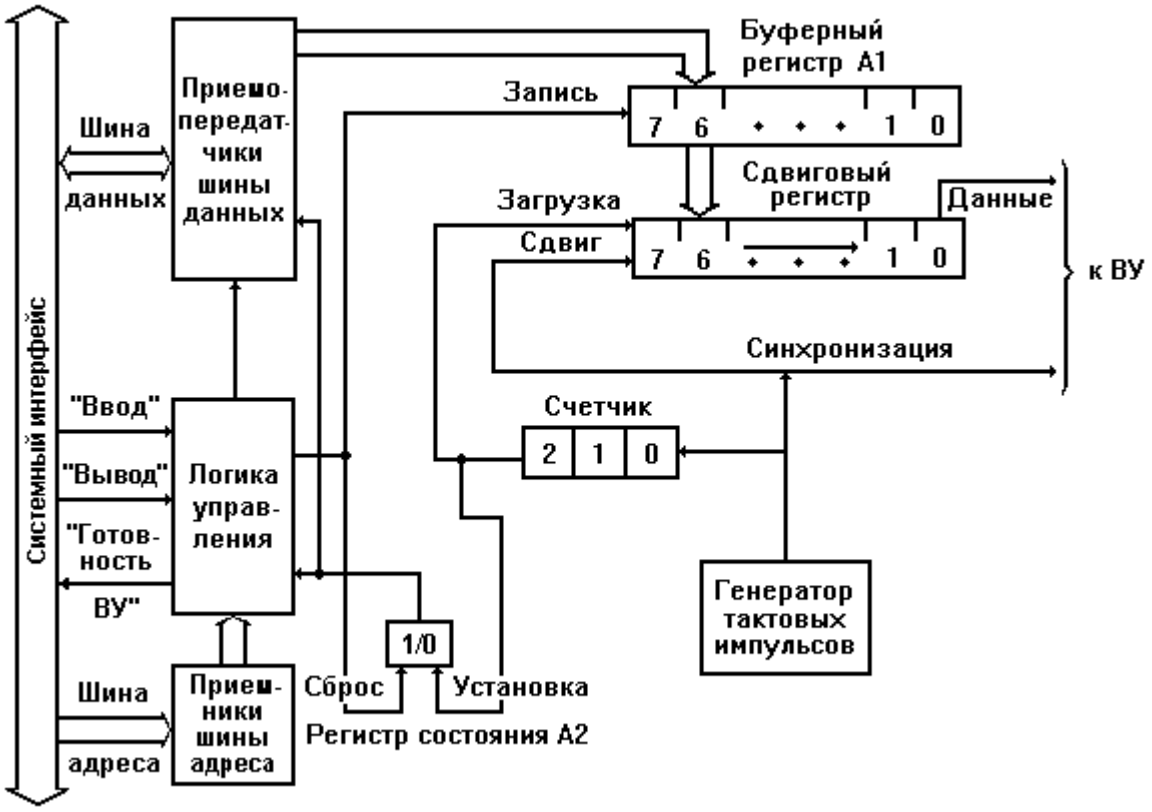


Рисунок 5.4 — Контроллер последовательной синхронной передачи

Восьмиразрядный адресуемый буферный регистр контроллера А1 служит для временного хранения байта данных до его загрузки в сдвиговый регистр. Запись байта данных в буферный регистр с шины данных системного интерфейса производится при наличии единицы в одноразрядном адресуемом регистре состояния контроллера А2.

Единица в регистре состояния указывает на готовность контроллера принять очередной байт в буферный регистр. Содержимое регистра А2 передается в процессор по одной из линий шины данных системного интерфейса и используется для формирования управляющего сигнала системного интерфейса «Готовность ВУ». При записи очередного байта в буферный регистр А1 обнуляется регистр состояния А2.

Преобразование данных из параллельного формата, в котором они поступили в буферный регистр контроллера из системного интерфейса, в последовательный и передача их на линию связи производятся в сдвиговом регистре с помощью генератора тактовых импульсов и двоичного трехразрядного счетчика импульсов следующим образом.

Последовательная линия связи контроллера с ВУ подключается к выходу младшего разряда сдвигового регистра. По очередному тактовому импульсу содержимое сдвигового регистра сдвигается на один разряд вправо и в линию связи «Данные» выдается значение очередного разряда. Одновременно со сдвигом в ВУ передается по отдельной линии «Синхронизация» тактовый импульс. Таким образом, каждый передаваемый по линии «Данные» бит информации сопровождается синхронизирующим сигналом по линии «Синхронизация», что обеспечивает его однозначное восприятие на приемном конце последовательной линии связи.

Количество переданных в линию тактовых сигналов, а следовательно, и переданных бит информации подсчитывается счетчиком тактовых импульсов. Как только содержимое счетчика становится равным 7, т.е. в линию переданы 8 бит (1 байт) информации, формируется управляющий сигнал «Загрузка», обеспечивающий запись в сдвиговый регистр очередного байта из буферного регистра. Этим же управляющим сигналом устанавливается в «1» регистр состояния. Очередным тактовым импульсом счетчик будет сброшен в «0», и начнется очередной цикл выдачи восьми битов информации из сдвигового регистра в линию связи.

Синхронная последовательная передача отдельных битов данных на линию связи должна производиться без какого-либо перерыва, и следующий байт данных должен быть загружен в буферный регистр из системного интерфейса за время, не превышающее времени передачи восьми битов в последовательную линию связи.

При записи байта данных в буферный регистр обнуляется регистр состояния контроллера. Нуль в этом регистре указывает, что в линию связи передается байт данных из сдвигового регистра, а следующий передаваемый байт данных загружен в сдвиговый регистр.

Контроллер для последовательного синхронного приема данных из ВУ состоит из тех же компонентов, что и контроллер для синхронной последовательной передачи, за исключением генератора тактовых импульсов.

Организация асинхронного последовательного обмена данными с внешним устройством осложняется тем, что на передающей и приемной стороне последовательной линии связи используются настроенные на одну частоту, но физически разные генераторы тактовых импульсов и, следовательно, общая синхронизация отсутствует. Рассмотрим на примерах организацию контроллеров последовательных интерфейсов для последовательных асинхронных передачи и приема данных.

Простейший контроллер для асинхронной передачи данных в ВУ по последовательной линии связи представлен на рис. 5.5. Он предназначен для передачи данных в формате с двумя стоповыми битами.

После передачи очередного байта данных в регистр состояния А2 записывается 1. Единичный выходной сигнал регистра А2 информирует процессор о готовности контроллера к приему следующего байта данных и

передаче его по линии связи в ВУ. Этот же сигнал запрещает формирование импульсов со схемы выработки импульсов сдвига — делителя частоты сигналов тактового генератора на 16. Счетчик импульсов сдвига (счетчик по mod 10) находится в нулевом состоянии, и его единичный выходной сигнал поступает на вентиль И, подготавливая цепь выработки сигнала загрузки сдвигового регистра.

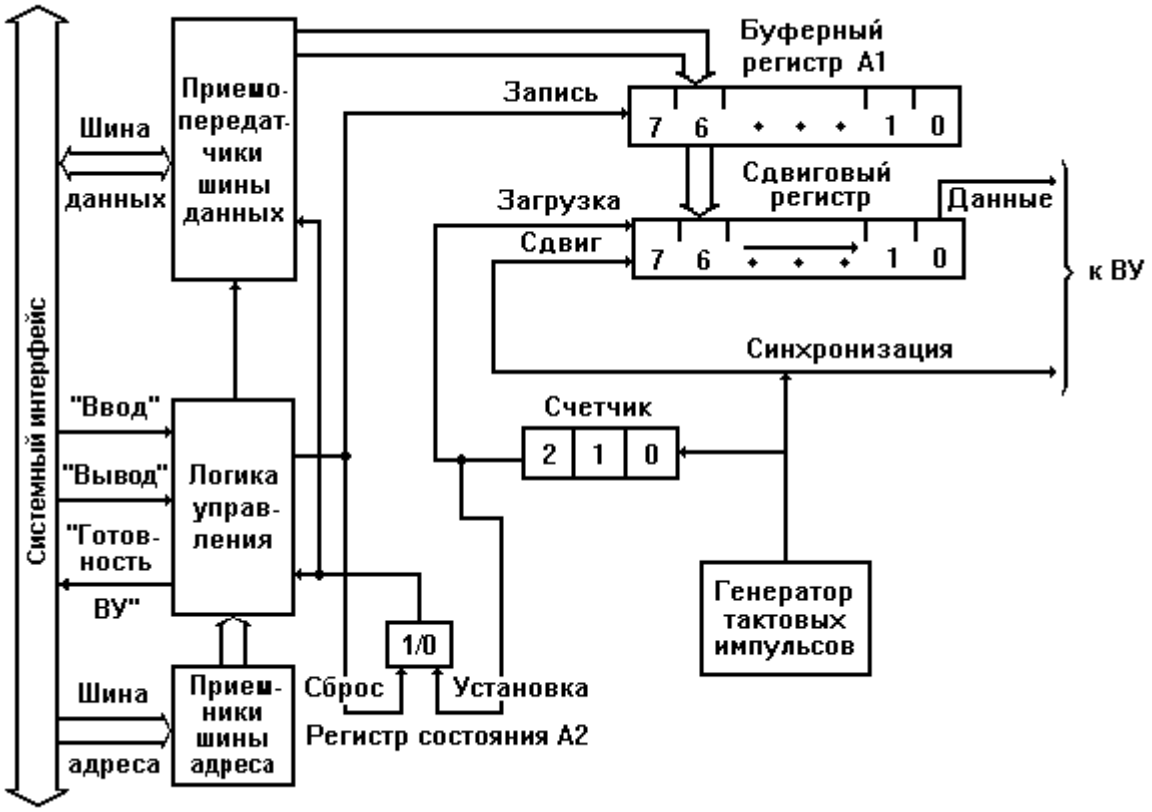


Рисунок 5.5 — Контроллер последовательной асинхронной передачи

Процесс передачи байта данных начинается с того, что процессор, выполняя команду «Вывод», выставляет этот байт на шине данных. Одновременно процессор формирует управляющий сигнал системного интерфейса «Вывод», по которому производятся запись передаваемого байта в буферный регистр А1, сброс регистра состояния А2 и формирование на вентиле И сигнала «Загрузка». Передаваемый байт переписывается в разряды 1,..., 8 сдвигового регистра, в нулевой разряд сдвигового регистра записывается 0 (стартовый бит), а в разряды 9 и 10—1 (стоповые биты). Кроме того, снимается сигнал «Сброс» с делителя частоты, он начинает накапливать импульсы генератора тактовой частоты и в момент приема шестнадцатого тактового импульса вырабатывает импульс сдвига.

На выходной линии контроллера «Данные» поддерживается состояние 0 (значение стартового бита) до тех пор, пока не будет выработан первый импульс сдвига. Импульс сдвига изменит состояние счетчика импульсов сдвига и перепишет в нулевой разряд сдвигового

регистра первый информационный бит передаваемого байта данных. Состояние, соответствующее значению этого бита, будет поддерживаться на линии «Данные» до следующего импульса сдвига.

Аналогично будут переданы остальные информационные биты, первый стоповый бит и, наконец, второй стоповый бит, при передаче которого счетчик импульсов сдвига снова установится в нулевое состояние. Это приведет к записи 1 в регистр состояния А2. Единичный сигнал с выхода регистра А2 запретит формирование импульсов сдвига, а также информирует процессор о готовности к приему нового байта данных. После завершения передачи очередного кадра (стартового бита, информационного байта и двух стоповых бит) контроллер поддерживает в линии связи уровень логической единицы (значение второго стопового бита).

Уровень логической единицы поступает по линии «Данные» в контроллер для асинхронного приема данных (рис. 5.6).

Этот уровень создает условия для выработки сигнала, запрещающего работу делителя частоты генератора тактовых импульсов. Действительно, после приема предыдущего байта данных счетчик импульсов сдвига (счетчик по mod 9) находится в нулевом состоянии и на вентиль И поступают два единичных сигнала: со счетчика сдвигов и из линии «Данные». На выходе вентиля И вырабатывается сигнал сброса делителя частоты сигналов тактового генератора, запрещающий формирование импульсов сдвига

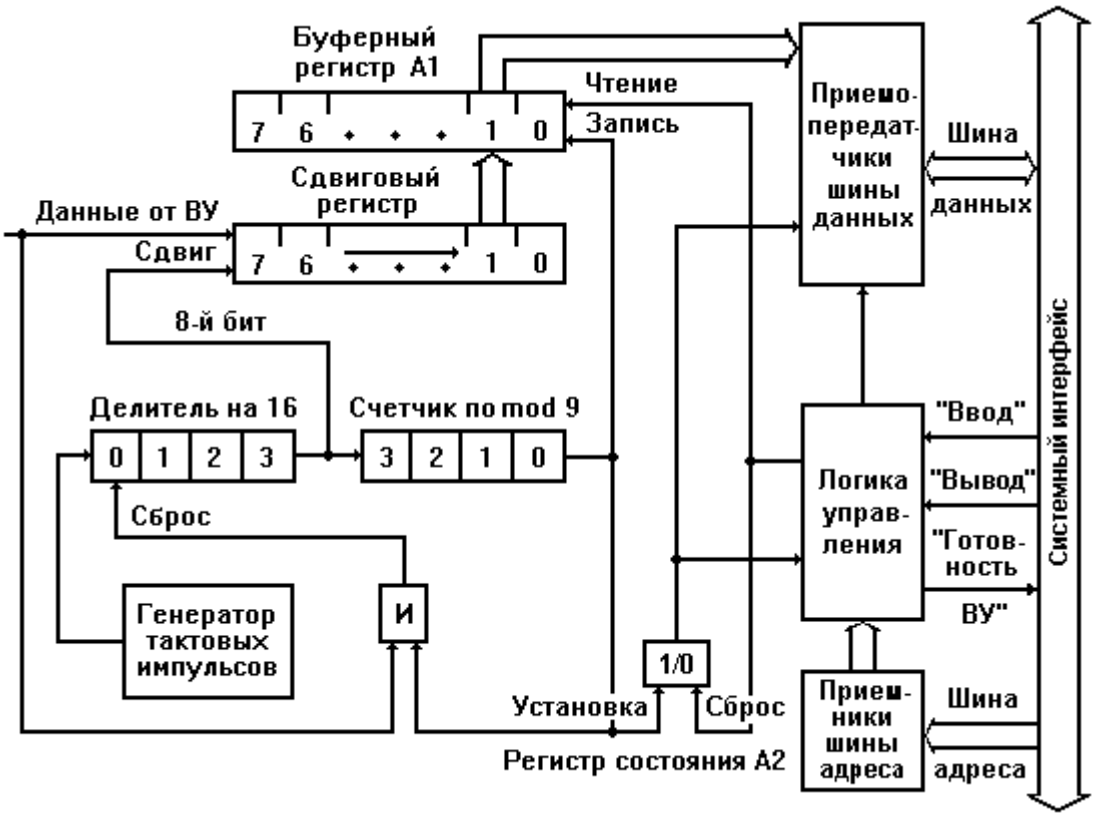


Рисунок 5.6 — Контроллер последовательного асинхронного приема

.

В момент смены стопового бита на стартовый бит (момент начала передачи нового кадра) на линии «Данные» появится уровень логического нуля, и тем самым будет снят сигнал сброса с делителя частоты. Состояние 4-разрядного двоичного счетчика (делителя частоты) начнет изменяться. Когда на счетчике накопится значение 8, он выдаст сигнал, поступающий на входы сдвигового регистра и счетчика импульсов сдвига. Так как частота сигналов генератора тактовых импульсов приемника должна совпадать с частотой генератора тактовых импульсов передатчика, то сдвиг (считывание) бита произойдет примерно на середине временного интервала, отведенного на передачу бита данных, т.е. времени, необходимого для выработки шестнадцати тактовых импульсов. Это делается для уменьшения вероятности ошибки из-за возможного различия частот генераторов передатчика и приемника, искажения формы передаваемых сигналов (переходные процессы) и т.п. Следующий сдвиг произойдет после прохождения шестнадцати тактовых импульсов, т.е. на середине временного интервала передачи первого информационного бита.

При приеме в сдвиговый регистр девятого бита кадра (восьмого информационного бита) из него «выдвинется» стартовый бит и, следовательно, в сдвиговом регистре будет размещен весь принятый байт информации. В этот момент счетчик импульсов сдвига придет в нулевое состояние и на его выходе будет выработан единичный сигнал, по которому содержимое сдвигового регистра перепишется в буферный регистр; в регистр состояния А2 запишется 1 и он будет информировать процессор об окончании приема очередного байта; вентиль И подготовится к выработке сигнала «Сброс» (этот сигнал сформируется после прихода первого стопового бита).

Получив сигнал готовности (1 в регистре А2), процессор выполнит команду «Ввод». При этом вырабатывается управляющий сигнал системного интерфейса «Ввод», по которому производятся пересылка принятого байта данных из буферного регистра в процессор (сигнал «Чтение») и сброс регистра состояния А2.

Отметим, что для простоты изложения в контроллере на рис. 5.6 не показаны схемы контроля стоповых бит принимаемого кадра. Не показаны также схемы контроля четности или нечетности (паритета) передаваемой информации (обычно в передаваемом байте восьмому биту придается значение 0 или 1, так чтобы в этом байте было четное количество единиц). В реальных контроллерах имеются такие схемы, и если контроллер не принимает из линии связи нужного количества стоповых бит или вырабатывается сигнал ошибки паритета в схеме контроля четности, то принятые в текущем кадре биты данных игнорируются и контроллер ожидает поступления нового стартового бита.

Обмен данными с ВУ по последовательным линиям связи широко используется в микроЭВМ, особенно в тех случаях, когда не требуется высокой скорости обмена. Вместе с тем применение в них

последовательных линий связи с ВУ обусловлено двумя важными причинами. Во-первых, последовательные линии связи просты по своей организации: два провода при симплексной и полудуплексной передаче и максимум четыре — при дуплексной. Во-вторых, в микроЭВМ используются внешние устройства, обмен с которыми необходимо вести в последовательном коде.

В современных микроЭВМ применяют, как правило, универсальные контроллеры для последовательного ВВ, обеспечивающие как синхронный, так и асинхронный режим обмена данными с ВУ.

Одной из разновидностей программно-управляемого обмена данными с ВУ в микроЭВМ является обмен с прерыванием программы, отличающийся от асинхронного программно-управ-ляемого обмена тем, что переход к выполнению команд, физически реализующих обмен данными, осуществляется с помощью специальных аппаратных средств. Команды обмена данными в этом случае выделяют в отдельный программный модуль — подпрограмму обработки прерывания. Задачей аппаратных средств обработки прерывания в процессоре микроЭВМ как раз и является приостановка выполнения одной программы (ее еще называют основной программой) и передача управления подпрограмме обработки прерывания. Действия, выполняемые при этом процессором, как правило, те же, что и при обращении к подпрограмме. Только при обращении к подпрограмме они инициируются командой, а при обработке прерывания — управляющим сигналом от ВУ, который называют «Запрос на прерывание» или «Требование прерывания».

Эта важная особенность обмена с прерыванием программы позволяет организовать обмен данными с ВУ в произвольные моменты времени, не зависящие от программы, выполняемой в микроЭВМ. Таким образом, появляется возможность обмена данными с ВУ в реальном масштабе времени, определяемом внешней по отношению к микроЭВМ средой. Обмен с прерыванием программы существенным образом экономит время процессора, затрачиваемое на обмен. Это происходит за счет того, что исчезает необходимость в организации программных циклов ожидания готовности ВУ, на выполнение которых тратится значительное время, особенно при обмене с медленными ВУ.

Прерывание программы по требованию ВУ не должно оказывать на прерванную программу никакого влияния, кроме увеличения времени ее выполнения за счет приостановки на время выполнения подпрограммы обработки прерывания. Поскольку для выполнения подпрограммы обработки прерывания используются различные регистры процессора (счетчик команд, регистр состояния и т.д.), то информацию, содержащуюся в них в момент прерывания, необходимо сохранить для последующего возврата в прерванную программу.

Обычно задача сохранения содержимого счетчика команд и регистра состояния процессора возлагается на аппаратные средства обработки прерывания. Сохранение содержимого других регистров процессора,

используемых в подпрограмме обработки прерывания, производится непосредственно в подпрограмме. Отсюда следует достаточно очевидный факт: чем больший объем информации о прерванной программе сохраняется программным путем, тем больше время реакции микроЭВМ на сигнал прерывания, и наоборот. Предпочтительными с точки зрения повышения производительности микроЭВМ (сокращения времени выполнения подпрограмм обработки, а следовательно, и основной программы) являются уменьшение числа команд, обеспечивающих сохранение информации о прерванной программе, и реализация этих функций аппаратными средствами.

Формирование сигналов прерываний — запросов ВУ на обслуживание происходит в контроллерах соответствующих ВУ. В простейших случаях в качестве сигнала прерывания может использоваться сигнал «Готовность ВУ», поступающий из контроллера ВУ в системный интерфейс микроЭВМ. Однако такое простое решение обладает существенным недостатком — процессор не имеет возможности управлять прерываниями, т.е. разрешать или запрещать их для отдельных ВУ. В результате организация обмена данными в режиме прерывания с несколькими ВУ существенно усложняется.

Для решения этой проблемы регистр состояния и управления контроллера ВУ (рис. 5.7) дополняют еще одним разрядом — «Разрешение прерывания».

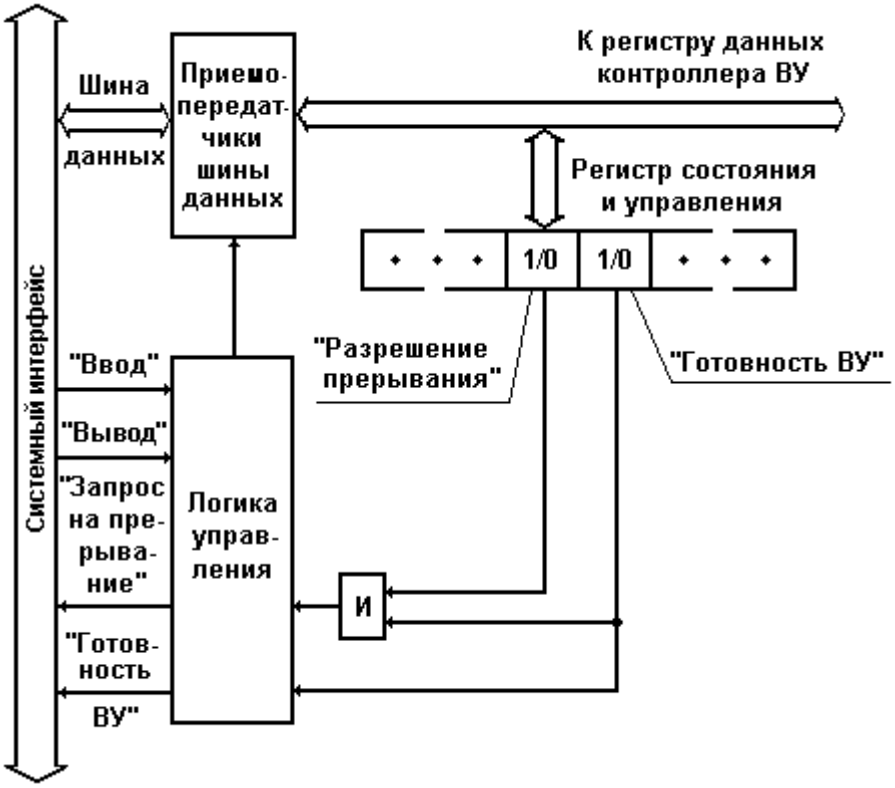


Рисунок 5.7 — Фрагмент блок-схемы контроллера ВУ с разрядом «Разрешение прерывания» в регистре состояния и управления

Запись 1 или 0 в разряд «Разрешение прерывания» производится программным путем по одной из линий шины данных системного интерфейса. Управляющий сигнал системного интерфейса «Запрос на прерывание» формируется с помощью схемы совпадения только при наличии единиц в разрядах «Готовность ВУ» и «Разрешение прерывания» регистра состояния и управления контроллера.

Аналогичным путем решается проблема управления прерываниями в микроЭВМ в целом. Для этого в регистре состояния процессора выделяется разряд, содержимое которого определяет, разрешены или запрещены прерывания от внешних устройств. Значение этого разряда может устанавливаться программным путем.

В микроЭВМ обычно используется одноуровневая система прерываний, т.е. сигналы «Запрос на прерывание» от всех ВУ поступают на один вход процессора. Поэтому возникает проблема идентификации ВУ, запросившего обслуживание, и реализации заданной очередности (приоритета) обслуживания ВУ при одновременном поступлении нескольких сигналов прерывания. Существуют два основных способа идентификации ВУ, запросивших обслуживания:

– программный опрос регистров состояния (разряд «Готовность ВУ») контроллеров всех ВУ;

– использование векторов прерывания.

Организация прерываний с программным опросом готовности предполагает наличие в памяти микроЭВМ единой подпрограммы обслуживания прерываний от всех внешних устройств. Структура такой подпрограммы приведена на рис. 5.8.

Обслуживание ВУ с помощью единой подпрограммы обработки прерываний производится следующим образом. В конце последнего машинного цикла выполнения очередной команды основной программы процессор проверяет наличие требования прерывания от ВУ. Если сигнал прерывания есть и в процессоре прерывание разрешено, то процессор переключается на выполнение подпрограммы обработки прерываний.

После сохранения содержимого регистров процессора, используемых в подпрограмме, начинается последовательный опрос регистров состояния контроллеров всех ВУ, работающих в режиме прерывания. Как только подпрограмма обнаружит готовое к обмену ВУ, сразу выполняются действия по его обслуживанию. Завершается подпрограмма обработки прерывания после опроса готовности всех ВУ и восстановления содержимого регистров процессора.

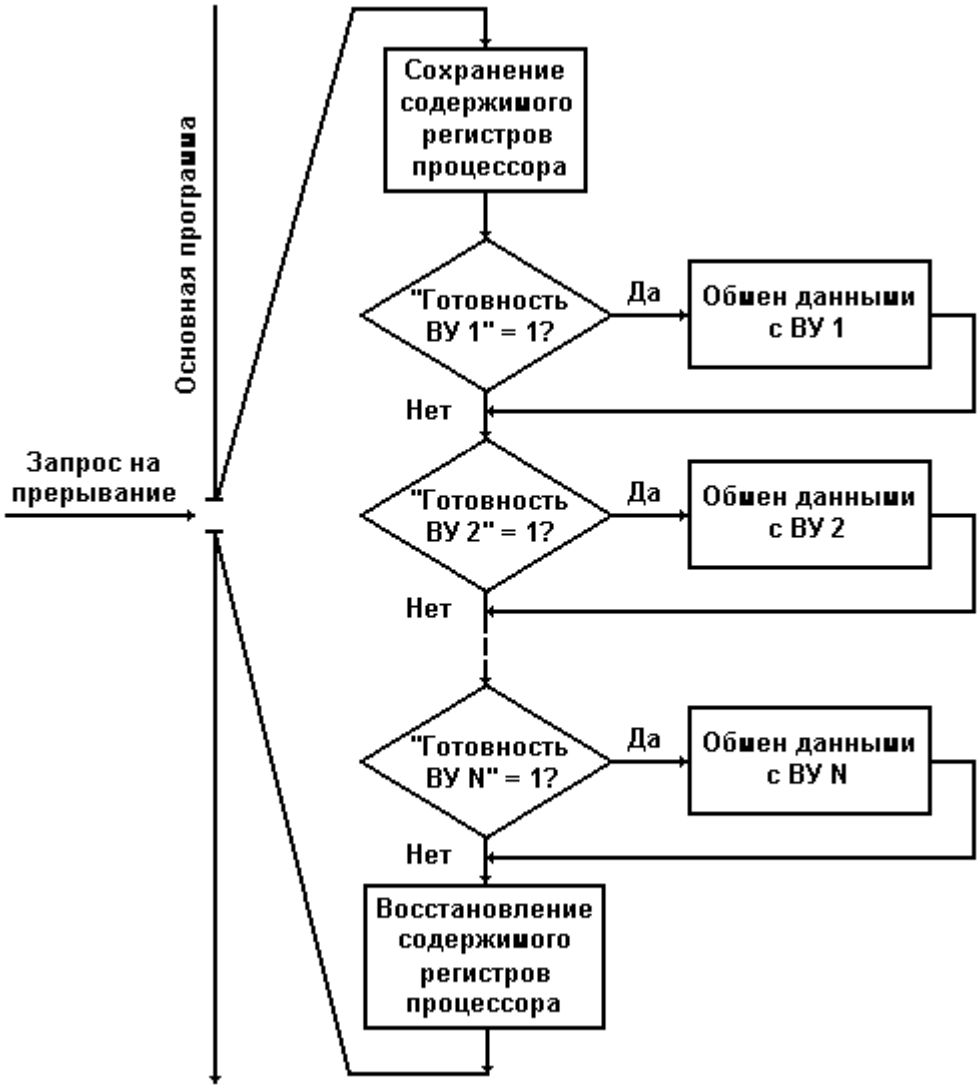


Рисунок 5.8 — Структура единой программы обработки прерываний и ее связь с основной программой

Приоритет ВУ в микроЭВМ с программным опросом готовности внешнего устройства однозначно определяется порядком их опроса в подпрограмме обработки прерываний. Чем раньше в подпрограмме опрашивается готовность ВУ, тем меньше время реакции на его запрос и выше приоритет. Необходимость проверки готовности всех внешних устройств существенно увеличивает время обслуживания тех ВУ, которые опрашиваются последними. Это является основным недостатком рассматриваемого способа организации прерываний. Поэтому обслуживание прерываний с опросом готовности ВУ используется только в тех случаях, когда отсутствуют жесткие требования на время обработки сигналов прерывания внешних устройств.

Организация системы прерываний в микроЭВМ с использованием векторов прерываний позволяет устранить указанный недостаток. При такой организации системы прерываний ВУ, запросившее обслуживания, само идентифицирует себя с помощью вектора прерывания — адреса

ячейки основной памяти микроЭВМ, в которой хранится либо первая команда подпрограммы обслуживания прерывания данного ВУ, либо адрес начала такой подпрограммы. Таким образом, процессор, получив вектор прерывания, сразу переключается на выполнение требуемой подпрограммы обработки прерывания. В микроЭВМ с векторной системой прерывания каждое ВУ должно иметь собственную подпрограмму обработки прерывания.

Различают векторные системы с интерфейсным и внеинтерфейсным вектором. В первом случае вектор прерывания (или его адрес) формирует контроллер ВУ, запросивший обслуживание, во втором — контроллер прерываний, общий для всех устройств, работающих в режиме прерываний (IBM-совместимые персональные компьютеры).

Рассмотрим организацию векторной системы с интерфейсным вектором. Вектор прерывания (или его адрес) выдается контроллером не одновременно с запросом на прерывание, а только по разрешению процессора, как это реализовано в схеме на рис. 5.9.

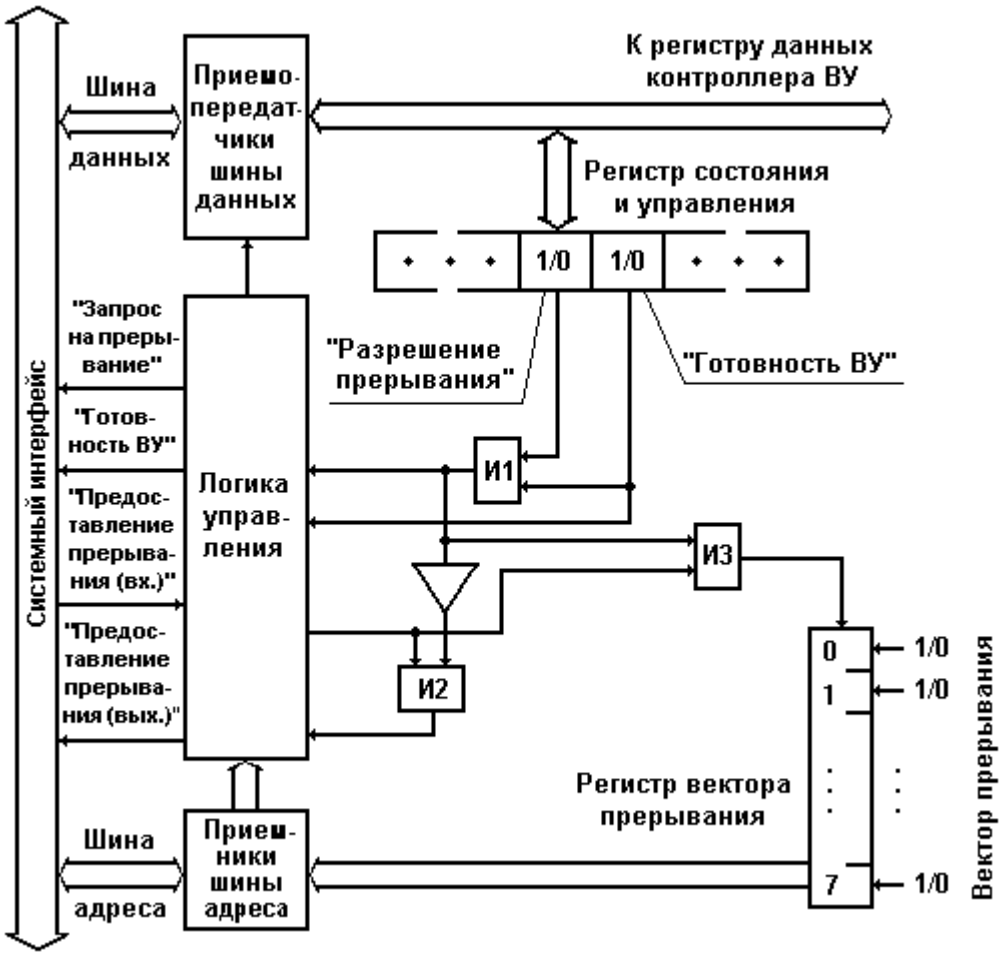
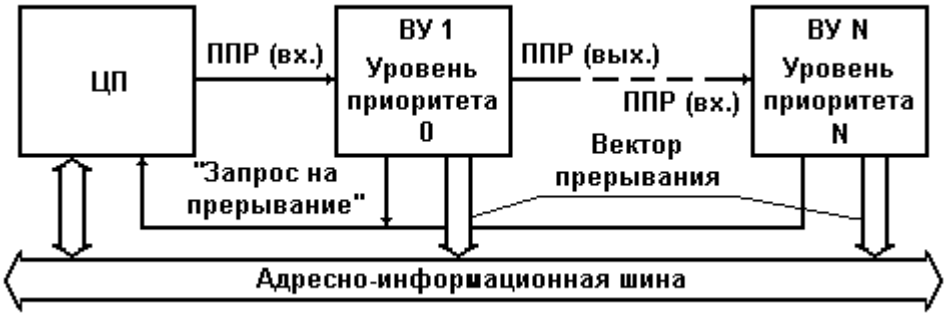


Рисунок 5.9 — Формирование векторов прерывания в контроллере ВУ

Это делается для того, чтобы исключить одновременную выдачу векторов прерывания от нескольких ВУ. В ответ на сигнал контроллера ВУ «Запрос на прерывание» процессор формирует управляющий сигнал «Предоставление прерывания (вх.)», который разрешает контроллеру ВУ,

запросившему обслуживание, выдачу вектора прерывания в шину адреса системного интерфейса. Для этого в контроллере используются регистр вектора прерывания и схема совпадения И3. Регистр вектора прерывания обычно реализуется с помощью перемычек или переключателей, что позволяет пользователю устанавливать для конкретных ВУ требуемые значения векторов прерывания.

Управляющий сигнал «Предоставление прерывания (вых.)» формируется в контроллере ВУ с помощью схемы совпадения И2. Этот сигнал используется для организации последовательного аппаратного опроса готовности ВУ и реализации тем самым требуемых приоритетов ВУ. Процессор при поступлении в него по общей линии системного интерфейса «Запрос на прерывание» сигнала прерывания формирует управляющий сигнал «Предоставление прерывания (вх.)», который поступает сначала в контроллер ВУ с наивысшим приоритетом (рис. 5.10). Если это устройство не требовало обслуживания, то его контроллер пропускает сигнал «Предоставление прерывания» на следующий контроллер, иначе дальнейшее распространение сигнала прекращается и контроллер выдает вектор прерывания на адресно-информационную шину.



ППР (вх.) — «Предоставление прерывания (входной)»;
«ППР (вых.) — Предоставление прерывания (выходной)»

Рисунок 5.10 — Реализация приоритетов ВУ в микроЭВМ с векторной системой прерываний с интерфейсным вектором

Аппаратный опрос готовности ВУ производится гораздо быстрее, нежели программный. Но если обслуживания запросили одновременно два или более ВУ, обслуживание менее приоритетных ВУ будет отложено на время обслуживания более приоритетных, как и в системе прерывания с программным опросом.

Рассмотренная векторная система прерываний практически полностью соответствует системе прерываний, реализованной в микроЭВМ «Электроника-60». Восьмиразрядный вектор прерывания в «Электронике-60» указывает одну из ячеек памяти с адресами от 0 до (376)8, в которой размещается адрес начала подпрограммы обработки прерывания. В следующей за указанной вектором прерывания ячейке

памяти хранится новое содержимое регистра состояния процессора, загружаемое в него при переключении на подпрограмму обработки прерывания. Один из бит нового содержимого регистра состояния процессора запрещает или разрешает прерывания от других ВУ, что позволяет ВУ с более высоким приоритетом прерывать подпрограммы обслуживания ВУ с меньшим приоритетом и наоборот.

Векторная система с внеинтерфейсным вектором прерывания используется в IBM-совместимых персональных компьютерах. В этих компьютерах контроллеры внешних устройств не имеют регистров для хранения векторов прерывания, а для идентификации устройств, запросивших обслуживания, используется общий для всех ВУ контроллер прерываний. Ниже приведен пример контроллера прерываний INTEL 8259A.

БИС программируемого контроллера прерываний (ПКП) представляет собой устройство, реализующее до восьми уровней запросов на прерывания с возможностью программного маскирования и изменения порядка обслуживания прерываний. За счет каскадного включения БИС ПКП число уровней прерывания может быть расширено до 64 (в архитектуре персонального компьютера IBM PC AT — 16).

Структурная схема ПКП приведена на рисунке 5.11.

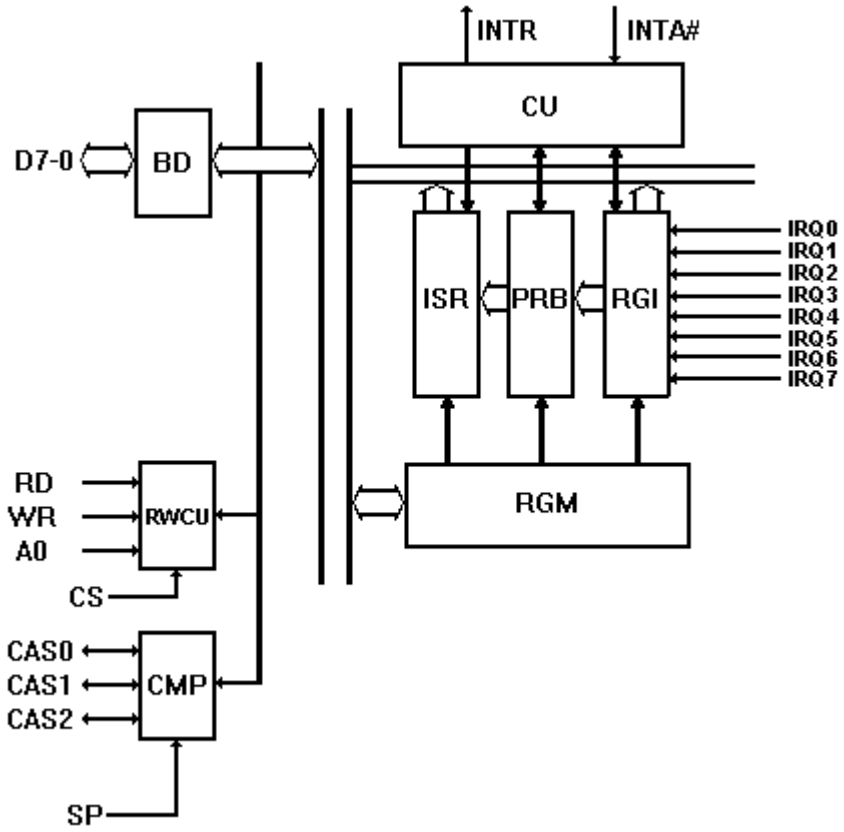


Рисунок 5.11 — Контроллер прерываний Intel 8259A

В состав БИС входят:

RGI — регистр запретов прерываний; хранит все уровни, на которые поступают запросы IRQx;

PRB — схема принятия решений по приоритетам; схема идентифицирует приоритет запросов и выбирает запрос с наивысшим приоритетом;

ISR — регистр обслуживаемых прерываний; сохраняет уровни запросов прерываний, находящиеся на обслуживании ПКП;

RGM — регистр маскирования прерываний; обеспечивает запрещение одной или нескольких линий запросов прерывания;

BD — буфер данных; предназначен для сопряжения ПКП с системной шиной данных;

RWCU — блок управления записью/чтением; принимает управляющие сигналы от микропроцессора и задает режим функционирования ПКП;

CMP — схема каскадного буфера-компаратора; используется для включения в систему нескольких ПКП;

CU — схема управления; вырабатывает сигналы прерывания и формирует трехбайтовую команду CALL для выдачи на шину данных.

Установка ПКП в исходное состояние и «настройка» его на определенный режим обслуживания прерываний происходит с помощью двух типов команд: команд инициализации (ICW) и команд управления операциями (OCW).

Программируемый контроллер прерываний (ПКП) имеет 16 входов запросов прерываний (IRQ 0—IRQ 15). Контроллер состоит из двух каскадно включенных контроллеров — выход INTR (запрос на прерывание) второго контроллера подключен ко входу IRQ 2 первого контроллера. В качестве примера отметим, что к линии IRQ 0 подключен системный таймер, к линии IRQ 1 — клавиатура, к линии IRQ 8 — часы реального времени и т.д.

Упрощенная схема взаимодействия контроллера прерываний с процессором и контроллером шины имеет следующий вид.

Рисунок 5.12 — Упрощенная схема взаимодействия контроллера

Эта схема функционирует следующим образом. Пусть в некоторый момент времени контроллер клавиатуры с помощью единичного сигнала по линии IRQ 1 известил контроллер прерываний о своей готовности к обмену. В ответ на запрос контроллер прерываний генерирует сигнал INTR (запрос на прерывание) и посылает его на соответствующий вход процессора. Процессор, если маскируемые прерывания разрешены (т.е. установлен флаг прерываний IF в регистре флагов процессора), посылает на контроллер шины сигналы R# — чтение, C# — управление и IO# — ввод/вывод, определяющие тип цикла шины. Контроллер шины, в свою очередь, генерирует два сигнала подтверждения прерывания INTA# и направляет их на контроллер прерываний. По второму импульсу контроллер прерываний выставляет на шину данных восьмибитный номер вектора прерывания, соответствующий данной линии IRQ.

В режиме реального адреса («реальном» режиме) векторы прерываний хранятся в таблице векторов прерываний, которая находится в первом килобайте оперативной памяти. Под каждый вектор отведено 4 байта (2 байта под адрес сегмента и 2 байта под смещение), т.е. в таблице может содержаться 256 векторов. Адрес вектора в таблице — номер вектора * 4.

Далее процессор считывает номер вектора прерывания. Сохраняет в стеке содержимое регистра флагов, сбрасывает флаг прерываний IF и помещает в стек адрес возврата в прерванную программу (регистры CS и IP). После этого процессор извлекает из таблицы векторов прерываний адрес подпрограммы обработки прерываний для данного устройства и приступает к ее выполнению.

Процедура обработки аппаратного прерывания должна завершаться командой конца прерывания EOI (End of Interruption), посылаемой контроллеру прерываний. Для этого необходимо записать байт 20h в порт 20h (для первого контроллера) и в порт A0h (для второго).

В IBM PC/XT/AT используется режим прерываний с фиксированными приоритетами. Высшим приоритетом обладает запрос по линии IRQ 0, низшим — IRQ 7. Так как второй контроллер подключен к линии IRQ 2 первого контроллера, то приоритеты линий IRQ в порядке убывания приоритета располагаются следующим образом: IRQ 0, IRQ 1, IRQ 8 — IRQ 15, IRQ 3 — IRQ 7. Если запрос на обслуживание посылают одновременно два устройства с разными приоритетами, то контроллер обслуживает запрос с большим приоритетом, а запрос с меньшим приоритетом блокирует. Блокировка сохраняется до получения команды EOI.

5.3 Последовательные каналы связи

Для передачи данных по электропроводным кабелям требуется выполнить следующие операции:

- синхронизировать тактовые частоты передатчика и приемника;
- преобразовать цифровые сигналы в аналоговые;
- сузить спектр электрических сигналов с помощью фильтров;
- обеспечить передачу отфильтрованных сигналов по линии связи;
- усилить принятые сигналы и восстановить их первоначальную форму;
- преобразовать аналоговые сигналы в цифровые.

Рассмотрим, как связаны между собой тактовая частота и последовательность битов. Тактовая частота, измеряемая в герцах, — это число изменений сигнала в единицу времени. Скорость же передачи данных равна числу передаваемых битов в секунду. Кажущаяся аналогия этих параметров породила множество ошибочных представлений. На самом деле, численное равенство скорости передачи данных и тактовой частоты является лишь частным случаем — когда применяется простейшее кодирование сигнала. Для увеличения скорости передачи используются более сложные способы кодирования — электрический сигнал может иметь два, три, пять и более уровней и передаваться по нескольким линиям параллельно, например по всем четырем витым парам кабеля. Каждому протоколу соответствует определенная ширина спектра сигнала и требуется некая пропускная способность информационной магистрали. Однозначного соответствия между мегагерцами и мегабитами в секунду не существует. Скорость передачи данных может равняться тактовой частоте либо в два, четыре и более раз превосходить ее, что зависит от метода кодирования. Рассмотрим методы кодирования в порядке их усложнения.

RZ

Сущность этого трехуровневого кода заключена в его названии — кодирование с возвратом к нулю (Return to Zero) (рис. 5.13). Логическому нулю соответствует положительный импульс напряжения, логической единице — отрицательный. Информационный переход осуществляется в начале бита, возврат к нулевому уровню — в середине.

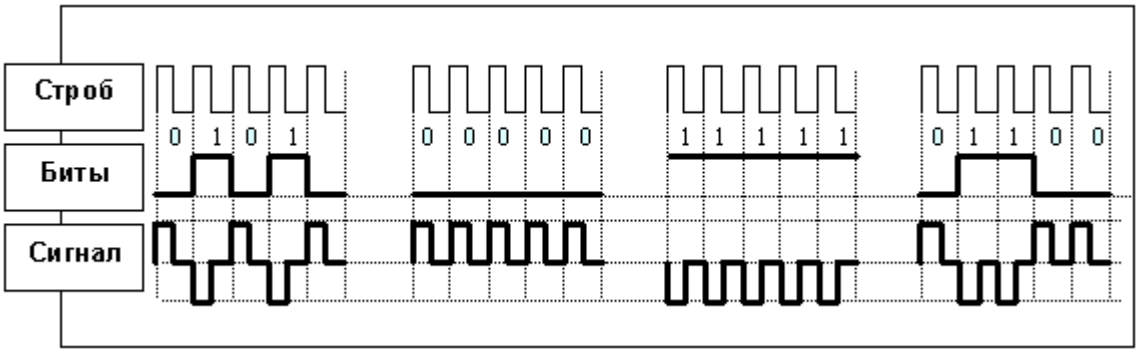


Рисунок 5.13 — Трехуровневый код RZ

Основной характеристикой кода RZ является то, что в середине каждого бита всегда есть переход (положительный или отрицательный), обозначающий каждый бит. Нужный для обработки сигнала синхроимпульс (строб) выделяется приемником из самого сигнала. Коды со стробом называются самосинхронизирующимися.

Код RZ не отличается высокой плотностью передачи данных — при тактовой частоте 10 МГц она равна всего 10 Мбит/с. К тому же, чтобы различать три уровня сигнала, на входе приемника необходимо обеспечить лучшее отношение сигнал/шум, чем при использовании двух уровней.

Наиболее часто код RZ применяется в оптоволоконных линиях связи. Однако при передаче используются три уровня мощности световых импульсов, поскольку оптические сигналы не бывают положительными или отрицательными.

Манчестерский код

Манчестерский код, или Манчестер-II, получил наибольшее распространение в локальных сетях. Он, как и RZ, является самосинхронизирующимся кодом, но в отличие от него имеет не три, а два уровня, что обеспечивает лучшую помехозащищенность канала. Логическому нулю соответствует переход на верхний уровень в центре битового интервала, логической единице — переход на нижний уровень.

Логика кодирования хорошо видна на примере передачи последовательности единиц или нулей (рис. 5.14). При передаче чередующихся битов частота следования импульсов уменьшается в два раза. Изменения сигнала в середине бита остаются, а на границе битовых интервалов отсутствуют. Эта логическая функция выполняется с помощью последовательности запрещающих импульсов, которые синхронизируются с информационными импульсами и запрещают нежелательные граничные переходы.

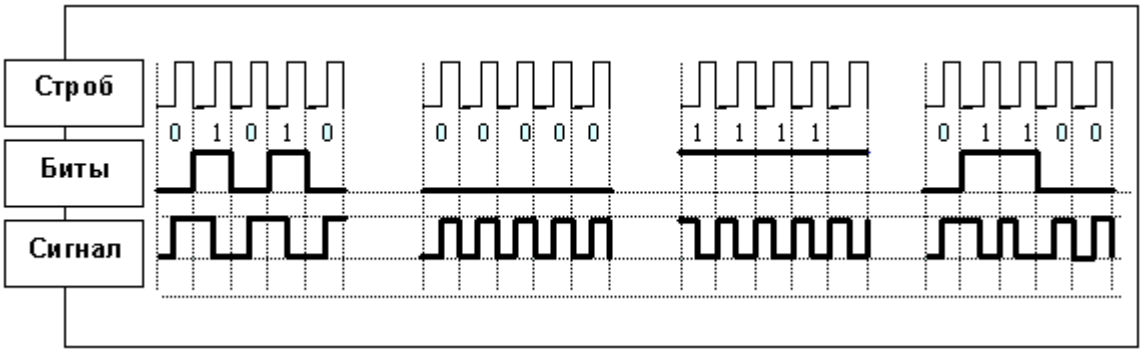


Рисунок 5.14 — Двухуровневый код Манчестер 2

Важная характеристика манчестерского кода — отсутствие у сигнала постоянной составляющей при передаче длинной последовательности

единиц или нулей. Благодаря этому передатчики и приемники можно «развязать» гальванически с помощью импульсных трансформаторов.

Спектр сигнала при манчестерском кодировании содержит только две частотные составляющие. Для десятимегабитового протокола это 10 МГц при передаче последовательности одних нулей или единиц и 5 МГц при их чередовании. Поэтому все другие частоты можно удалить с помощью полосовых фильтров.

Код Манчестер-II нашел применение в оптоволоконных и электропроводных сетях. Самый распространенный протокол локальных сетей — Ethernet 10 Мбит/с использует именно его.

Код NRZ

Код NRZ (Non Return to Zero), т.е. без возврата к нулю, является простейшим двухуровневым кодом (рис. 5.15). Нулю здесь соответствует нижний уровень сигнала, единице — верхний. Информационные переходы совпадают с границей битов. Вариант кода — NRZI (Non Return to Zero Inverted) имеет обратную полярность.

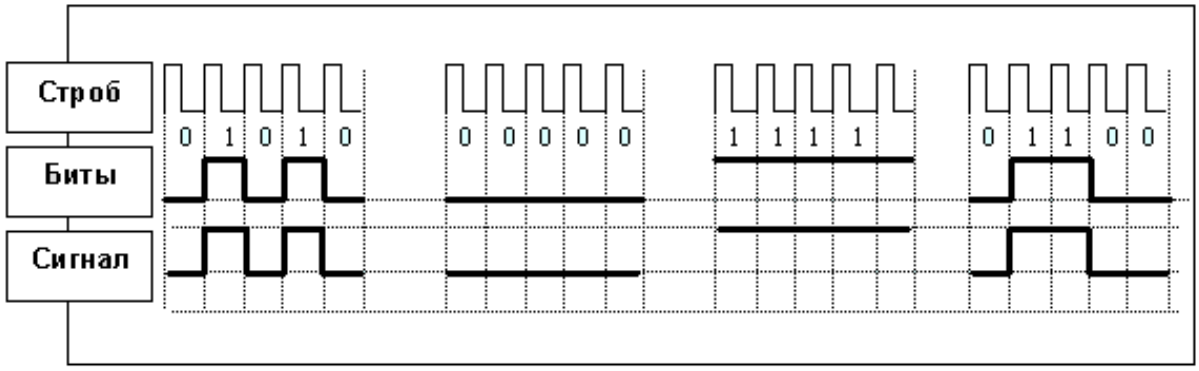


Рисунок 5.15 — Двухуровневый код NRZ

Несомненное достоинство кода — его простота: сигнал не надо кодировать и декодировать. Кроме того, числовое значение скорости передачи данных вдвое превышает тактовую частоту. Максимальная же частота спектра соответствует чередованию нулей и единиц. Для других комбинаций частота будет меньше, а при передаче последовательности одинаковых битов сигнал вовсе отсутствует.

Код NRZ и его разновидность NRZI не обеспечивают синхронизации между передатчиком и приемником, и это является самым большим его недостатком.

Для синхронизации начала приема пакета используется стартовый служебный бит, например единица. Наиболее известное современное применение кода NRZI — стандарт ATM155. Многие годы популярным был протокол связи через последовательный порт компьютеров — RS232A, тоже использующий код NRZ.

MIL-STD-1553 (MIL-STD-1553B)

Стандарт Министерства обороны США; распространяется на магистральный последовательный интерфейс (МПИ) с централизованным управлением, применяемый в системе электронных модулей.

Изначально разрабатывался по заказу МО США для использования в военной бортовой авионике, однако позднее спектр его применения существенно расширился, стандарт стал применяться и в гражданских системах.

Особенностью интерфейса является двойная избыточная линия передачи информации, полудуплексный протокол «команда-ответ» и до 31 удалённого абонента (оконечного устройства). Каждая линия управляется своим контроллером канала.

Стандарт устанавливает требования:

- к составу технических средств интерфейса;
- организации контроля передачи информации;
- характеристикам линии передачи информации (ЛПИ);
- характеристикам устройств интерфейса;
- интерфейсу с резервированием.

Впервые опубликован в США как стандарт ВВС в 1973 году, применён на истребителе F-16. Принят в качестве стандарта НАТО — STANAG 3838 AVS. Вытеснен более новым стандартом FireWire (MIL-STD-1394B), хотя старый всё ещё применяется (в том числе и для разработки).

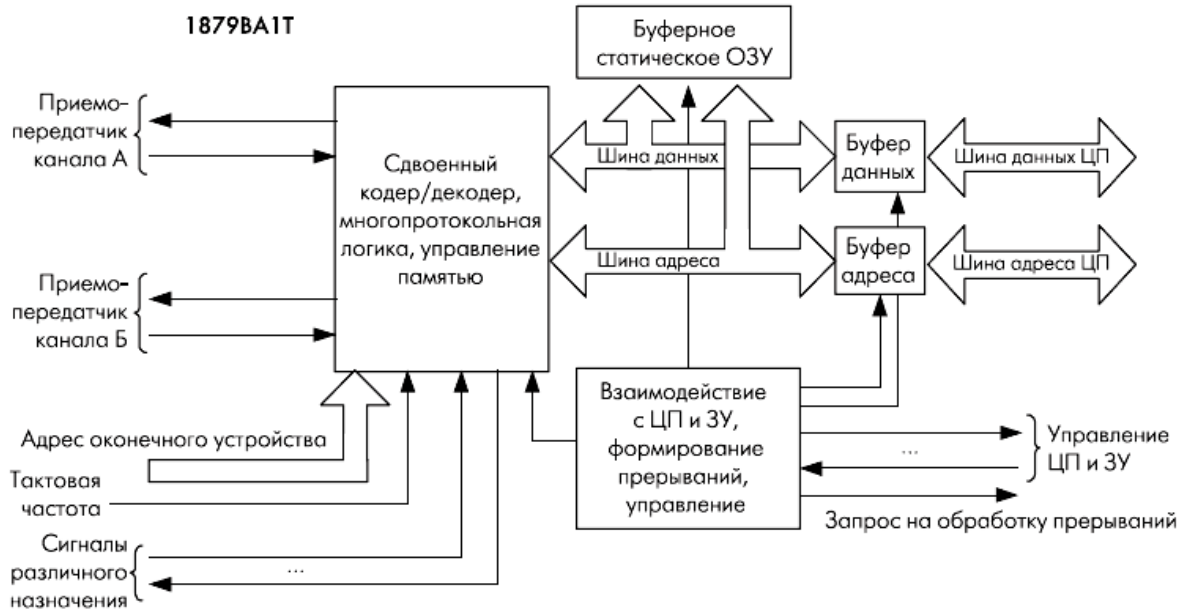


Рисунок 5.16 — Функциональная схема БИС 1879ВА1Т

Из статьи журнала «ЭЛЕКТРОНИКА: Наука, Технология, Бизнес 2/2005»:

«Известный российский дизайн-центр НТЦ «Модуль» в феврале 2005 года анонсировал свою новую БИС 1879ВА1Т — терминал мультиплексного канала обмена по ГОСТ Р 52070-2003. Однако гораздо шире этот стандарт известен под названием MIL-STD-1553B (обозначение военного стандарта США)».

5.4 FIFO-буфер

В вычислительных системах используются подсистемы с различным быстродействием и, в частности, с различной скоростью передачи данных (рис. 5.17). Обычно обмен данными между такими подсистемами реализуется с использованием прерываний или канала прямого доступа к памяти. В первую очередь подсистема 1 формирует запрос на обслуживание по мере готовности данных к обмену. Однако обслуживание прерываний связано с непроизводительными потерями времени, и при пакетном обмене производительность подсистемы 2 заметно уменьшается. При обмене данными с использованием канала прямого доступа к памяти подсистема 1 передает данные в память подсистемы 2. Данный способ обмена достаточно эффективен с точки зрения быстродействия, но для его реализации необходим довольно сложный контроллер прямого доступа к памяти.

Рисунок 5.17 — Применение буферной памяти

Наиболее эффективно обмен данными между подсистемами с различным быстродействием реализуется при наличии между ними специальной буферной памяти. Данные от подсистемы 1 временно запоминаются в буферной памяти до готовности подсистемы 2 принять их. Емкость буферной памяти должна быть достаточной для хранения тех блоков данных, которые подсистема 1 формирует между считываниями их подсистемой 2. Отличительной особенностью буферной памяти является запись данных с быстродействием и под управлением подсистемы 1, а считывание — с быстродействием и под управлением подсистемы 2 («эластичная память»). В общем случае память должна выполнять операции записи и считывания совершенно независимо и даже одновременно, что устраняет необходимость синхронизации подсистем. Буферная память должна сохранять порядок поступления данных от подсистемы 1, т.е. работать по принципу «первое записанное слово

считывается первым» (First Input First Output — FIFO). Таким образом, под буферной памятью типа FIFO понимается запоминающее устройство с произвольной выборкой (ЗУПВ), которое автоматически следит за порядком поступления данных и выдает их в том же порядке, допуская выполнение независимых и одновременных операций записи и считывания.

5.5 Модем

Модем — это устройство, которое позволяет обмениваться данными по телефонной линии.

Если компьютеры расположены слишком далеко и их нельзя соединить стандартным сетевым кабелем, связь между ними осуществляется с помощью модема. В сетевой среде модемы служат для соединения отдельных сетей между собой или между ЛВС и остальным миром. Осуществлять связь напрямую через телефонную линию компьютеры не могут, так как обмениваются данными с помощью цифровых электронных импульсов, а по телефонной линии можно передавать только аналоговые сигналы (звуки).

Цифровой сигнал может принимать лишь два значения — 0 или 1. Аналоговый сигнал — это плавная кривая, которая может иметь бесконечное множество значений. Модем на передающей стороне преобразует цифровые сигналы в аналоговые и передаёт их по телефонной линии. Модем на принимающей стороне преобразует приходящие аналоговые сигналы в цифровые для компьютера-получателя. Другими словами, передающий модем преобразует цифровой сигнал в аналоговый, а принимающий модем преобразует аналоговый сигнал в цифровой.

Аппаратное обеспечение модемов

Модемы имеют два стандартных физических интерфейса:

– последовательный интерфейс передачи данных (RS-232);

– интерфейс с телефонной линией RG-11 (четырёхконтактный телефонный разъём).

Существуют внутренние и внешние модемы. Внутренние модемы устанавливаются в слоты расширения на материнской плате подобно другим платам.

Внешний модем представляет собой коробочку, подключаемую к компьютеру с помощью последовательного (RS-232) кабеля. Этот кабель соединяет последовательный порт компьютера с тем разъёмом модема, который предназначен для связи с компьютером. Для подключения модема к телефонной линии используется кабель с разъёмом RG-11.

Стандарты модемов

Промышленные стандарты существуют практически для каждой области сетевых технологий и модемы не являются исключением.

Стандарты обеспечивают взаимодействие модемов от разных производителей. Спецификации, известные как V-серии, включают номер стандарта. Иногда включается так же слово «bis». Оно указывает, что данный стандарт — пересмотренная версия более раннего стандарта. Если в названии присутствует слово «terbo», это означает, что второй-«bis» стандарт также был модифицирован.

Производительность модема

Изначально скорость модемов измерялась в битах в секунду или в единицах, называемых «бод». Многие путали их, считая, что они обозначают одно и то же. На самом деле бод относится к частоте осцилляций звуковой волны, переносящих биты данных по телефонной линии. В начале 1980-х годов скорость в бодах равнялась скорости передачи модемов. Затем инженеры разработали методы сжатия и кодирования информации. В результате каждая модуляция звука могла переносить больше одного бита информации, следовательно, скорость передачи в битах в секунду может быть больше, чем скорость в бодах, поэтому необходимо сначала обратить внимание на скорость в битах в секунду, а затем в бодах. Например модем на скорости 28 800 бод в действительности может передавать данные со скоростью 115 200 бит/с. Современные модемы имеют такие индустриальные стандарты сжатия данных, как V.42bis/MNP5, и имеют скорость передачи данных 57 600 бит/с, а некоторые — 76 800 бит/с.

Типы модемов

Существуют разные типы модемов, так как существуют разные среды передачи, для которых требуются разные методы передачи. Эти типы можно грубо разделить, взяв за основу критерий синхронизации связи. Связь бывает асинхронная и синхронная. Тип модема будет зависит от среды и от назначения сети.

Асинхронная связь

Асинхронная связь — самая распространённая форма передачи данных. Причина такой популярности заключается в использовании этим методом стандартных телефонных линий. При асинхронной передаче данные передаются последовательным потоком. Каждый символ — буква, число или знак — раскладывается в последовательность битов. Каждая такая последовательность отделяется от другой стартовым и стоповым битами. Передающее и принимающее устройства должны согласовывать последовательность стартовых и стоповых битов. Связь этого типа не синхронизируется, передающий компьютер передаёт, а принимающий получает без координации взаимодействия устройств. Затем принимающий компьютер проверяет полученные данные на наличие ошибок и принимает следующий блок информации. 25 % трафика уходит на передачу согласующей информации.

Контроль ошибок

Вероятность ошибок никогда не исключена, поэтому в асинхронной передаче используется специальный бит-бит чётности. Схема проверки и коррекции ошибок, которая его применяет, называется контролем чётности. При контроле чётности количество посланных и принятых единичных битов должно совпадать.

Стандарт модемов V.32 не предусматривал контроль ошибок. Чтобы решить эту проблему, компания Microcom создала собственный стандарт асинхронного контроля ошибок данных, который был назван Microcom Network Protocol (MNP). Этот метод оказался настолько удачным, что и другие компании заимствовали не только начальную версию его, но и другие версии, называемые классами. В настоящее время используется MNP классов 2, 3 и 4.

В 1989 г. комитет CCITT опубликовал схему асинхронного контроля ошибок, названную V.42. Этот стандарт аппаратной коррекции ошибок включает в себя два протокола. Основная схема контроля ошибок — это Link Acces Procedure for Modem (LAPM), однако V.42 так же использует MNP4. Протокол LAPM используется для соединения модемов по стандарту V.42, однако если один из модемов поддерживает только стандарт MNP4, будет использоваться MNP4.

Увеличение скорости передачи

Алгоритм коррекции/сжатия

При передаче информации с использованием протокола коррекции (MNP4, v.42) происходит обрезание 10 бит, полученных из компьютера, до 8-ми информационных (удаляются стартовый и стоповый биты) (10 бит = старт_бит + 8 информационных + стоп_бит — см. Асинхронный протокол RS232). И наоборот, при получении из линии 8-ми информационных бит модем их преобразует в 10 и передает в компьютер. Таким образом, по линии идет информации меньше, чем модем получил из компьютера. Но это еще не все. При использовании протокола сжатия (MNP5, v.42bis) происходит еще и уменьшение объема полезной информации, так что от тех 10-ти бит, что модем получил от компьютера, в линию (и на удаленный модем) попадет только часть.

На производительность канала связи оказывают влияние два фактора:

– скорость канала — характеризует, насколько быстро биты кодируются и передаются по каналу связи;

– пропускная способность — характеризует долю полезной информации, передаваемой по каналу.

Скорость передачи и пропускная способность — не одно и то же. За счёт сжатия данных можно увеличить пропускную способность — сжатие уменьшает время, необходимое для передачи данных (за счёт удаления избыточных элементов и пустых участков). Одним из распространенных

протоколов сжатия данных является MNP5 — время передачи может быть сокращенно наполовину.

При использовании стандарта V.42bis можно добиться наибольшей производительности, так как он описывает аппаратную реализацию непрерывного сжатия информации. Пропускная способность на скорости 9600бит/с может достигать 38400бит/с. В настоящее время используются такие высокоскоростные протоколы, как x2 и V.90.

Комбинирование стандартов

Для увеличения производительности используют комбинацию протоколов передачи данных и коррекции ошибок. Например, при асинхронной передаче хорошие результаты даёт комбинация:

V.32bis — передача;

V.42 — коррекция ошибок;

V.42bis — сжатие.

Синхронная связь

Синхронная связь основана на схеме синхронизации, согласованной между двумя устройствами. Её цель — выделить биты из группы при передаче их блоками. Эти блоки называются кадрами. Для установления синхронизации и проверки правильности её работы используются специальные символы. Поскольку биты передаются в синхронном режиме, стартовые и стоповые биты не нужны. Передача завершается в конце одного кадра и начинается в начале другого. Этот метод более эффективен, чем асинхронная передача. В случае ошибки синхронная схема распознавания и коррекции ошибок повторяет передачу кадра.

Синхронные протоколы выполняют следующие действия, не предусмотренные асинхронными протоколами:

- разбивают данные на блоки;
- добавляют управляющую информацию;
- проверяют данные на наличие ошибок.

Основные протоколы синхронной передачи:

- SDLC — протокол синхронного управления каналом.
- HDLC — протокол высокоуровнего управления каналом.
- BISYNC — протокол двоичной синхронизированной связи.

Синхронная связь используется, в основном, на выделенных цифровых линиях и в домашних условиях, как правило, не применяется.

5.6 Pager

Протокол POCSAG

Одним из самых распространенных на сегодняшний день форматов пейджинговой передачи является протокол POCSAG (Post Office Code Standartisation Advisory Group), разработанный Британским почтовым

ведомством. Он предусматривает скорость передачи информации 512, 1200 и 2400 бит/сек. Сообщения передаются в асинхронном режиме: пакет сообщения может стартовать в любой момент времени и длина его не определена. Общая структура сигнала в формате POCSAG приведена на рис. 5.18.

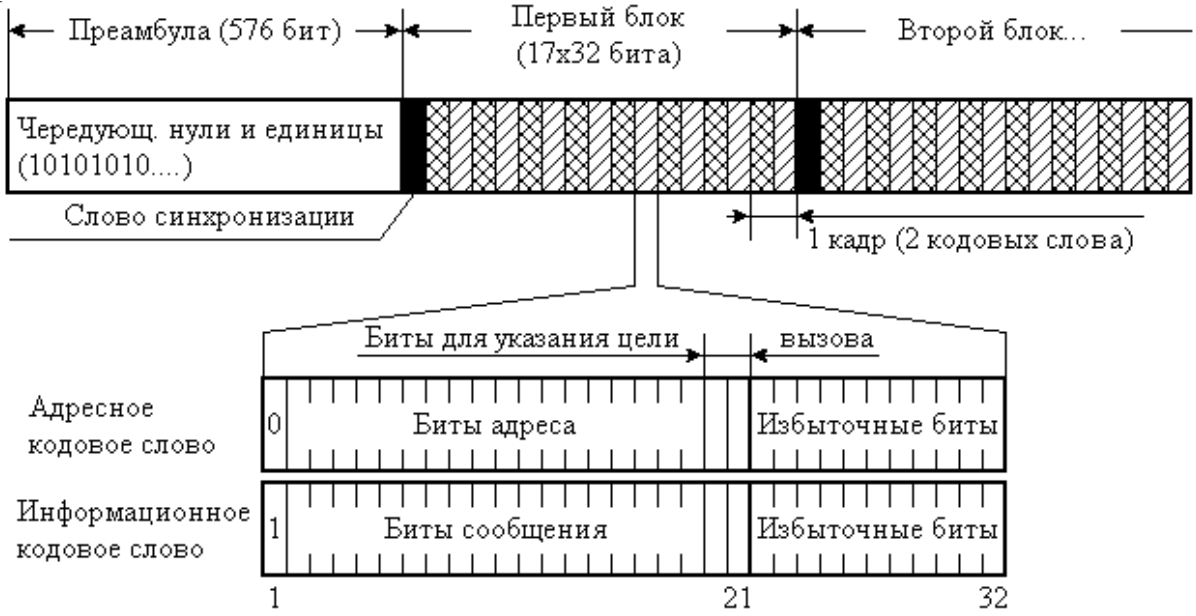


Рисунок 5.18 — Общая структура сигнала в формате POCSAG

Сигнал в формате POCSAG начинается с преамбулы, состоящей из 576 бит чередующихся 0 и 1. Преамбула служит для вывода приемного устройства (пейджера) из «спящего» состояния в режим «приема» и его тактовой синхронизации.

После преамбулы следует поток блоков, содержащих физические адреса пейджеров и тексты сообщений. Длина кодовой последовательности в формате POCSAG не определена, блоки следуют один за другим, каждый со своим кодовым словом синхронизации — для подстройки синхронизации приемников (при передаче длинных сообщений).

Каждый блок состоит из 17-ти 32-битных слов. Первое из них является словом синхронизации (фиксированная последовательность 32 бит), далее идет последовательность из восьми двойных слов или кадров (фреймов), нумеруемых с 0-го по 7-й.

Каждое 32-разрядное слово содержит 21 информационный бит и 11 избыточных (контрольных) бит, которые служат для определения и корректировки ошибок. Протокол предусматривает коррекцию ошибок по алгоритму БЧХ(32,12), при котором в одном 32-битном кодовом слове корректируется ошибка при приеме одного неверного бита (либо 2, если расстояние между ними не превышает 6 бит), а возникновение от 2 до 5 ошибок детектируется (т.е. это слово исключается и в большинстве моделей текстовых пейджеров эта часть обозначается скобками).

В зависимости от функционального назначения блоки делятся на адресные, в которых передается физический адрес пейджера, и информационные, содержащие текст сообщения.

Основное отличие протокола POCSAG от других протоколов пейджинговой передачи заключается в способе приема содержащегося в начале каждого пейджингового сообщения физического адреса пейджера — кэпкода (CapCode), которому оно адресовано. Все возможные 2 млн физических адресов разбиты на 8 групп, соответствующих 8 кадрам (frames) адресного блока.

Адресный блок состоит из адресного кодового слова и предшествующих «пустых» слов (специальные фиксированные 32-битовые последовательности) и формируется следующим образом: физический адрес пейджера делится на 8. Остаток от деления дает номер фрейма, в первое слово которого записывается частное от деления. Во все предыдущие фреймы записываются «пустые» слова (специальные фиксированные 32-битовые последовательности), а все оставшиеся до конца адресного блока слова пропускаются, т.е. сразу за адресным словом начинается следующий блок. Фактически остаток от деления является номером интервала времени, в котором данный пейджер будет вести прием и распознавание своего номера.

Пейджер принимает только кадры, соответствующие его адресу. Это позволяет в восемь раз увеличить адресную емкость системы и начительно повысить срок службы элементов питания.

Информационный блок служит для передачи цифровой и алфавитно-цифровой информации на пейджер, заданный адресным блоком. Он содержит слово синхронизации, информационные слова и, если сообщение закончилось, «пустые» слова до конца блока.

В протоколе POCSAG не оговаривается, какие физические значения сигнала принимаются за 0, а какие за 1. Поэтому различные пейджеры (или режимы приема пейджера) воспринимают эту кодировку с точностью до наоборот. Отсюда появилось понятие инверсной кодировки POCSAG. Инверсная кодировка POCSAG полностью совпадает с описанной выше, за исключением того, что нулевые биты заменяются единичными, а единичные биты — нулевыми.

Увеличение скорости передачи сообщений ведет к увеличению пропускной способности системы, однако при этом снижается устойчивость к помехам, а главное — снижается чувствительность радиоприема, т.е. фактически — радиус рабочей зоны приема сообщений. Для подавляющего большинства пейджеров чувствительность в зависимости от скорости передачи равна следующим значениям: 512 бит/сек — 5 мкВ/м; 1200 бит/сек — 7 мкВ/м; 2400 бит/сек — 9 мкВ/м.

Протокол FLEX

Протокол пейджинговой связи FLEX разработан формой Motorola. Основным достоинством этого протокола является высокая скорость

передачи данных — 1600, 3200 и 6400 бит/сек, а следовательно, высокая пропускная способность. Так, если в стандарте POCSAG ресурс частоты составляет 10—15 тысяч абонентов, то во FLEX-системах ресурс частотного канала лежит в пределах 20—80 тысяч абонентов. В отличие от протокола POCSAG протокол FLEX использует синхронную передачу данных, т.е. синхронизация передатчика и приемника производится по абсолютному значению времени.

Структура формата FLEX приведена на рис. 5.19.

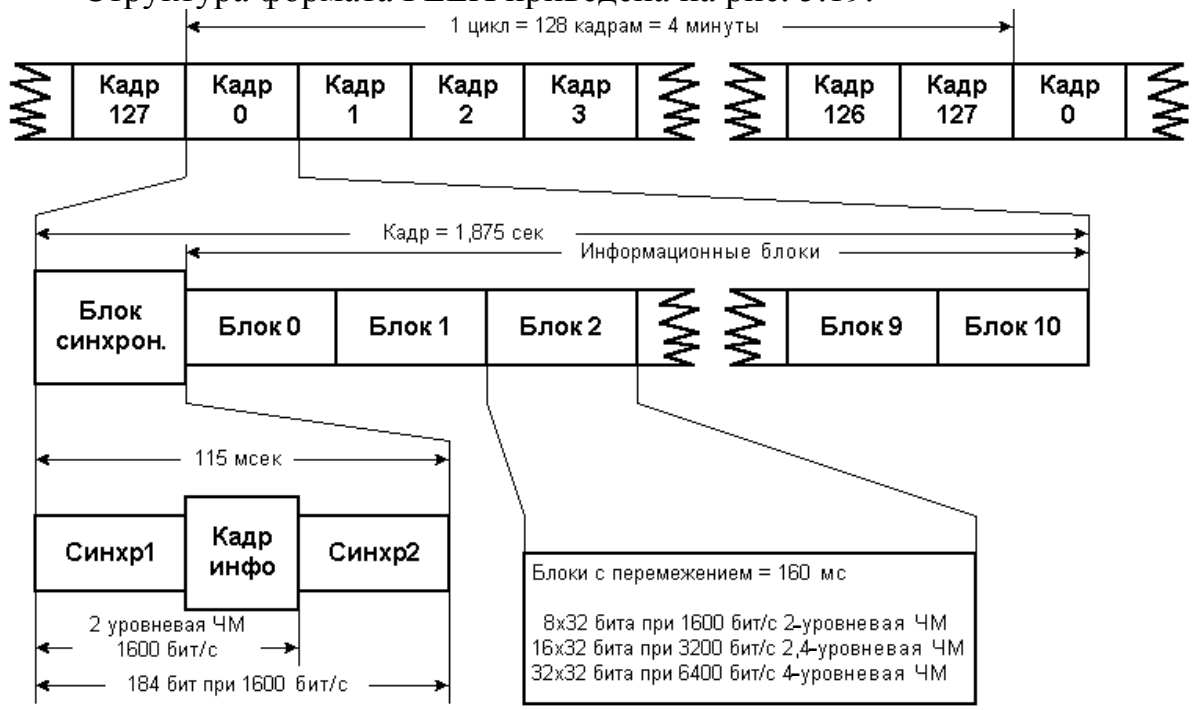


Рисунок 5.19 — Структура формата FLEX

При скорости 1600 бит/сек используется 2-уровневая частотная модуляция. При скорости 3200 бит/сек может использоваться как 2-уровневая, так и 4-уровневая частотная модуляция. При скорости 6400 бит/сек используется 4-уровневая частотная модуляция. Значения девиации для различных двоичных данных при 2-уровневой и 4-уровневой частотной модуляции приведены на диаграмме, рис. 5.20.

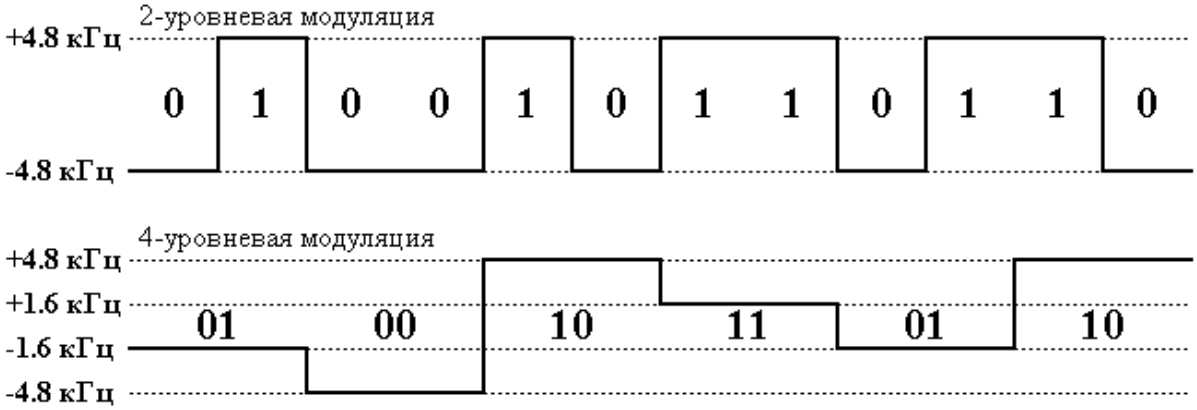


Рисунок 5.20 — Значение девиации для различных двоичных данных

Данные в протоколе FLEX формируются в кадры, которые передаются последовательно со скоростью 32 кадра в минуту (1,875 сек на кадр). Полный цикл протокола FLEX состоит из 128 кадров, которые нумеруются от 0 до 127 и передаются ровно 4 минуты. Каждый час делится на 15 циклов, пронумерованных от 0 до 14.

Так как протокол FLEX является синхронным, для его синхронизации используются сигналы точного времени, передаваемые в начале каждого часа в кадре 0 цикла 0. При передаче этого кадра осуществляется синхронизация приемников.

Каждый кадр протокола FLEX передается 1,875 сек и состоит из блока синхронизации (115 мсек) и 11 информационных блоков (по 160 мсек на блок).

Блок синхронизации обеспечивает синхронизацию кадра и настройку пейджеров (фрагменты «Синхрон. 1» и «Синхрон. 2»), а также несет информацию о номере цикла и кадра (фрагмент «Кадр инфо»).

Информационные блоки содержат служебную информацию, адресное поле, задающее адреса пейджеров, которым адресованы сообщения, векторное поле, указывающее, где расположены сообщения в поле сообщений и их длину, и непосредственно поле сообщений, содержащее сами сообщения.

Последовательность расположения полей в кадре показана на рис. 5.21.

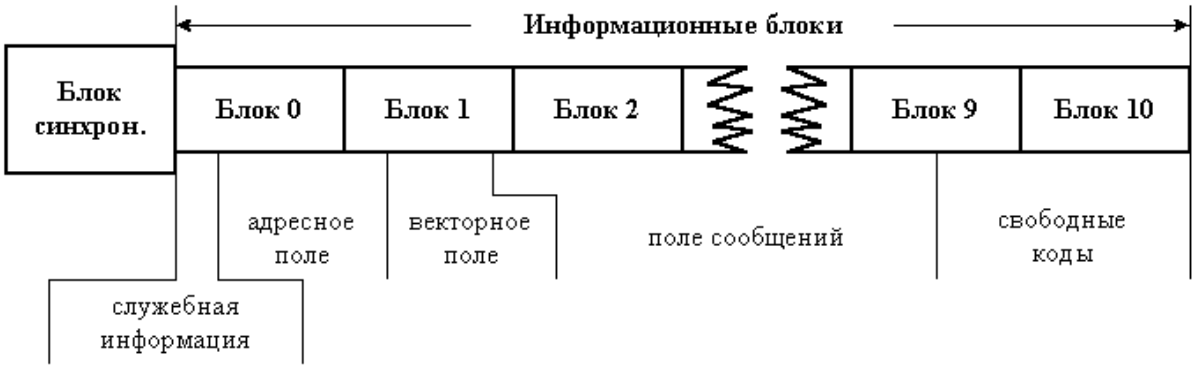


Рисунок 5.21 — Структура формата FLEX

Поля не привязаны к границам блока. Порядок расположения адресов пейджеров в адресном поле должен соответствовать порядку расположения векторов в векторном поле. Адреса пейджеров могут задаваться одним кодовым словом (короткий адрес), поддерживая при этом до 2 миллионов адресов, или двумя кодовыми словами (длинный адрес), поддерживая до 5 миллиардов адресов.

При кодировании информации используется код БЧХ, позволяющий восстанавливать единичные ошибки передачи данных. Кроме того, используемая в протоколе последовательность передачи сформированных

бит информации позволяет восстанавливать принятые данные при пропадании сигнала на интервале до 10 мсек.

Каждый пейджер, работающий с протоколом FLEX, может принимать сообщения на любой из допустимых скоростей передачи данных (1600, 3200 или 6400 бит/с). Одним из важных следствий синхронного протокола является то, что сообщения для каждого конкретного пейджера можно помещать в кадр с определенным номером. Это позволяет пейджеру избирательно принимать один или несколько кадров из всего четырехминутного цикла протокола FLEX, в которые помещаются сообщения на его адрес. Если пейджер не обнаруживает своего адреса в своем кадре, он прекращает прием. Такая организация связи позволяет резко повысить срок службы батареек пейджера.

Еще одной важной отличительной особенностью протокола FLEX является возможность работы совместно с другими протоколами связи. Для этого в цикле выделяются определенные кадры для работы по протоколу FLEX, а промежутки между ними отдаются для работы по другим протоколам, например POCSAG. Это позволяет компании-оператору, не создавая новой инфраструктуры, постепенно перейти от работы в протоколе POCSAG на работу в протоколе FLEX.

К достоинствам протокола FLEX следует отнести:

– повышенную скорость передачи данных, а следовательно, повышенную пропускную способность на один частотный канал;

– возможность поддержания большого количества адресов (до 5 миллиардов);

– улучшенные характеристики помехоустойчивости канала передачи;

– обеспечение энерго-экономичного режима работы пейджеров;

– возможность совместной работы с другими протоколами.

Протокол ERMES

Протокол ERMES был разработан как общеевропейский протокол пейджинговой связи. Он включает в себя, кроме собственно протокола передачи данных, ряд организационных положений и технических решений в рамках Меморандума о взаимопонимании, подписанного руководителями администраций 16 стран Европы в январе 1990 г. с целью координации усилий по созданию общеевропейской системы персонального радиовызова (СПРВ). К достоинствам протокола ERMES следует отнести следующее:

– повышенную скорость передачи данных, а, следовательно, повышенную пропускную способность на один канал;

– обеспечение энерго-экономичного режима работы пейджеров;

– возможность передачи произвольного набора данных объемом до 64 кбит;

– возможность удобной организации роуминга во всех регионах,

охваченных сетью ERMES.

Для функционирования СПРВ по протоколу связи ERMES выделяется единый диапазон частот (или его часть) 169,4—169,8 МГц, в котором организуются 16 частотных каналов с разносом частот в 25 кГц. Для приема сигнала используются сканирующие по частоте абонентские приемники (пейджеры). Скорость передачи данных составляет 6,25 кбит/сек.

Системы персонального радиовызова на базе протокола ERMES обеспечивают следующие услуги:

– передачу цифровых сообщений длиной до 1600 знаков;

– передачу буквенно-цифровых сообщений длиной до 9000 символов;

– передачу произвольного набора данных объемом до 64 кбит;

– возможность приема вызова и сообщений одним унифицированным приемником (пейджером) во всех странах, входящих в объединенную СПРВ ERMES.

Структура протокола ERMES приведена на рис. 5.22.

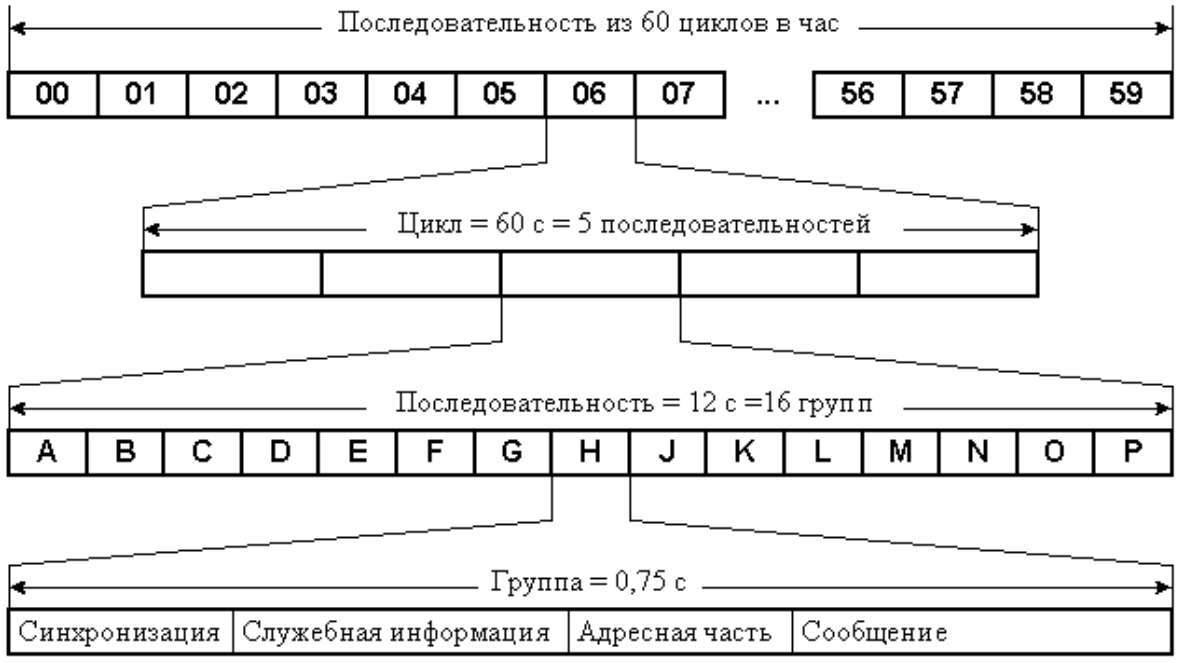


Рисунок 5.22 — Структура протокола ERMES

Каждый час передается 60 циклов по одной минуте каждый. Каждый минутный цикл содержит 5 последовательностей по 12 сек. Каждая из последовательностей включает в себя 16 типов групп, которые обозначаются буквами латинского алфавита от A до P. Каждая группа имеет длину 0,75 сек и состоит из четырех блоков: синхронизации; служебной информации; адреса; информационного сообщения.

Порядок расположения групп внутри последовательности для каждого частотного канала свой и устроен так, что позволяет пейджеру в

пределах последовательности просмотреть свою группу в режиме сканирования на всех 16 частотах.

Протокол ERMES использует помехоустойчивое кодирование передаваемой информации с прямой коррекцией ошибок (FEC), циклический код (30,18), кодовое расстояние Хемминга — 6.

Приемники персонального вызова (пейджеры) в системе ERMES работают следующим образом. Находясь в зоне приема «своей» базовой станции, пейджер принимает сообщения на ее частоте. При попадании в другой регион пейджер, не «слыша» сигнал на своей частоте, переходит в режим сканирования по каналам ERMES и, обнаружив сигнал, начинает принимать информацию на частоте базовой станции данного региона.

5.7 GPS

GPS (Global Positioning System) — это спутниковая система для высокоточного определения координат статичных и движущихся объектов. Разработана она и обслуживается Министерством обороны США, также известна у военных под кодовым названием NAVSTAR (Navigation Satellite Timing and Ranging).

Проект запущен в 1978 г. и вышел на запланированную мощность в 1994 г., получив высокие оценки военных во время войны в Заливе (особенно им были довольны группы специального назначения, уходившие из-под огневых налетов собственной артиллерии и авиации).

Система GPS в целом состоит из трех сегментов — космического, управляющего и пользовательского. К первому относятся 24 спутника, запущенных по шести различным орбитам таким образом, чтобы из любой точки земной поверхности были видны от четырех до двенадцати таких спутников. Срок службы каждого из них составляет 10 лет, их заменяют по мере выхода из строя. В управляющий (спутниками) сегмент GPS входят 5 контрольных центров (включая мастер-центр), дислоцированных на американских военных базах. И нетрудно догадаться, что к пользовательскому сегменту относятся десятки и сотни тысяч персональных GPS-приемников, которые продаются в виде автономных устройств, модулей расширения к портативным компьютерам или же встраиваются в определенные виды оборудования.

Приемник системы GPS представляет собой крошечный узкоспециализированный компьютер, способный вычислять свое местоположение по радиосигналам, принимаемым со спутника. И чем больше спутников может отслеживать такой приемник одновременно и чем больше разнесены эти спутники на небесной полусфере, тем быстрее пойдет процесс вычисления координат и тем более точными могут быть его показания. Способность приемника обрабатывать сигналы от нескольких спутников определяется числом его каналов, в современных устройствах их почти всегда 12. А для отслеживания спутников нужно быть под открытым небом — в помещении под крышей или в тесном

окружении высотных домов антенна приемника фактически беспомощна. Кстати, именно поэтому комплект для использования в транспортных средствах чаще всего снабжается внешней антенной, которая крепится снаружи, а многие модели GPS оснащены MMCX-штекером для их подключения. Встроенная антенна приемника обычно работает по узкой диаграмме направленности (patch-антенны), но ряд производителей освоил выпуск приемников с антеннами, которые имеют широкую диаграмму (multi-directional). Облачность влияния на сигнал не оказывает, стекло и пластик — тоже не помеха, поэтому GPS-приемник может спокойно пеленговать спутники с застекленного балкона, но при особо «удачном» выборе места и времени сигналы со спутников может блокировать даже... собственно владелец приемника!

Процесс определения координат GPS-приемником выглядит примерно так. При включении устройства после достаточно длительного перерыва приемник начинает принимать сигналы со спутников и тем самым определять, какие из них сейчас доступны из этой локации. Такое состояние приемника называется «холодным стартом», а группу запеленгованных спутников часто именуют «альманахом». После выключения приемник некоторое время держит в памяти последний альманах, и в случае повторного включения после кратковременного перерыва время пеленга существенно возрастает (имеет место «теплый старт»), а если перерыв был совсем кратким, то это «горячий старт». Термин TTFF (Time To First Fix), коим часто пользуются при описании этого этапа работы, как раз и означает время, необходимое для захвата того минимального числа спутников, которого достаточно для дальнейших вычислений, и оно указывается раздельно для холодного (обычно 1—2 минуты), теплого (до минуты) и горячего (до десяти секунд) старта. Но это — в тепличных условиях. В Киеве меньше 3 минут «холодный старт» вообще не происходил, а при особенно неудачном раскладе он может длиться от 5 до 50 минут.

Сами сигналы со спутников бывают двух видов (L1 и L2), все GPS гражданского назначения используют частоту L1=1575,42 MHz. Содержит такой сигнал, согласно текущей версии 2.2 стандарта NMEA 0183, три составляющие: псевдослучайный код (идентификатор спутника), собственно данные в формате GGA (статус готовности спутника, дата и время) и позиции всех спутников в течение дня в форматах GSA (Global Satellites Active — активные спутники), GSV (Global Satellites in View — спутники в зоне видимости) и RMC (Recommended Minimum speCific data — служебные данные о них). В рамках стандарта NMEA оговорены также дополнительные форматы — GLL и VTG, имеющие ограниченное применение, например VTG используется только при работе с оборудованием фирмы Garmin — лидера на рынке классических (не компьютерных) GPS-приемников. Кроме того, способами обмена могут быть двоичный SiRF-код, управляющие коды формата Trimble Standard

Interface Protocol (TSIP) и некоторые другие — они указываются в описаниях после NMEA.

Итак, наш GPS-приемник, получив со спутников точное время отправки сигнала (на них установлены атомные часы), по фактической задержке прохождения сигналов вычисляет физические расстояния до спутников (скорость распространения радиоволн известна). Посредством метода триангуляции, реализованного в прошивке приемника, определяется точное его положение (широта и долгота) на поверхности Земли минимум по трем спутникам (рис. 5.23).

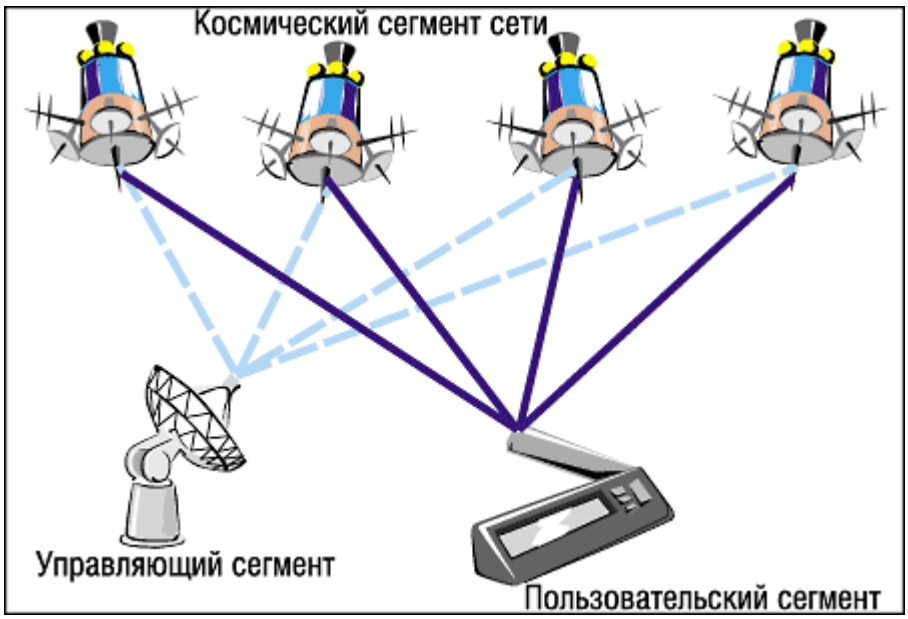


Рисунок 5.23 — Схема организации системы GPS

А запеленговав четыре или более спутников, приемник может также определить и высоту абонента над уровнем моря (altitude). Время, за которое способен это сделать приемник, никогда не указывается без привязки к условиям, при которых происходил прием сигнала. Обычно даются четыре параметра: время TTFF при холодном, теплом и горячем старте, и отдельно — reacqusition time, т.е. время, за которое приемник может восстановить связь со спутниками после временного выхода из зоны покрытия.

А дальше приемник может вычислять максимальную и среднюю скорость движения, поработает компасом, покажет направление на цель и примерное время, через которое вы там окажетесь, двигаясь с текущей скоростью, расстояние до пункта назначения, время ожидаемого восхода и заката солнца (весьма полезно туристам при планировании привалов и дневок) и многое другое — это уже зависит от встроенного софта. Данные постоянно обновляются — обычно раз в секунду.

Точность (Accuracy) определения координат, бесспорно, является важнейшим параметром GPS. Как правило, в характеристиках изделия она указывается для горизонтальных координат и довольно редко — для

высоты (alteration). Сравнительно недавно точность показаний коммерческих GPS искусственно уменьшалась посредством введения так называемых Selective Availability (SA), когда помехи намеренно вводились в показания спутников для занижения точности определения координат бытовыми (не военными) устройствами. Это давало погрешность в пределах 100 м (что автоматически делало применение систем на улицах городов довольно проблемным), хотя базовые возможности GPS-системы позволяют вычислить пределы ее точности от 5 до 25 м и 10 см/с — при нахождении составляющих вектора скорости (Velocity). Продолжалось так довольно долго, однако перспективы полноценного коммерческого использования GPS оказались столь многообещающими, что в мае 2000 г. специальным решением президента США были сняты все SA-ограничения по точности, так что теперь ее обычно считают равной 15 м (рис. 5.24). Для дальнейшего же ее повышения необходимо введение дополнительных поправок и усовершенствованных алгоритмов. Нужно ли это? Для гуляющего по городу туриста, пожалуй, нет. А вот для фаната-рыбака, отмечающего на карте места поклевок, — в самый раз.

Рисунок 5.24 — Точность GPS в различных режимах

Итак, при внедрении Differential GPS (DGPS) используются координаты от двух GPS-приемников, одного — рабочего, второго — эталонного (стационарно установленного в месте, координаты которого измерены с высокой точностью), и оба пеленгуют GPS-спутники в один и тот же промежуток времени, что дает возможность вычислить поправку и довести точность до 3—5 метров. Так действует служба береговой охраны США, содержащая сеть башен, принимающих сигналы GPS и передающих скорректированные сигналы посредством маячных (beacon) передатчиков. Дело в том, что стандартным GGA-форматом сигнала со спутника предусмотрены особые поля для использования DGPS, они после

ретрансляции заполнены дополнительными данными, а состояние DGPS отмечается специальным значением триггера «Position Fix Indicator» в GGA-сообщении. Эти поправки может принять любой желающий на побережье и прилегающей к нему территории страны. Но необходимо иметь соответствующий приемник с забавным названием BoB (Beacon-on-a-Belt) в дополнение к стандартному. А вот система трансляции поправок через FM-радиостанции является платной услугой, и только ее подписчикам выдается FM-приемник размером с пейджер, тоже работающий в связке с GPS-приемником.

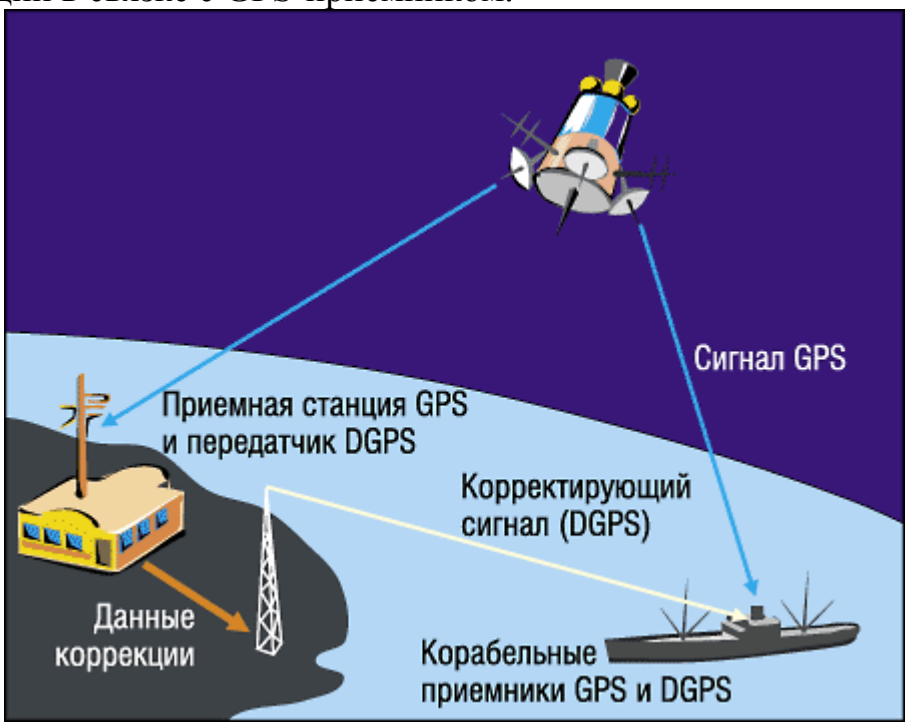


Рисунок 5.25 — Принцип работы системы DGPS

В дополнительном оборудовании не нуждается другая система внесения поправок, называемая WAAS (Wide Area Augmentation System), которая дает точность даже меньше 3 метров. Ее разработчиком является Федеральное управление авиации США. Собственно WAAS охватывает только США и включает 25 наземных станций, отслеживающих сигналы со спутников, а также две мастер-станции (по одной на Западное и Восточное побережье США), которые на основе данных от всех остальных вырабатывают поправки (рис. 5.25). Корректирующая информация постоянно транслируется через один из 2 геостационарных (т.е. висящих практически неподвижно над экватором) спутников в стандартном GGA-формате (в нем предусмотрены специальные поля для этого) и воспринимается теми приемниками, которые разрабатывались как WAAS-enabled (т.е. их владельцам дополнительное оборудование не требуется).

Можно возразить, что эта система, равно как и аналогичный японский проект MSAS (Multi-Functional Satellite Augmentation System), имеет для нас чисто познавательное значение. А созданная с той же целью

на закате СССР российская система ГЛОНАСС используется исключительно в военных целях (в 2000 г. Министерство обороны России, чтобы не отстать от США, приняло решение о подготовке системы к гражданскому использованию, но, похоже, воз и ныне там). Скорее всего, без дополнительных денежных вливаний ГЛОНАСС не сможет удовлетворить запросы коммерческих пользователей, ведь ее восемь спутников обеспечивают точность от 50 до 70 м, так что расчет поправок жизненно необходим.

К счастью, ситуация постепенно улучшалась — стартовал европейский проект EGNOS, полноценное использование которого предполагается начать в 2004 году. Любопытно, что в нем принимаются сигналы с обеих разновидностей спутников — и GPS, и ГЛОНАСС. EGNOS является совместным проектом Европейского космического агентства (ESA), Еврокомиссии (EC) и Eurocontrol — организации, отвечающей за аэронавигацию в Европе, и предшественником Galileo — первой действительно глобальной системы спутниковой навигации, которая разрабатывается в Европе.

Транслируются поправки EGNOS через три геостационарных спутника. Два из них относятся к семейству Inmarsat и висят над Атлантическим и Индийским океанами, а спутник ESA Artemis — над Африкой. Помимо них, в систему входят 4 Master Control Centres (MCC), которые управляют этими спутниками и вырабатывают поправки, 34 Ranging and Integrity Monitoring Stations (RIMS) — данные с них используются в мастер-центрах для уточнения поправок, а также передающие эту информацию на спутники станции Navigation Land Earth Stations (NLES), их на первой стадии проекта открыто семь. Сейчас EGNOS позволяет пользователям на территории Европы получать точность в пределах 5 м.

Архитектура SiRFstar

За реализацию всех рассмотренных выше операций отвечает чипсет GPS-приемника и его ПЗУ, причем абсолютное большинство реализаций GPS для карманных ПК построено на чипсетах американской компании SiRF Technology — ее костяк составляют экс-сотрудники таких авторитетов в области геодезии, как Trimble Navigation, с богатым научным прошлым. Уже в GPS на чипсете SiRF — starI (с 1997 г.) были реализованы запатентованные алгоритмы SnapLock (захват спутника за десятую долю секунды вместо 2—3 секунд после выхода из перекрытой области — это очень важно для автомобилей, проходящих под мостами и через туннели), SingleSat (прогнозирование местоположения приемника, когда виден хотя бы один спутник; обычно если их меньше трех, расчеты не производятся), Dual Multipath (игнорирование паразитных сигналов, образованных отражением основного сигнала от высотных зданий, скал и других поверхностей, которое замедляет его попадание на приемник и вносит распределенную случайным образом довольно значительную

погрешность), а также FoliageLoc — алгоритм, позволяющий принимать очень ослабленный сигнал, — как правило, при нахождении в лесу. Реализующие последний алгоритм GPS-приемники принимают сигнал на 20 дБ ниже типичного уровня в –160 дБ, который может быть распознан обычным приемником. А ведь еще совсем недавно любого могла удивить сама возможность при попадании в проблемную ситуацию мгновенно передать координаты в службу помощи с устройства, весившего «всего» около фунта!

Хотя до сих пор на рынке продается немало устройств на чипсете SiRFstarI предыдущего поколения, компания активно продвигает на мировой рынок архитектуру GPS-чипсетов SiRFstarII, которая дополняет ядро SiRFstarI рядом полезных возможностей. В новую же архитектуру добавлены механизмы отслеживания спутников без задействования процессорного модуля (кстати, внутри приемника установлен процессор 50 MHz ARM-архитектуры и память on-chip), реализована поддержка WAAS и DGPS, а также продвинутый режим энергосбережения (Advanced Trickle Power Mode), когда электроника приемника «засыпает» на 800 мс из каждой секунды, а за оставшиеся 200 мс выполняется собственно поиск спутников, прием данных и вычисления.

Семейство GPS-чипсетов SiRFstarII сейчас составляют SiRFstarIIe (основной чипсет, аппаратно реализованы WAAS, EGNOS, DGPS, есть свой ARM-процессор ARM7TDMI и 1 MB EDO DRAM on-chip), SiRFstarIIe/LP (чипсет с пониженным энергопотреблением, 60 мА на полной мощности и 20 мА в режиме TricklePower), а также бюджетный SiRFstarIIt, изюминка которого состоит в отсутствии встроенного процессора в приемнике и перекладывании всех вычислений на CPU того компьютера, на котором работает GPS-приемник. За их выполнение и обмен данными отвечает приложение SiRFNav с реализациями под различные платформы.

5.8 Транковая, сотовая, спутниковая связь

В современном деловом мире всё больше внимание уделяется средствам мобильной связи: пейджерам, аппаратам сотовой и спутниковой связи, персональным коммуникаторам и тому подобным устройствам. В самом деле, для того чтобы быть конкурентоспособными, современным компаниям требуется постоянно поддерживать связь со своими клиентами и, что не менее важно, между сотрудниками своей организации. В последнее время некоторые операторы мобильной связи предлагают так называемые «корпоративные» тарифы, которые как раз предназначены для создания «виртуальной телефонной сети» для сотрудников компании. Однако подобные программы — это не самое дешёвое решение проблемы коммуникации, но, к счастью, не единственно возможное.

Для компании, решившей «соединить» своих мобильных сотрудников, существует альтернативное решение — использование

транковой связи. Возможно, многие видят словосочетание «транковая связь» впервые. Действительно, системам транковой связи сейчас уделяется меньшее внимание, чем даже пейджинговым системам. В какой-то мере это связано с тем, что системы транковой связи предназначены прежде всего для использования крупными организациями, а не массовыми пользователями. Несмотря на это, данная технология имеет свои достоинства и заслуживает рассмотрения.

Итак, что же скрывается за термином «транковая система»? Как ни парадоксально, но мы пользуемся ею каждый день, даже не задумываясь об этом. Именно на принципе транкинга основано действие современных АТС. Давайте проследим, что же происходит, когда вы пытаетесь позвонить с домашнего телефона, допустим, своему другу. Вы снимаете трубку, дожидаетесь сигнала «линия свободна», затем набираете номер и ждёте ответа. Все остальные действия выполняет АТС: она выбирает один из свободных каналов связи и коммутирует (связывает) ваш телефонный аппарат с телефонным аппаратом друга. По окончании разговора линия, которая была использована, освобождается и становится доступной для использования уже другими людьми. Как вы догадываетесь, число линий связи ограничено и заведомо меньше необходимого для соединения всех телефонных аппаратов в городе. Таким образом, АТС контролирует распределение ограниченного числа линий между большим количеством абонентов. Предполагается, что ситуация, когда все абоненты вдруг решат одновременно связаться друг с другом, не возникнет. Следовательно, необходимо правильно рассчитать минимально необходимое число каналов связи, чтобы в процессе работы не возникали проблемы, связанные с их нехваткой. Этот вопрос эффективно решается с использованием математической теории систем массового обслуживания.

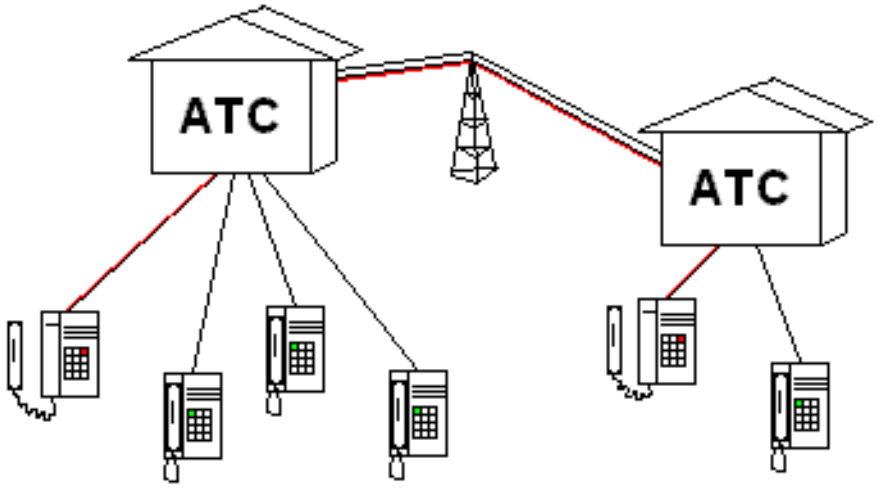


Рисунок 5.26 — Транковая связь

Транковые радиосистемы — это системы подвижной радиосвязи, которые основаны на тех же принципах, что и обычные телефонные сети.

Иными словами, в системе транковой радиосвязи имеется ограниченное число радиоканалов (как правило, от двух до двадцати), которые по мере надобности выделяются центральным контроллером для ведения переговоров.

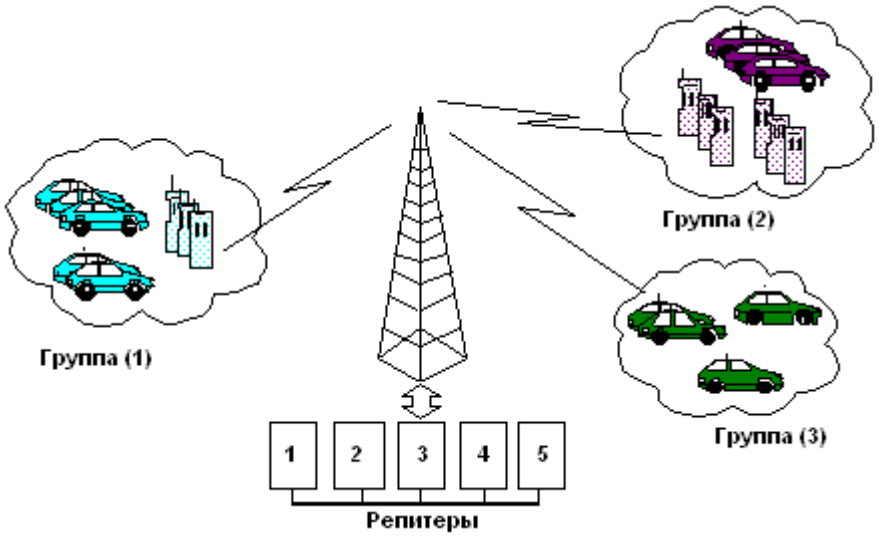


Рисунок 5.27 — Транковая радиосвязь

В обычных системах радиосвязи пользователю приходится вручную перенастраиваться на свободный радиоканал, в системах транковой связи эту работу берёт на себя центральный контроллер, который сам выделяет двум радиостанциям свободный канал. Таким образом, пользователю нужно просто набрать номер вызываемого абонента, остальное система сделает сама. Транковой системе можно дать следующее определение: автоматическое и динамическое распределение небольшого числа каналов среди большого числа радиопользователей.

Структурная схема базовой станции для системы транковой радиосвязи

На рисунке 5.28 приведена структурная схема базовой станции в случае использования одного канала.

Рисунок 5.28 — Структурная схема базовой станции

Репитер состоит из ретранслятора, предназначенного для приёма сигналов абонентских радиостанций, его усиления и передачи, и контроллера транкового канала, который выполняет управляющие функции.

Дуплексный фильтр — устройство, позволяющее использовать одну антенну для приёма и передачи. В принципе, ничто не мешает использовать для приёма и передачи две разные антенны, но в этом случае может возникнуть ситуация, когда в некоторых местах будет возможен приём, но невозможна передача либо наоборот. Кроме того, излучаемая передатчиком мощность влияет на приёмник, поэтому при наличии двух антенн их нужно устанавливать на достаточном расстоянии друг от друга.

Источник питания предназначен для репитера. Как правило, он допускает возможность перехода на аккумуляторную батарею при отключении питания.

Рассмотренная схема является достаточно простой и эффективной, однако в реальных условиях одного транкового канала оказывается недостаточно. Поэтому применяют системы, содержащие два и более каналов. На рисунке 5.29 показана схема системы, содержащей четыре независимых канала. Как видно, основное отличие от предыдущего варианта заключается в антенно-фидерном тракте, где появляются ещё два устройства: приёмная распредпанель и комбайнер.

Приёмная распредпанель обеспечивает одинаковый входной сигнал для каждого репитера в системе, как если бы репитер был подключен напрямую к антенне.

Комбайнер — это устройство, позволяющее комбинировать выходы определённого количества передатчиков без интерференции друг с другом.

Также отдельно вынесен источник бесперебойного питания, который просто обязан присутствовать в системе, ибо отсутствие связи в чрезвычайных обстоятельствах может привести к непредсказуемым последствиям.

Рассмотренную систему легко расширить, то есть в случае правильного проектирования число каналов можно увеличить достаточно безболезненно.

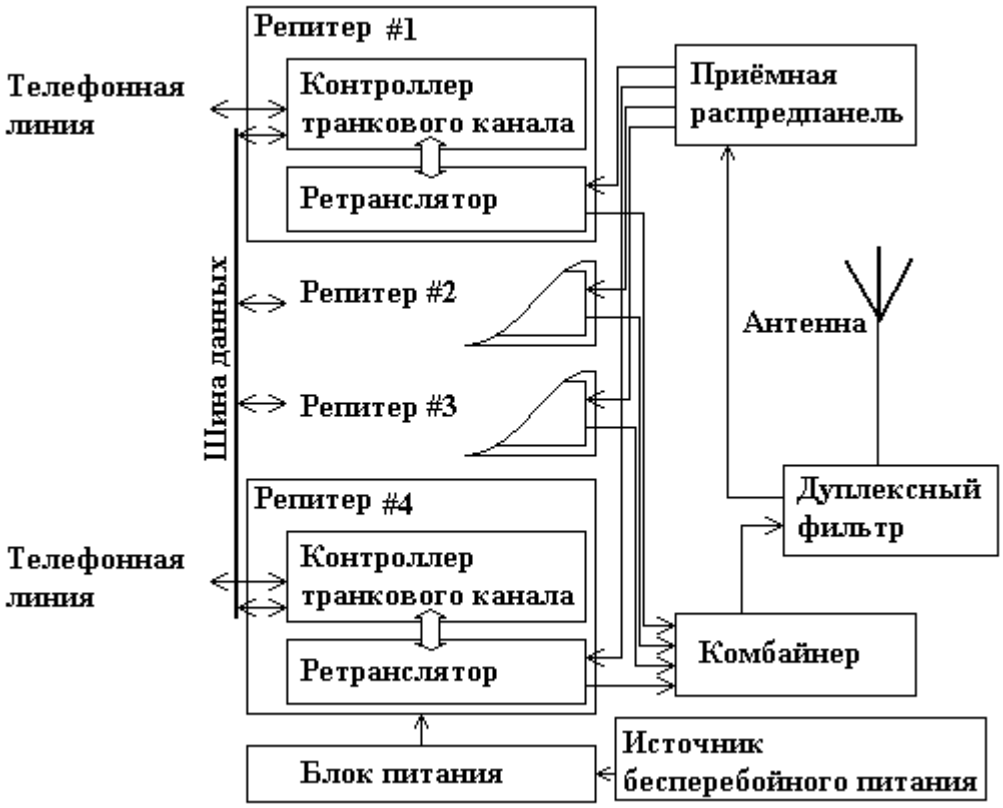


Рисунок 5.29 — Схема станции с четырьмя каналами

Сотовая связь — один из видов мобильной радиосвязи, в основе которого лежит сотовая сеть. Ключевая особенность заключается в том, что общая зона покрытия делится на ячейки (соты), определяющиеся зонами покрытия отдельных базовых станций (БС). Соты частично перекрываются и вместе образуют сеть. На идеальной (ровной и без застройки) поверхности зона покрытия одной БС представляет собой круг, поэтому составленная из них сеть имеет вид сот с шестиугольными ячейками (сотами).

Примечательно, что в английском варианте связь называется «ячеистой» или «клеточной» (cellular), что не учитывает шестиугольности сот.

Базовая станция (применительно к сотовой связи) — комплекс радиопередающей аппаратуры (ретрансляторы, приёмо-передатчики), осуществляющий связь с конечным абонентским устройством — сотовым телефоном. Одна базовая станция стандарта GSM обычно способна поддерживать до 12-ти передатчиков, а каждый передатчик способен одновременно поддерживать связь с 8-ю общающимися абонентами. Зона покрытия от антенн базовой станции — образует соту, или группу сот (обычно 3 соты в 900 MHz и 3 соты в 1800 MHz). Базовые станции обычно посредством радиорелейной связи (возможны и другие способы, например медный кабель, оптика и т.д.) соединены с коммутатором сотовой сети (через контроллер) или между собой. Базовые станции размещаются в

основном в административных зданиях либо обособленно. Антенны базовых станций можно часто видеть на крышах и фасадах — это узкие прямоугольные конструкции бежевого либо белого цвета, от 20 до 3200 см в длину и 20 см в ширину. Базовые станции также применяются в сотовом телевидении, сетях WiFi, WiMAx и других технологиях.

Сеть составляют разнесённые в пространстве приёмопередатчики, работающие в одном и том же частотном диапазоне, и коммутирующее оборудование, позволяющее определять текущее местоположение подвижных абонентов и обеспечивать непрерывность связи при перемещении абонента из зоны действия одного приёмопередатчика в зону действия другого.

Основные составляющие сотовой сети — это сотовые телефоны и базовые станции. Базовые станции обычно располагают на крышах зданий и вышках. Будучи включённым, сотовый телефон прослушивает эфир, находя сигнал базовой станции. После этого телефон посылает станции свой уникальный идентификационный код. Телефон и станция поддерживают постоянный радиоконтакт, периодически обмениваясь пакетами. Связь телефона со станцией может идти по аналоговому протоколу (AMPS, NAMPS, NMT-450) или по цифровому (DAMPS, CDMA, GSM, UMTS). Если телефон выходит из поля действия базовой станции, он налаживает связь с другой (англ. handover).

Сотовые сети могут состоять из базовых станций разного стандарта, что позволяет оптимизировать работу сети и улучшить её покрытие.

Сотовые сети разных операторов соединены друг с другом, а также со стационарной телефонной сетью. Это позволяет абонентам одного оператора делать звонки абонентам другого оператора, с мобильных телефонов на стационарные и со стационарных на мобильные.

Операторы разных стран могут заключать договоры роуминга. Благодаря таким договорам абонент, находясь за границей, может совершать и принимать звонки через сеть другого оператора (правда, по повышенным тарифам).

Спутниковая связь

Неизвестно, когда точно была придумана система радиосвязи, использующая объекты, движущиеся над Землей, но вполне возможно, что произошло это в 1945 году.

Именно тогда известный писатель-фантаст Артур Кларк придумал геостационарный спутник связи в виде гигантской платформы на орбите Земли, перемещающейся по ней синхронно с вращением планеты (на платформе размещалось оборудование, обеспечивающее передачу телефонных и телевизионных сигналов на всю планету). Научное сообщество высоко оценило столь гениальное предвидение и дало геостационарной орбите второе имя — Clarke Belt («Пояс Кларка»).

Первой успешной попыткой создать систему мобильной спутниковой связи стала международная система Inmarsat. Сама идея

такой системы (поначалу она предназначалась только для обеспечения связи с морскими судами) была выдвинута еще в 1966 году. Однако ее реализация потребовала такой работы, что коммерческая эксплуатация системы была начата только в 1982 году. При этом стоимость, размеры и вес первых абонентских комплектов оборудования были столь велики, что они могли размещаться только на очень крупных морских судах.

В начале 90-х наступил новый этап в развитии систем мобильной спутниковой связи. А в качестве первой реальной системы персональной спутниковой связи можно назвать разработку совместной американо-канадской компании Orbcomm. Предельно снизив размеры, вес и стоимость спутников и их запуска, создатели системы сделали мобильную связь доступной для многих, и в 1995 году Orbcomm начала предоставлять услугу обмена данными и электронной почтой.

Принцип действия

Любая система спутниковой связи состоит из 3-х основных элементов: космического сегмента (спутники, вращающиеся на орбитах вокруг Земли), наземных станций (служат для управления полетом спутников и передачи сигналов со спутников в наземные сети связи и обратно) и абонентских терминалов. Порядок работы спутниковых систем сходен с обычными сотовыми, где функции центральных коммутаторов выполняют наземные станции, а роль базовых станций играют спутники.

По типу используемых орбит спутниковые системы связи делятся на два класса: уже упоминавшиеся системы со спутниками на геостационарной орбите (расстояние до поверхности Земли — около 36 000 км) и негеостационарные.

Достоинством геостационарной орбиты является то, что угловая скорость вращения спутников на ней точно совпадает со скоростью вращения Земли и каждый спутник оказывается как бы «висящим» над заданной точкой на экваторе. При этом один спутник может охватывать связью примерно треть всей поверхности планеты, за исключением полюсов. Основные недостатки геостационарных систем обусловлены большой удаленностью спутников от Земли и проявляются в сильном ослаблении принимаемых сигналов и в довольно большой их задержке при распространении — 0,24 с, что становится заметным даже при обычном телефонном разговоре.

Негеостационарные спутниковые системы обычно используют круговые орбиты двух типов: средневысотные (MEO, высота — 5000—15 000 км) и низкоорбитальные (LEO, высота — 500—2000 км). При этом один MEO-спутник способен охватить связью до 25 % поверхности Земли, а для построения глобальной системы связи требуется около 10 ИСЗ. Зона покрытия LEO-спутника значительно меньше — 3—7 %, и глобальная система должна содержать уже порядка 50 ИСЗ.

Такой выбор орбит не прихоть разработчиков — он продиктован расположением в околоземном пространстве так называемых зон Ван

Аллена — поясов заряженных частиц, удерживаемых магнитным полем Земли. Оборудование ИСЗ, расположенных вне указанных орбит, будет подвергаться сильной «бомбардировке» частицами и быстро выйдет из строя.

Inmarsat

Как уже упоминалось, Inmarsat использует геостационарные ИСЗ и обеспечивает связь на широтах до 70°. За время своего существования система сильно изменилась. Так, на смену первой системе, Inmarsat-A (телефонная, телеграфная, телексная и факсимильная связь), в 1993 году пришла Inmarsat-B, основанная на цифровых способах передачи информации и совместимая с сетями ISDN.

Третья система, Inmarsat-C, находится в коммерческой эксплуатации с 1991 года, использует небольшие абонентские терминалы с ненаправленными антеннами и обеспечивает двухстороннюю низкоскоростную передачу данных и телексной информации.

Inmarsat-Aero, работающая с 1990 года, применяется в гражданской авиации для передачи данных и телефонной связи, в том числе и через пассажирские телефоны в самолетах.

Inmarsat-M, введенная в эксплуатацию в 1993 году, стала первой в мире системой, предоставляющей услуги персонального спутникового телефона. Система базируется на переносных станциях (типа чемоданчиков «дипломат» весом 10—15 кг), обеспечивающих цифровую телефонную и факсимильную связь и низкоскоростную передачу данных.

В 1995 году в эксплуатацию была введена система односторонней пейджинговой связи Inmarsat-D, а впоследствии и Inmarsat-D+ (пейджер с ответом).

Система Inmarsat-mini-M, пришедшая в 1997 году на смену Inmarsat-M, использует малогабаритные терминалы размером с ноутбук и весом около 2-х кг.

Последняя разработка, Inmarsat-M4, введенная в эксплуатацию в конце 1999 года, обеспечивает голосовую связь, факс, передачу данных на скорости до 64 кбит/с, электронную почту, доступ в интернет и др.

Следует отметить, что за все время своего существования Inmarsat всегда оставалась стабильно действующей системой мобильной связи. В настоящее время в ней работают более 150 тыс. абонентских станций различных типов, из которых около 10 тыс. — российские, что вполне закономерно: Россия участвовала в создании Inmarsat изначально.

Omnitracs и Euteltracs

Спутниковая система мобильной связи Omnitracs, также на основе геостационарных ИСЗ, работает на североамериканском континенте с 1988 года, а ее аналог, Euteltracs, — с 1991 года в Европе и России. Системы обеспечивают связь с подвижными объектами (преимущественно с большегрузными автомобилями) путем передачи буквенно-цифровых

сообщений от объекта в центр и из центра на выбранный объект или группу объектов.

Iridium

Название этой системы глобальной подвижной персональной спутниковой связи произошло от химического элемента, атом которого содержит 77 электронов. Именно такое количество низкоорбитальных спутников должна была содержать система, создававшаяся компанией Motorola начиная с 1989 года для предоставления услуг телефонной связи на всей территории земного шара. В процессе разработки число ИСЗ удалось сократить до 66, расположенных на 6 орбитах высотой 780 км над поверхностью Земли.

Iridium обеспечивает услуги радиотелефонной связи, передачу данных, факсов, персонального вызова (пейджинг) и определение местоположения. Система начала работать в сентябре 1998 года.

Globalstar

Еще одна система на основе низкоорбитальных ИСЗ, разработанная рядом компаний во главе с американской промышленной группой Loral и предназначенная для предоставления услуг связи на территории земного шара между 70° с. ш. и 70° ю. ш. Космический сегмент системы представляет собой группу из 48 спутников, размещенных на 8 круговых орбитах на высоте 1414 км над поверхностью Земли. В системе предусмотрены услуги телефонной и факсимильной связи, передачи данных, определение местоположения и др. Абонентское оборудование состоит из портативных, мобильных и стационарных терминальных устройств, включая специализированные таксофоны. Портативные устройства выпускаются в нескольких модификациях, способных работать как в системе Globalstar, так и в сетях наземной сотовой связи стандартов GSM, AMPS, CDMA. В России система Globalstar предоставляет услуги связи (через свою дочернюю компанию «ГлобалТел») с мая 2000 года.

Проблемы и перспективы

Общее количество спутниковых систем связи, заявлявшихся к реализации в последнее десятилетие, уже давно исчисляется десятками: ECCO, Ellipso (приблизительный аналог отечественной системы «Орбита» со спутниками «Молния»), Archimedes, Odyssey, Celestri, Skybridge, SECOMS, Starsys, Voicespan, Astrolink, Cyberstar, KaStar, Spaseway, Teledesic и др. Существуют и проекты развертывания отечественных систем, среди которых наиболее известны «Сигнал», «Гонец», «Марафон», «Полярная звезда». Однако большинство пока так и остается на различных стадиях «бумажного» проектирования. И по вполне объективным причинам.

Все геостационарные системы первых поколений не позволяли сделать абонентский аппарат малогабаритным (наилучший результат — формат ноутбука). Это позволяют достигнуть системы на низкоорбитальных ИСЗ, но здесь возникает другая проблема. Технически

такие системы сегодня вполне реализуемы, они обеспечивают устойчивую связь с помощью малогабаритных абонентских трубок. Но стоимость создания и эксплуатации системы, содержащей десятки спутников, весьма высока, что не позволяет установить низкие цены на обслуживание. Ну а здесь возникает порочный круг: в развитых странах (где такие цены могли бы быть допустимы) спутниковая связь не нужна, так как практически вся территория уже охвачена сотовыми сетями. Там же, где у спутниковой связи фактически нет конкурентов (в «заповедных уголках» Африки, Азии и т.п.), отсутствуют платежеспособные потребители.

Впрочем, оценивая перспективы спутниковых систем мобильной связи, нельзя исключить и то, что они могут получить широкое распространение. Пример тому — система Thuraya, также использующая геостационарный ИСЗ. Она была введена в эксплуатацию только в конце 2000 года. Реализованные в ней последние научные достижения (особо мощные передатчики и антенная решетка диаметром более 12 м с цифровым управлением, обеспечивающая одновременное формирование до 300 остронаправленных лучей) позволили предоставлять услуги мобильной связи с помощью малогабаритных радиотелефонов, подобных обычным сотовым, весом всего около 200 г. Данная система, принадлежащая Thuraya Satellite Telecommunications Company из ОАЭ и разработанная компанией Boeing Satellite Systems, в настоящее время способна предоставлять связь в 99 странах Европы, Азии и Африки, включая и Россию.

И у этого проекта есть последователи. По пятам за ним следует разработка японской компании Mitsubishi Electric, предусматривающая вывод в космос 3-х спутников связи с антеннами диаметром уже 45 м. Планируется, что данная система, способная одновременно предоставлять 100 тыс. телефонных разговоров, сможет обслуживать до 5 млн абонентов.

Таким образом, вполне возможно, что именно спутниковые системы смогут справиться с этой нелегкой задачей — обеспечить все население нашей планеты персональной связью.

Системы с ШПС

Весьма важными достижениями инженеров XX столетия явились изобретение широкополосных сигналов (ШПС) и создание на их основе новых систем радиосвязи и радиолокации. Широкое использование этих сигналов для создания радиосистем массового применения началось в последние 10—15 лет. Однако к разработке идей их использования для повышения помехоустойчивости приема приступили еще в 40-х годах. В обычных видах модуляции информационный поток изменяет амплитуду, фазу или частоту гармонического колебания — несущей частоты. При этом ширина спектра излучаемого в эфир сигнала соизмерима с шириной спектра модулирующего сигнала. Однако переносчиком информации может быть не только гармонический, но и сложный широкополосный сигнал. Такой сигнал может быть сформирован разным образом: несущая

может быть модулирована вспомогательной кодовой последовательностью или вспомогательным аналоговым сигналом по фазе, частоте или амплитуде.

Первые работы, связанные с ШПС, были направлены на разработку методов борьбы с мощными радиопомехами, мешающими приему радиолокационных сигналов. В последующие годы были разработаны и нашли применение в системах радиосвязи три основных способа формирования ШПС.

Первые идеи построения ШПС были связаны с ЧМ несущей частоты вспомогательным сигналом, структура которого должна была быть известна на приеме. Одно из первых изобретений, позволяющих выделить ШПС на фоне шума и мощной помехи, представляющее, по сути, полосовой коррелятор, было сделано еще в 1942 году. В этом же году было сделано еще одно пионерское изобретение — способ формирования ШПС методом скачкообразного изменения частоты несущего колебания за время передачи одного информационного символа (FH-SS — Frequency Hopping Spread Spectrum). Честь этого изобретения, долгое время остававшегося секретным, принадлежит известной американской актрисе Хэди Ламар, признанной в 1940 году на конкурсе красоты самой красивой женщиной мира, и ее мужу — композитору Георгу Атсейлу. Изобретенный ими принцип формирования широкополосных сигналов сегодня находит применение в ряде систем связи. В отечественной литературе такие сигналы называют сигналами с частотно-временной матрицей (ЧВМ).

Другим методом формирования ШПС, дуальным к методу ЧВМ и разработанным в 1946 году, является метод скачкообразного изменения относительного временного положения коротких кодовых импульсов за время передачи одного информационного символа (TH-SS — Time Hopping Spread Spectrum).

Один из наиболее широко применяемых сегодня на практике способов формирования ШПС, который называется методом прямого расширения спектра путем непосредственной фазовой модуляции несущей определенной кодовой последовательностью (DS-SS — Direct Sequence Spread Spectrum), был изобретен американскими специалистами Дж. Г. Грином и М. Г. Никольсоном в 1957 году. Ими был предложен метод построения бинарной кодовой последовательности с хорошими корреляционными свойствами. Позже было выполнено значительное число работ, посвященных синтезу подобных псевдослучайных последовательностей с помощью регистров сдвига. К пионерским теоретическим работам в этом направлении относятся исследования С. Голомба (1955 г.) и Н. Цирлера (1959 г.).

Во всех указанных случаях возможно создание в общей полосе частот больших ансамблей сигналов, которые отличаются либо законом чередования фазы несущей частоты для сигналов DS-SS, либо законом изменения значения несущей частоты для сигналов FH-SS, либо определенной временной расстановкой коротких импульсов для сигналов

TH-SS. Аналогично тому, как в системах с ЧУ и ВУ сигналы разных каналов могут быть разделены по частоте либо по временному положению, которое они занимают в общей временной последовательности, возможно разделение и разных сигналов ансамбля ШПС по индивидуальной кодовой структуре каждого из этих сигналов. Таким образом, ШПС могут использоваться в качестве переносчиков информации подобно тому, как используются гармонические колебания в обычных системах связи. При этом аналогом АМ является передача одного из ШПС, принадлежащего к определенному ансамблю сигналов, с определенной амплитудой, аналогом ФМн является манипуляция фазы ШПС, а аналогом обычной ЧМн является передача одного из двух возможных сигналов ШПС по линии связи.

Сигналы, не перекрывающиеся по спектру или времени, являются полностью ортогональными. Их применение в качестве переносчиков информации в многоканальных системах позволяет полностью разделить соответствующие каналы связи. В отличие от таких сигналов разные ШПС, принадлежащие к одному ансамблю, не являются полностью ортогональными, и поэтому при их разделении возникают дополнительные шумы. Однако их замечательное свойство состоит в том, что в системах связи, использующих ШПС, которые называются системами с кодовым разделением каналов (CDMA — Code Division Multiple Access), или асинхронно-адресными системами, поступающие на вход помехи подавляются в B = WIF раз, где B — база ШПС, W — полоса частот, занимаемая ШПС в канале связи, F — полоса частот информационного сигнала. В широкополосных системах связи B = 100—10 000, и в них обеспечивается весьма значительное подавление помех, действующих в той же самой полосе частот, в которой работает данная система. Данное свойство ШПС является уникальным и позволяет многократно использовать один и тот же частотный канал для связи разных абонентов на ограниченной территории. В традиционных системах связи для исключения возможности возникновения помех между зонами, в которых используется один и тот же частотный канал, должен быть обеспечен весьма значительный территориальный разнос. Таким образом, в системах CDMA достигается весьма высокая эффективность использования радиочастотного спектра. Кроме того, ШПС позволяют путем специальной обработки принимаемых сигналов эффективно бороться с замираниями сигналов в многолучевых каналах связи, разделяя отдельные лучи и осуществляя их когерентное сложение.

Первой системой, в которой начали применяться ШПС, явилась созданная в 1946 году система гиперболической навигации «Лоран», в которой около десятка пар станций работали в общем частотном канале независимо друг от друга, используя сигналы TH-SS. В 1952 году на этом же принципе было создано связное оборудование для передачи сигналов телефонии.

В 1958 году была создана первая система коротковолновой связи «Рейк» для работы в многолучевом канале, в которой ШПС применялись

для разделения отдельных лучей и устранения замираний, вызванных их интерференцией.

Первые системы, использующие сигналы с FH-SS, появились в начале 60-х годов. В 1963 году была создана наземная система связи RACEP (Random Access and Correlation for Extended Perfomance), в которой для передачи полезных сообщений применялась ФИМ. Система занимала полосу частот 4 МГц и работала в диапазоне 140 МГц. Она позволяла осуществлять передачу сигналов телефонии и цифровой информации и давала возможность организации на одной территории сети связи с емкостью 700 абонентов. Наибольшее число одновременно работающих абонентов составляло 35. Несколько позже была создана аналогичная система RADAS — Random Access Adress System, в которой для передачи информации применялась ДМ.

С 1963 году на основе ШПС начинают создаваться спутниковые системы связи со свободным доступом к общему каналу связи, тропосферные радиорелейные системы связи с разделением отдельных лучей. Исследования эффективности применения ШПС в сравнении с другими методами модуляции в системах связи различных назначений начались с 1965 года.

В 90-х годах системы с ШПС начинают внедряться в системы сотовой подвижной связи. Подобные системы будут применяться в сотовых системах подвижной связи, широкое внедрение которых начнется в XXI веке. Исследования вопросов эффективности использования РЧС в таких системах и разработка методов их частотного планирования были начаты российским ученым Л.Е. Варакиным.

6 УСТРОЙСТВА ВВОДА, ОТОБРАЖЕНИЯ ИНФОРМАЦИИ

Человек взаимодействует с информационными системами главным образом через устройства ввода-вывода (input-output devices). Прогресс в области информационных технологий достигается не только благодаря возрастающей скорости процессоров и емкости запоминающих устройств, но также за счет совершенствования устройств ввода и вывода данных. Устройства ввода-вывода называются также периферийными устройствами (peripheral devices).

6.1 Устройства ввода данных

Клавиатура

Клавиатура (keyboard) — традиционное устройство ввода данных в компьютер. Клавиатурами оснащены как персональные компьютеры, так и терминалы мэйнфреймов. Клавиатура современного компьютера содержит обычно 101 или 102 клавиши, разделенные на 4 блока:

– алфавитно-цифровой блок — содержит клавиши латинского и национального алфавитов, а также клавиши цифр и специальных

символов;

- блок управляющих клавиш;
- блок расширенной цифровой клавиатуры;
- блок навигации.

Компьютерная мышь

Мышь (mouse) была разработана довольно давно (в 60-х годах), но стала широко использоваться только с приходом в мир персональных компьютеров графического пользовательского интерфейса.

Прямой привод

Изначальная конструкция датчика перемещения мыши, изобретённой Дугласом Энгельбартом в Стэнфордском исследовательском институте в 1963 году, состояла из двух перпендикулярных колес, выступающих из корпуса устройства. При перемещении мыши колеса крутились каждое в своем измерении.

Такая конструкция имела много недостатков и довольно скоро была заменена на мышь с шаровым приводом.

Шаровой привод

В шаровом приводе движение мыши передается на выступающий из корпуса гуммированный стальной шарик (его вес и резиновое покрытие обеспечивают хорошее сцепление с рабочей поверхностью). Два прижатых к шарику ролика снимают его движения по каждому из измерений и передают их на датчики, преобразующие эти движения в электрические сигналы.

Основной недостаток шарового привода — загрязнение шарика и снимающих роликов, приводящее к заеданию мыши и необходимости в периодической её чистке. Несмотря на недостатки, шаровой привод долгое время доминировал, успешно конкурируя с альтернативными схемами датчиков. В настоящее время шаровые мыши почти полностью вытеснены оптическими мышами второго поколения.

Контактные датчики

Контактный датчик представляет из себя текстолитовый диск с лучевидными металлическими дорожками и тремя контактами, прижатыми к нему. Такой датчик достался шаровой мыши «в наследство» от прямого привода.

Основными недостатками контактных датчиков являются окисление контактов, быстрый износ и невысокая точность. Поэтому со временем все мыши перешли на бесконтактные оптопарные датчики.

Оптопарные (оптомеханические) датчики

Оптронный датчик состоит из двойной оптопары — светодиода и двух фотодиодов (обычно — инфракрасных) и диска с отверстиями или

лучевидными прорезями, перекрывающего световой поток по мере вращения. При перемещении мыши диск вращается, и с фотодиодов снимается сигнал с частотой, соответствующей скорости перемещения мыши.

Второй фотодиод, смещённый на некоторый угол или имеющий на диске датчика смещённую систему отверстий/про-резей, служит для определения направления вращения диска (свет на нём появляется/исчезает раньше или позже, чем на первом, в зависимости от направления вращения).

Оптические мыши первого поколения

Оптические датчики призваны непосредственно отслеживать перемещение рабочей поверхности относительно мыши. Исключение механической составляющей обеспечивало более высокую надёжность и позволяло увеличить разрешающую способность детектора.

Первое поколение оптических датчиков было представлено различными схемами оптопарных датчиков с непрямой оптической связью — светоизлучающих и воспринимающих отражение от рабочей поверхности светочувительных диодов. Такие датчики имели одно общее свойство — они требовали наличия на рабочей поверхности (мышином коврике) специальной штриховки (перпендикулярными или ромбовидными линиями). В некоторых моделях мышей эти штриховки выполнялись красками, невидимыми в обычном свете (такие коврики даже могли иметь рисунок).

Недостатками таких датчиков обычно называют:

– необходимость использования специального коврика и невозможность его замены другим. Кроме всего прочего, коврики разных оптических мышей часто не были взаимозаменяемыми и не выпускались отдельно;

– необходимость определённой ориентации мыши относительно коврика, в противном случае мышь работала неправильно;

– чувствительность мыши к загрязнению коврика (ведь он соприкасается с рукой пользователя) — датчик неуверенно воспринимал штриховку на загрязнённых местах коврика;

– высокую стоимость устройства.

В СССР и России оптические мыши первого поколения, как правило, встречались только в зарубежных специализированных вычислительных комплексах.

Оптические мыши второго поколения

Оптические мыши второго поколения сделаны на базе микросхемы, содержащей фотосенсор и процессор обработки изображения. Удешевление и миниатюризация компьютерной техники позволили уместить всё это в одном элементе за доступную цену. Фотосенсор периодически сканирует участок рабочей поверхности под мышью. При

изменении рисунка процессор определяет, в какую сторону и на какое расстояние сместилась мышь. Сканируемый участок подсвечивается светодиодом (обычно — красного цвета) под косым углом.

Предполагалось, что такой датчик позволит оптической мыши работать на произвольной поверхности, однако скоро выяснилось, что многие продаваемые модели (в особенности первые широко продаваемые устройства) не так уж и безразличны к рисункам на коврике. На некоторых участках рисунка графический процессор способен сильно ошибаться, что приводит к хаотичным движениям указателя, абсолютно неадекватным реальному перемещению. Для склонных к таким сбоям мышей необходимо подобрать коврик с иным рисунком или вовсе с однотонным покрытием.

Отдельные модели также склонны к детектированию мелких движений при нахождении мыши в состоянии покоя, что проявляется дрожанием указателя на экране, иногда с тенденцией сползания в ту или иную сторону.

Датчики второго поколения постепенно совершенствуются, и в настоящее время мыши, склонные к сбоям, встречаются гораздо реже. Кроме совершенствования датчиков, некоторые модели оборудуются двумя датчиками перемещения сразу, что позволяет, анализируя изменения сразу на двух участках поверхности, исключать возможные ошибки. Такие мыши иногда способны работать на стеклянных, оргстеклянных и зеркальных поверхностях (на которых не работают другие мыши).

Также выпускаются коврики для мышей, специально ориентированные на оптические мыши. Например, коврик, имеющий на поверхности силиконовую плёнку с взвесью блёсток (предполагается, что оптический сенсор гораздо четче определяет перемещения по такой поверхности).

Лазерные мыши

В последние годы была разработана новая, более совершенная разновидность оптического датчика, использующего для подсветки полупроводниковый лазер.

О недостатках таких датчиков пока известно мало, но известно об их преимуществах:

- более высокой надёжности и разрешении;
- успешной работе на стеклянных и зеркальных поверхностях (недоступных оптическим мышам);
- отсутствии сколько-нибудь заметного свечения;
- низком энергопотреблении.

Индукционные мыши

Индукционные мыши используют специальный коврик, работающий по принципу графического планшета, или, собственно, входят в комплект графического планшета. Некоторые планшеты имеют в своем составе

манипулятор, похожий на мышь со стеклянным перекрестием, но работающий по несколько иному принципу.

Индукционные мыши имеют хорошую точность, и их не нужно правильно ориентировать. Индукционная мышь может быть «беспроводной» (к компьютеру подключается планшет, на котором она работает) и иметь индукционное же питание, следовательно, не требовать аккумуляторов, как обычные беспроводные мыши.

Мышь в комплекте графического планшета позволит сэкономить немного места на столе (при условии, что на нём постоянно находится планшет).

Индукционные мыши редки, дороги и не всегда удобны. Мышь для графического планшета практически невозможно поменять на другую (например, больше подходящую по руке и т.п.).

Инерционные мыши

Инерционные мыши используют акселерометры для определения движений мыши по каждой из осей. Обычно инерционные мыши являются беспроводными и имеют выключатель для отключения детектора движений, для перемещения мыши без влияния на указатель.

Патент на инерционную мышь утверждает, что такие мыши имеют меньшее энергопотребление, чем оптические, обладают лучшей чувствительностью, меньшим весом и более просты в использовании.

Сенсорные экраны

Сенсорные экраны (touch screens) предназначены для тех, кто не может пользоваться обычной клавиатурой. Пользователь может ввести символ или команду прикосновением пальца к определенной области экрана. Сенсорные экраны используются в основном на сладах продукции, в ресторанах, супермаркетах. К примеру, в магазинах Muse Inc. (Бруклин), продающих компакт-диски, можно прослушать желаемую композицию, прикоснувшись пальцем к ее названию на экране компьютера. Слушая выбранную мелодию, вы можете одним прикосновением вызвать список других композиций исполнителя.

Устройства автоматизированного ввода информации

Устройства этого типа считывают информацию с носителя, где она уже имеется. Примерами таких систем могут служить кассовые терминалы, сканеры штрих-кодов и другие системы оптического распознавания символов. Одно из преимуществ устройств автоматизированного ввода данных состоит в том, что при их использовании исключаются некоторые ошибки, неизбежные при вводе информации с клавиатуры. Сканер штрих-кодов делает менее чем одну ошибку на 10000 операций, в то время как обученный наборщик ошибается один раз при вводе каждых 1000 строк.

Основные вида устройств автоматизированного ввода информации — системы распознавания магнитных знаков, системы оптического

распознавания символов, системы ввода информации на базе светового пера, сканеры, системы распознавания речи, сенсорные датчики и устройства видеозахвата.

Системы распознавания магнитных знаков (Magnetic Inc Character Recognition, MICR) используются в основном в банковской сфере. В нижней части обычного банковского чека находится код, нанесенный специальными магнитными чернилами. В коде содержится номер банка, номер расчетного счета и номер чека. Система считывает информацию, преобразовывает ее в цифровую форму и передает в банк для обработки.

Системы оптического распознавания символов (Optical Character Recognition, OCR) преобразуют специальным образом нанесенную на носитель информацию в цифровую форму. Наиболее широко используемые устройства этого типа — сканеры штрих-кодов (bar-code scanners), которые применяются в кассовых терминалах магазинов. Эти системы используются также в больницах, библиотеках, на военных объектах, складах продукции и в компаниях по перевозке грузов. В дополнение к данным, идентифицирующим предмет, на который нанесен штрих-код, последний может содержать информацию о времени, дате и физическом положении предмета; таким образом можно, например, отслеживать передвижение груза.

Ручные устройства распознавания информации, такие, как перьевые планшеты, особенно полезны для людей, работающих в сферах сбыта продукции и сервиса, — такие работники избегают «общения» с клавиатурой. Устройства перьевого ввода обычно содержат плоский экран и световое перо, похожее на шариковую ручку. Перьевые планшеты преобразуют буквы и цифры, написанные пользователем на экране, в цифровую форму и передают эти данные в компьютер для обработки. Например, United Parcel Service (UPS), известнейшая в мире компания по доставке грузов, заменила обычные планшеты с листками бумаги, использовавшиеся водителями, на портативные перьевые планшеты. Эти устройства используются для подтверждения заказов и передачи другой информации, необходимой для погрузки и доставки грузов. К недостаткам систем данного вида следует отнести недостаточную точность распознавания информации, написанной от руки.

Сканеры (scanners) преобразуют в цифровую форму графическую информацию (рисунки, чертежи и пр.) и большие объемы текстовой информации. Системы распознавания речи (voice input devices) преобразуют в цифровую форму произносимые пользователем слова. Существует два режима работы подобных устройств. В режиме управления (command mode) вы произносите команды (такие, как «открыть документ», «запустить программу» и т.д.), которые выполняются компьютером. В режиме диктовки (dictation mode) можно надиктовывать компьютеру любой текст. К сожалению, точность распознавания речи таких систем оставляет желать лучшего. Человеческий голос имеет

множество оттенков, на точность распознавания может повлиять интонация, громкость речь, окружающий шум, даже банальный насморк. Тем не менее работа над совершенствованием этих устройств ввода информации продолжается и, несомненно, у них большое будущее. Некоторые отделения Почтовой службы США используют системы распознавания речи для повышения эффективности труда работников, занятых упаковкой и сортировкой почтовых грузов. Вместо того чтобы вводить ZIP-код, работник произносит его, в то время как его руки заняты упаковкой.

Сенсорные датчики (sensors) — это устройства для ввода в компьютер пространственной информации. Например, корпорация General Motors использует сенсоры в своих легковых автомобилях для передачи в бортовой компьютер машины данных об окружающем пространстве и маршруте. Сенсорные датчики также нашли применение в системах виртуальной реальности, игровых приставках и симуляторах.

Устройства видеозахвата (video capture devices) представляют собой небольшие цифровые видеокамеры, соединенные с компьютером. Устройства видеозахвата применяются в основном в системах видеоконференций, которые получают все большее распространение. Благодаря развитию локальных сетей и Интернета, появилась возможность организовывать видеоконференц-связь, находясь в любой точке планеты.

Трекбол (англ. trackball) — указательное устройство ввода информации об относительном перемещении для компьютера. Аналогично мыши по принципу действия и по функциям. Трекбол функционально представляет собой перевернутую мышь. Шар находится сверху или сбоку, и пользователь может вращать его ладонью или пальцами, при этом не перемещая корпус устройства. Несмотря на внешние различия, трекбол и мышь конструктивно похожи — при движении шар приводит во вращение пару валиков или, в более современном варианте, его сканируют оптические датчики перемещения (как в оптической мыши).

Световое перо

Для ввода рисунков в ПК может использоваться так называемое световое перо. Оно применяется сравнительно редко, так как пригодно для работы с крупными объектами, но очень ненадежно при выборе малых объектов.

Световое перо получило дальнейшее развитие при его совместном использовании с дигитайзером (диджитайзером), где пером просто пишут, затем специальные программы переводят рукописный текст или рисунок в цифровой код. Профессиональные световые перья могут определить толщину линий, силу нажатия на перо и другие параметры.

Дигитайзер

Является стандартным устройством ввода для профессиональных графических работ. С помощью программного обеспечения движение руки преобразовывается в формат векторной графики. Дигитайзер способен определять и обрабатывать абсолютно точные координаты, что недоступно другим устройствам ввода.

Планшет (или дигитайзер, диджитайзер, от англ. digitizer) — это устройство для ввода рисунков от руки непосредственно в компьютер. Состоит из пера и плоского планшета, чувствительного к нажатию или близости пера. Также может прилагаться специальная мышь.

В современных планшетах основной рабочей частью также является сеть из проводов (или печатных проводников), подобная той, что была в «Графаконах». Эта сетка имеет достаточно большой шаг (3—6 мм), но механизм регистрации положения пера позволяет получить шаг считывания информации намного меньше шага сетки (до 100 линий на мм).

По принципу работы и технологии есть разные типы планшетов. В электростатических планшетах регистрируется локальное изменение электрического потенциала сетки под пером. В электромагнитных перо излучает электромагнитные волны, а сетка служит приёмником. В обоих случаях на перо должно быть подано питание.

Фирма Wacom (англ.) создала технологию на основе электромагнитного резонанса, когда сетка и излучает, и принимает сигнал, а перо лишь отражает его. Поэтому в таком устройстве запитывать перо не нужно. Но при работе электромагнитных планшетов возможны помехи от излучающих устройств, в частности мониторов. На таком же принципе действия основаны некоторые тачпэды.

Также есть планшеты, в которых нажим пера улавливается за счёт пьезоэлектрического эффекта. При нажатии пера в пределах рабочей поверхности планшета, под которой проложена сетка из тончайших проводников, на пластине пьезоэлектрика возникает разность потенциалов, что позволяет определять координаты нужной точки. Такие планшеты вообще не требуют специального пера и позволяют чертить на рабочей поверхности планшета как на обычной чертёжной доске.

Кроме координат пера, в современных графических планшетах также могут определяться давление пера на рабочую поверхность, наклон, направление и сила сжатия пера рукой.

Также в комплекте графических планшетов совместно с пером может поставляться мышь, которая, однако, работает не как обычная компьютерная мышь, а как особый вид пера. Такая мышь может работать только на планшете. Поскольку разрешение планшета гораздо выше, чем разрешение обычной компьютерной мыши, то использование связки мышь+планшет позволяет достичь значительно более высокой точности при вводе.

6.2 Устройства вывода информации

Жидкокристаллические индикаторы

ЖК-индикаторы — пассивные устройства. Они не генерируют свет и требуют дополнительной подсветки, сами же выполняют роль модулятора, работая в режиме пропускания или отражения света.

Жидкие кристаллы (ЖК) представляют собой органические жидкости, имеющие удлиненные стержнеобразные молекулы. Различают ЖК трех типов (рис. 6.1): смектические, нематические и холестерические.

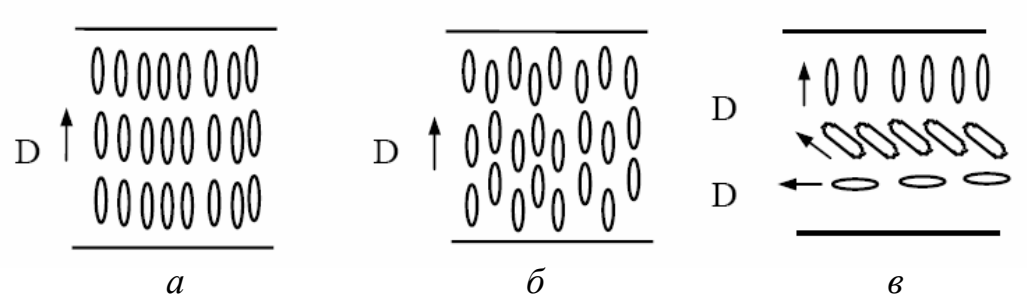


Рисунок 6.1 — Типы жидкокристаллических индикаторов:
а — смектические; *б* — нематические; *в* — холестерические

В смектических ЖК сильно вытянутые молекулы располагаются слоями одинаковой толщины, близкой к длине молекул. Ориентированы молекулы параллельно друг другу. У нематических ЖК отсутствует слоистая структура, а молекулы также ориентированы параллельно друг другу своими длинными осями. Холестерические ЖК имеют структуру слоистую, но в каждом слое молекулы вытянуты в некотором преимущественном направлении.

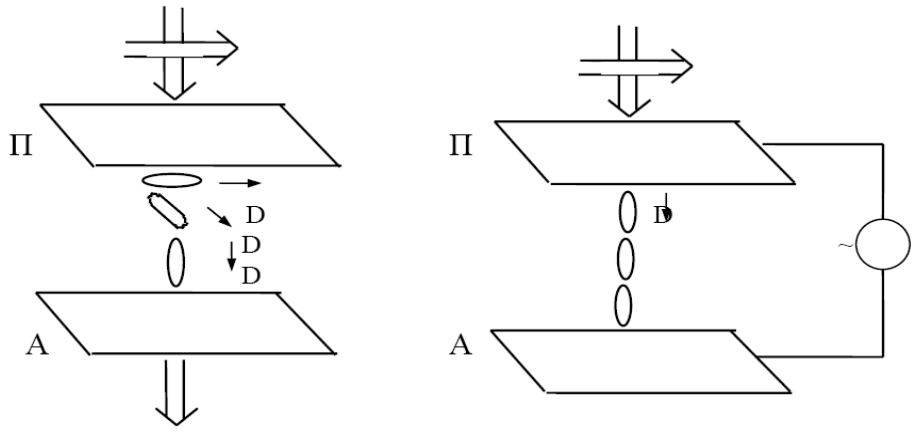
Ориентация отдельной молекулы ЖК подвергается непрерывным тепловым флюктуациям, однако в любой точке жидкости существует средняя ориентация, характеризуемая единичным вектором, называемым директором D. Когда ЖК-вещество занимает большой объем, то в молекуле появляются области с независимыми ориентациями директора. Для придания одинаковой ориентации во всем рабочем пространстве ЖК заключают в узкое (несколько десятков микрометров) пространство между подложками. В результате специфическая ориентация молекул ЖК определяется и соседними молекулами, и граничной поверхностью подложки. Ориентирующее действие достигается напылением на подложки тонких пленок $Si0_2$.

Молекулы ЖК представляют собой индивидуальные диполи. Ориентация молекул может меняться в результате различных электрогидродинамических эффектов, обусловленных протеканием даже небольшого тока или под действием электрического поля.

Конструкция элементарной ячейки ЖК-индикатора проста и содержит две стеклянные пластины, имеющие на внутренней стороне

прозрачное проводящее покрытие. Между пластинами залит ЖК. Толщина ЖК лежит в пределах от 6 до 25 мкм. Такая конструкция по сути представляет собой плоский конденсатор. При отсутствии напряжения на ячейке ЖК-вещество однородно и прозрачно. При приложении к ячейке порогового напряжения возникает волнистая доменная структура. При превышении порогового напряжения доменная структура превращается в ячеистую, затем в жидкости возникает вихревое движение. ЖК теряет оптическую однородность и рассеивает свет во всех направлениях. Этот эффект называют динамическим рассеиванием. В настоящее время распространены индикаторы на основе эффекта динамического рассеивания, а также индикаторы, использующие полевой твист-эффект (закручивание) и эффект типа «гость-хозяин».

В настоящее время наиболее распространены индикаторы, использующие полевой твист-эффект (от англ. twist — закручивание). Работа ячейки со скрещенными поляризатором П и анализатором А показана на рис. 6.2.



а *б*

Рисунок 6.2 — Работа ЖК-индикатора на твист-эффекте при напряжениях: *а* — нулевом; *б* — превышающем пороговое

В отсутствие напряжения питания на ячейке молекулы ЖК закручены приблизительно на 90° благодаря ориентирующему действию подложек П и А.

Поляризатор — это оптический элемент, пропускающий свет, поляризованный в одном направлении, и гасящий свет, поляризованный в противоположном направлении, в зависимости от ориентации поляризатора. Если оси второго поляризатора, называемого анализатором, параллельны осям первого, то свет проходит через второй поляризатор; если же оси анализатора перпендикулярны, излучение гасится.

Свет, падающий сверху, поляризуется таким образом, что его вектор поляризации совпадает с направлением директора D у верхней подложки. При прохождении через ЖК плоскость поляризации света вращается (как директор у молекул ЖК) и свет проходит через анализатор. При питании

ячейки напряжением выше порогового вектор поляризации ЖК приобретает вертикальное направление и ЖК не вращают плоскость поляризации, а анализатор не пропускает свет.

ЖК-индикаторы имеют преимущества по сравнению с индикаторами на эффекте динамического рассеяния (меньше рабочие токи: 1—3 мкА/см^2 вместо 10 мкА/см^2, и поэтому — большая долговечность). Быстродействие ЖК на твист-эффекте гораздо выше, чем при использовании динамического рассеяния.

К недостаткам ЖК-индикаторов на твист-эффекте относится меньший, чем у индикаторов на эффекте динамического рассеяния, угол обзора, что связано с узкой диаграммой направленности света при твист-эффекте и влиянием поляризаторов. Применение поляризаторов приводит к потерям до 50 % света, а также повышает стоимость индикаторов.

Индикаторы без поляризаторов могут быть созданы на основе эффекта «гость-хозяин». Стержневидные молекулы красителя (гость) вводятся в ЖК (хозяин). Молекулы красителя стремятся ориентироваться параллельно осям молекул ЖК (рис. 6.3).

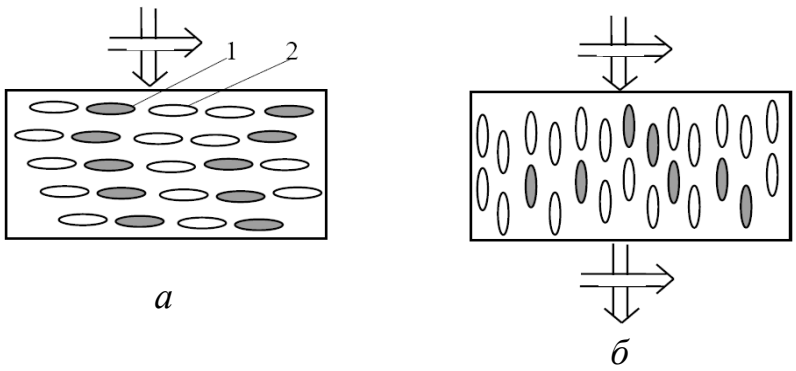


Рисунок 6.3 — Работа ЖК-ячейки на эффекте «гость-хозяин» при напряжениях: *а* — нулевом, *б* — превышающем пороговое; 1 — молекулы красителя; 2 — молекулы ЖК

В начальном состоянии, при нулевом напряжении на ЖК-ячейке, свет с любым направлением поляризации поглощается (рис. 6.3, *а*). При наложении достаточно сильного электрического поля ЖК-вещество переходит в состояние, в котором все молекулы красителя ориентированы вертикально, а падающий на ячейку свет свободно проходит сквозь нее (рис. 6.3, *б*).

Описанная система перспективна, так как позволяет получить почти черное позитивное изображение на белом фоне при высокой яркости и достаточно широком угле обзора. Контраст у индикаторов на эффекте «гость-хозяин» несколько хуже вследствие поглощения света красителем.

Достоинства ЖК-индикаторов заключаются в следующем:

- малая потребляемая мощность (1÷10 мкВт/ см^2);
- работа при высоком уровне внешней освещенности;
- простота конструкции и технологии изготовления;

– низкая стоимость, низкое рабочее напряжение.

К основным недостаткам ЖК-индикаторов следует отнести узкий диапазон рабочих температур (от –10 до +60° С), длительные переходные процессы, к тому же зависящие от температуры.

ЭЛТ

Электронно-лучевые трубки были достаточно полно рассмотрены в соответствующем курсе электронных приборов, поэтому здесь мы только напомним, что они из себя представляют. В качестве примера можно рассмотреть осциллографическую трубку с электростатическим управлением.

Осциллографическая трубка — это прибор, предназначенный для преобразования электрического сигнала в световое изображение с помощью тонкого электронного луча, направленного на люминесцирующий экран.

На рис. 6.4 показана схема устройства трубки с электростатической фокусировкой и электростатическим отклонением луча.

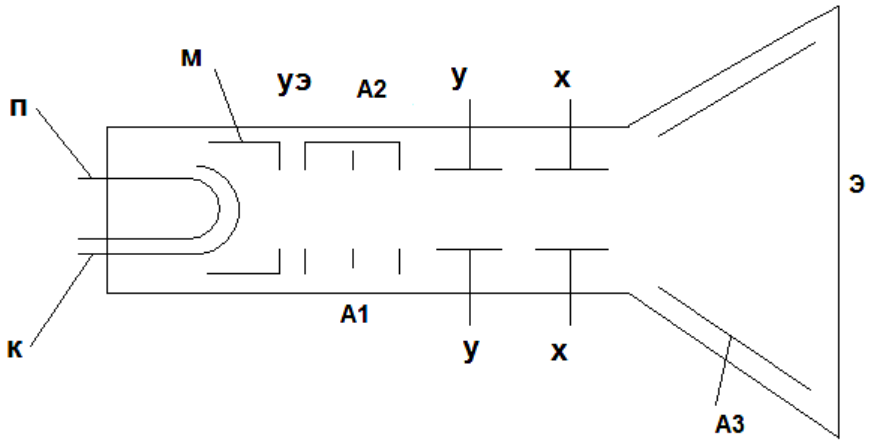


Рисунок 6.4 — Схема устройства осциллографической трубки:
П — подогреватель; К — катод; М — модулятор; УЭ — ускоряющий электрод; А1 — первый анод; А2 — второй анод; У — пластины вертикального отклонения; Х — пластины горизонтального отклонения; А3 — третий анод; Э — экран

Транспаранты

Известно, что массовое создание больших плоских экранов на жидких кристаллах сталкивается с трудностями не принципиального, а чисто технологического характера. Хотя принципиально возможность создания таких экранов продемонстрирована, однако в связи со сложностью их производства при современной технологии их стоимость оказывается очень высокой. Поэтому возникла идея создания проекционных устройств на жидких кристаллах, в которых изображение, полученное на жидкокристаллическом экране малого размера, могло бы быть спроектировано в увеличенном виде на обычный экран, подобно тому, как это происходит в кинотеатре с кадрами кинопленки. Оказалось,

что такие устройства могут быть реализованы на жидких кристаллах, если использовать сэндвичные структуры, в которые наряду со слоем жидкого кристалла входит слой фотополупроводника. Причем запись изображения в жидком кристалле, осуществляемая с помощью фотополупроводника, производится лучом света.

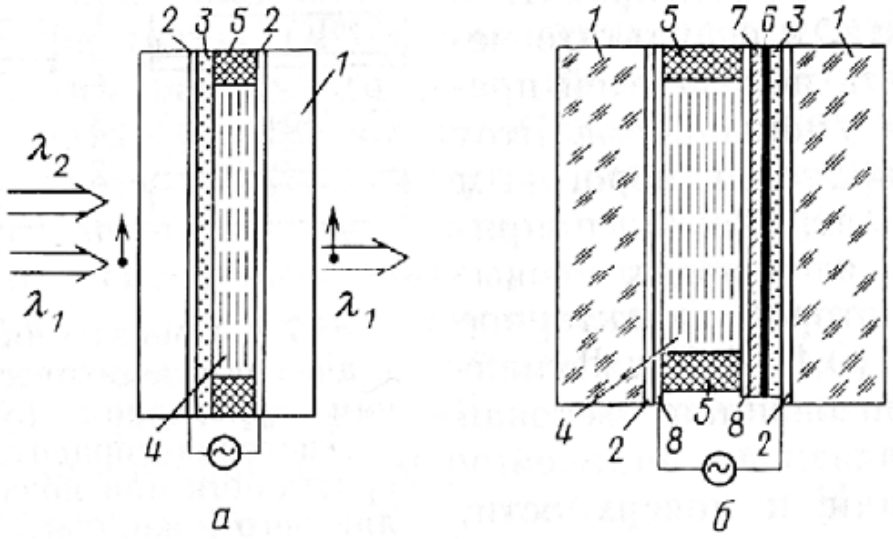


Рисунок 6.5 — Оптически управляемые транспаранты на основе структуры фотопроводник — ЖК: *а* — с модуляцией проходящего света; *б* — с модуляцией отраженного света; 1 — подложки; 2 — прозрачные электроды; 3 — фотопроводящий слой; 4 — ЖК; 5 — прокладки; 6 — светоблокирующий слой; 7 — диэлектрическое зеркало; 8 — ориентирующий слой

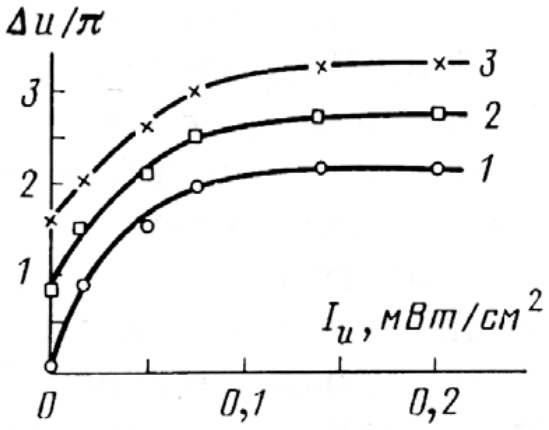


Рисунок 6.6 — Статические модуляционные характеристики оптически управляемого транспаранта на основе нематического ЖК (S — эффект). Амплитуда управляющего напряжения: 1 — 3.2 В, 2 — 3.6 В, 3—4 В

Полупроводниковые индикаторы (светодиодные индикаторы)

Полупроводниковые индикаторы (ППИ) представляют собой твердотельные приборы, работающие на p-n-переходах, и применяются для отображения знаковой информации при относительно небольших размерах символа и ограниченном числе знакомест. Одной из отличительных особенностей ППИ является их совместимость по электрическим характеристикам с обычными интегральными микросхемами. При напряжении питания 3÷5 В ППИ обладают малой инерционностью (менее 50 нс) и небольшими габаритами.

Первые полупроводниковые светоизлучающие приборы (на основе карбида кремния — SiC) были созданы советским ученым О.В. Лосевым в 1922 г., который также заложил основы современных представлений о механизме излучения (электролюминесценции) и возможных применениях светоизлучающих p-n-переходов.

В таблице приведены характеристики некоторых материалов, способных создавать излучение в области видимого спектра.

Материал	Цвет излучения	Длина волны, мкм	Яркость, кд/м^2
GaP: ZnO	Красный	0,699	350
GaP: N	Зеленый	0,55	470
SiC	Желтый	0,59	10
$GaAs_{0.35}P_{0.65}$:N	Оранжевый	0,632	–
GaN	Голубой	0,44	–

Для изготовления цифровых и цифро-буквенных ППИ используют технологические методы, широко применяемые в производстве интегральных микросхем. В зависимости от размеров ППИ изготовляются по монолитной и по гибридной технологии. В первом случае — это интегральный блок светодиодов, выполненный на одном полупроводниковом кристалле. Так как размеры кристалла ограничены, то ППИ имеют малые размеры. Во втором случае излучающая часть индикатора представляет собой сборку дискретных светодиодных индикатор (СИД) на миниатюрной печатной плате. Гибридный вариант является основным для средних и больших ППИ.

Матричные ППИ по конструкции делятся на два типа: отражающие и не отражающие. Конструкции таких ППИ приведены на рис. 6.7 и рис. 6.8.

В индикаторе отражающего типа на подложке из алюминия продольно расположены катодные электроды в виде параллельных линий шириной 0.5 мм и шагом 0.8 мм между ними. Анодные электроды сформированы отдельно на краю подложки ортогонально катодным электродам. Отражающая стеклянная пластина имеет отверстия с шагом 0.8 мм, в которые устанавливают СИД, образующие матрицу

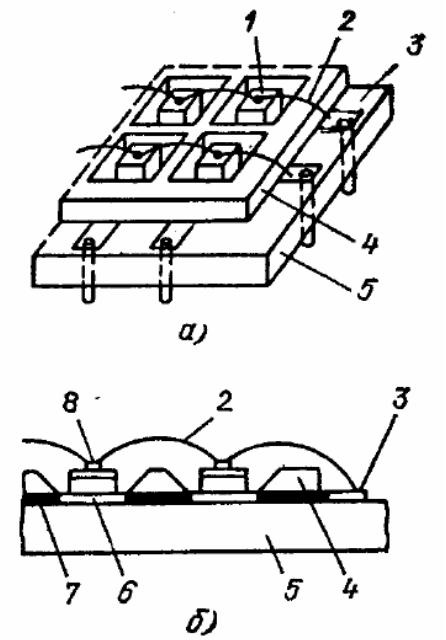


Рисунок 6.7 — Структура ППИ отражающего типа (*а*) и поперечный разрез такого индикатора (*б*): 1 — светодиод; 2 — проводник из золота; 3 — контактная площадка; 4 — отражающая стеклянная пластинка; 5 — алюминиевая подложка; 6 — катод; 7 — черная смола; 8 — анод

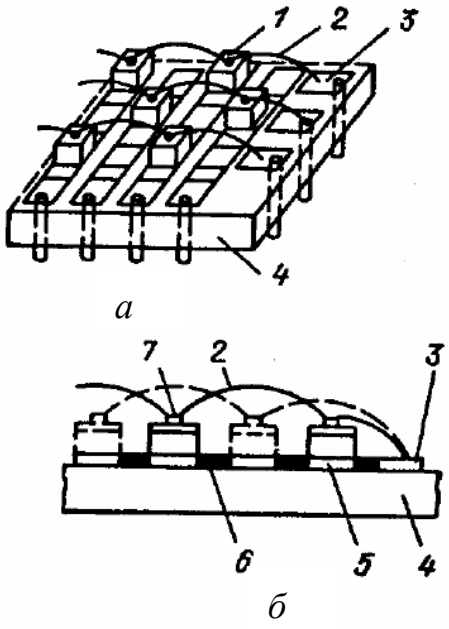


Рисунок 6.8 — Структура ППИ не отражающего типа (*а*) и поперечный разрез такого индикатора (*б*): 1 — светодиод; 2 — проводник из золота; 3 — контактная площадка; 4 — алюминиевая подложка; 5 — катод; 6 — черная смола; 7 — анод

. Стенки отверстий в отражающей стеклянной пластине имеют обращенный вверх наклон около 45°. Они покрыты тонкой отражающей пленкой золота. На поверхность алюминиевой подложки нанесен слой черной смолы, за исключением мест расположения контактных площадок

диодных кристаллов, контактных площадок для проводников и электродов-выводов (анодных электродов). В индикаторах неотражающего типа катодные электроды шириной 0.3 мм имеют шаг 0.5 мм. Светодиоды расположены зигзагообразно на алюминиевой подложке.

Лазерные УО

В последние годы активно проводятся исследования возможности использования лазеров для построения системы отображения информации (СОИ) коллективного пользования. Применение лазеров для отображения информации считается перспективным.

Главной проблемой, стоящей перед разработчиками лазерных устройств отображения, является проблема управления световым лучом. Известные методы отклонения светового луча не обеспечивают такой гибкости управления, как, например, метод управления электронным лучом. Механические способы отклонения луча с помощью качающихся и вращающихся зеркал и призм обладают большой инерционностью и используются сравнительно редко. Электрические способы имеют малую чувствительность, но считаются более перспективными.

Представляют интерес следующие свойства излучения лазеров: пространственная когерентность, временная когерентность, цвет и яркость.

При разработке СОИ на лазерах используются следующие методы: визуальная лазерная индикация, когда на экран направляется собственный свет лазера; индикация с активным экраном, когда луч лазера применяется лишь для управления световым излучением некоторого активного материала экрана; лазерно-лучевой световой клапан, когда луч лазера обеспечивает местное управление оптическими параметрами некоторого материала (его коэффициентом отражения или коэффициентом пропускания), а отдельный источник обычного типа дает свет для проекции на экран; лазерный генератор изображения с непосредственным воздействием на объемный резонатор (такой генератор позволяет получать двумерное изображение непосредственно от лазерного источника).

При отображении информации используют способ «последовательной выдачи», когда луч лазера последовательно обходит все точки поверхности экрана, либо способ «выборочного отображения», когда луч лазера направляется только на те элементы экрана, в которые вводится информация.

Информацию можно накладывать на лазерный луч путем изменения его интенсивности. Для изменения интенсивности луча лазера используются различные способы. Необходимость воспроизведения широкой полосы частот с целью получения высокой разрешающей способности требует быстродействующих устройств, в качестве которых используют электрооптические модуляторы с линейным или квадратичным эффектом.

Дефлекторы, осуществляющие управление лучом, основаны на различных способах отклонения луча: механическом, рефракционном,

дифракционном, когерентной оптической фазовой решетки, двоичного электрооптического управления положением луча.

Механический способ реализуется с помощью применения двух вращающихся многогранных призм или зеркал с весьма высоким коэффициентом отражения, перемещаемых по горизонтали и вертикали пьезоэлектрическим и гальванометрическим приводами. Способ обеспечивает относительно большие рабочие углы отклонения (до 10°—12°) и достаточно высокий оптический коэффициент полезного действия. Быстродействие таких устройств мало, поэтому их можно использовать лишь при режиме последовательной выдачи. Кроме того, им свойственны нестабильность, жесткие допуски на элементы, трудности синхронизации и т.д.

Рефракционный способ реализует известное оптическое свойство — отклонение светового луча вследствие преломления (рефракции) на границе двух прозрачных сред. В этом случае применяют электрооптическую призму или ультразвуковую рефракционную ячейку.

Дифракционный способ может быть использован, если диаметр падающего светового пучка существенно больше длины ультразвуковой волны, когда возникает дифракция света (при растровой развертке). Он обеспечивает малые рабочие углы (до нескольких градусов) и низкую эффективность отклонения.

Способ когерентной оптической фазовой решетки основан на свойстве излучения лазера, характеризующемся высокой степенью временной и пространственной когерентности. Это свойство используется для отклонения лазерного луча за счет разделения его на множество параллельных лучей и изменения относительных фаз между соседними лучами в ближней зоне поля. Этот способ требует высокой стабильности как источника света, так и дефлектора и имеет ряд других ограничений.

Способ двоичного электрооптического управления световым лучом основан на использовании свойства двойного лучепреломления некоторых веществ. В таких веществах обычный неполяризованный луч света расщепляется на два луча. Один из лучей называется обыкновенным, а другой — необыкновенным. Эти лучи линейно поляризованы, причем плоскости их поляризации взаимно ортогональны. Если свет, падающий на вещество с двойным лучепреломлением (по нормали), полностью линейно поляризован и его плоскость поляризации совпадает с плоскостью поляризации обыкновенного луча, то свет проходит не отклоняясь. Если свет линейно поляризован в плоскости необыкновенного луча, выходной луч оказывается смещенным относительно точки выхода обыкновенного луча. Величина такого смещения пропорциональна толщине кристалла с двойным лучепреломлением (КДП). В качестве такого вещества используют кальцит. Кристалл такого рода может выполнять функцию двоичного переключения линейно поляризованного света, преобразующего обыкновенный О-луч в необыкновенный Н-луч путем

введения фазового запаздывания на 180° при воздействии на кристалл напряжения полуволнового запаздывания.

6.3 Речь техническими средствами

Природа наделила человека даром речи. Тысячелетиями люди мечтали поделиться эти даром с объектами живого и неживого мира. Лишь в наше время, постигнув физиологический механизм действия голосового тракта, люди научились конструировать «говорящие машины». Благодаря современным электронным технологиям доступны компактные и недорогие синтезаторы речи и устройства ее распознавания.

Существует три основных технологически различных подхода к проблеме синтеза речи:

1. Метод кодирования-восстановления формы сигналов.
2. Аналоговый метод синтеза формантных частот.
3. Цифровое моделирование голосового тракта.

Причем первый является одним из самых основных и элементарных подходов к созданию говорящего устройства. Но у этого метода есть основной недостаток — для хранения речевых сообщений необходима память значительного объема.

Второй метод принципиально отличается от первого. Синтезатор на основе этого метода, как правило, имеет неестественное звучание, голос робота, т.к. выходная речь не имеет в своей основе естественной человеческой речи. В основу такого синтезатора положены принципы акустического моделирования голосового тракта человека. Такой метод дает возможность иметь неограниченный словарь, но иногда нелегко понять, что говорит формантный синтезатор.

Формантный метод синтеза речи точнее всего можно описать как цифровое моделирование голосового тракта человека. Наиболее распространенная реализация этого метода известна под названием линейного предиктивного кодирования (ЛПК) речи. Преимущества синтезаторов третьего типа обусловлены простотой их реализации в виде цифровых интегральных микросхем, вытекающей отсюда меньшей себестоимостью производства и меньшей эквивалентной скоростью передачи информации.

Есть некоторые основные правила «этикета», которым должен следовать говорящий компьютер:

1. Если компьютер может сказать что-нибудь не вовремя, он это сделает.

2. Если вы будете неоднократно демонстрировать говорящий компьютер одним и тем же людям, то каждый раз они будут надеяться услышать речь лучшего качества.

3. В то время как человека можно попросить помолчать, к запрограммированному компьютерному синтезатору речи бесполезно обращаться с подобной просьбой.

4. Компьютеры, прерывающие говорящего, как и люди, поступают бестактно. «Идеально воспитанный» компьютер должен вступать в разговор, лишь когда в помещении тишина.

5. Короткая, неожиданно сказанная компьютером фраза потеряется в шуме обычного разговора.

6. Если компьютер намерен сообщить что-то важное, он должен заранее оповестить об этом сигналом или каким-либо звуком, привлекающим внимание слушателей.

7. Говорящий компьютер похож на ребенка: он знает, как говорить, но не знает когда.

8. Компьютер, говорящий слишком много, будет выключен.

9. Компьютер, говорящий слишком мало, останется без внимания.

10. Слушатель должен быть готов услышать искусственный голос компьютера, если, конечно, последний не обладает натуральным голосом.

Эти правила особенно важны, когда вы только начинаете «выводить» ваш компьютер в общество.

Метод непосредственного кодирования-восстановления

Этот простейший метод похож на цифровую запись. Одной из задач здесь ставится выборка аналоговых сигналов. В процессе преобразования звуков речи микрофоном генерируется электрический сигнал, который содержит как медленно, так и весьма быстро меняющиеся компоненты. Если бы мы стали делать выборки из этого сигнала с относительно низкой частотой, то самое большое, что нам бы удалось установить, — это то, что в сигнале действительно существуют медленно меняющиеся компоненты. Если же мы начнем осуществлять выборку электрических сигналов микрофона со все возрастающей частотой, то будем обнаруживать все более высокочастотные компоненты сигнала. Этот пример иллюстрируется рисунком 6.9.

Процесс удвоения частоты можно проводить до тех пор, пока дальнейшее повышение ничего более не добавит к выборке. Далее для обработки сигнала компьютером используются АЦП и далее для вывода на динамик — ЦАП.

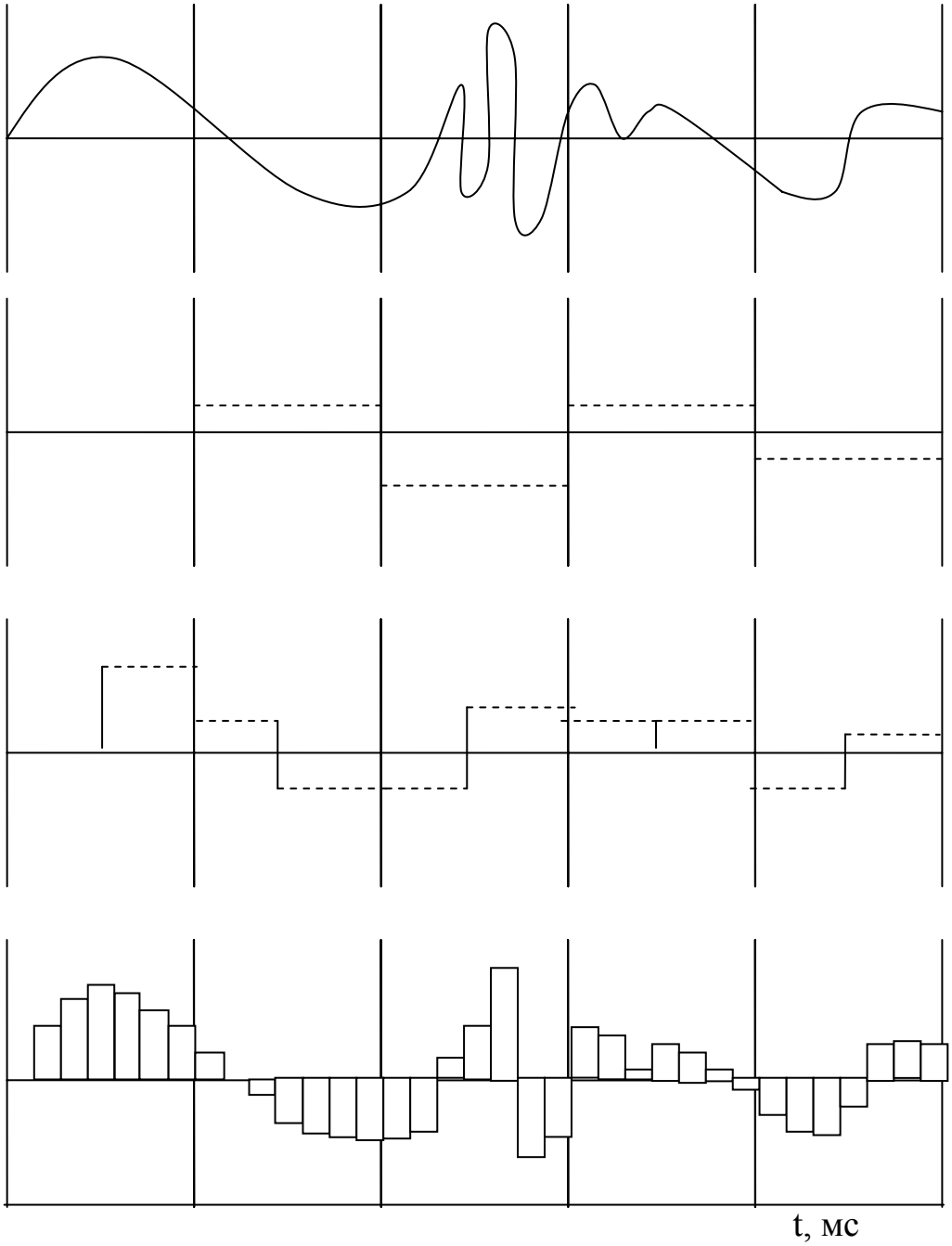


Рисунок 6.9 — Увеличение частоты выборки сигнала

Аналоговый синтез формантных частот

В отличие от предыдущего метода синтез формантных частот представляет собой способ выражения той же речи искусственным путем.

Блок-схема расширенного формантного синтезатора приведена на рис. 6.10

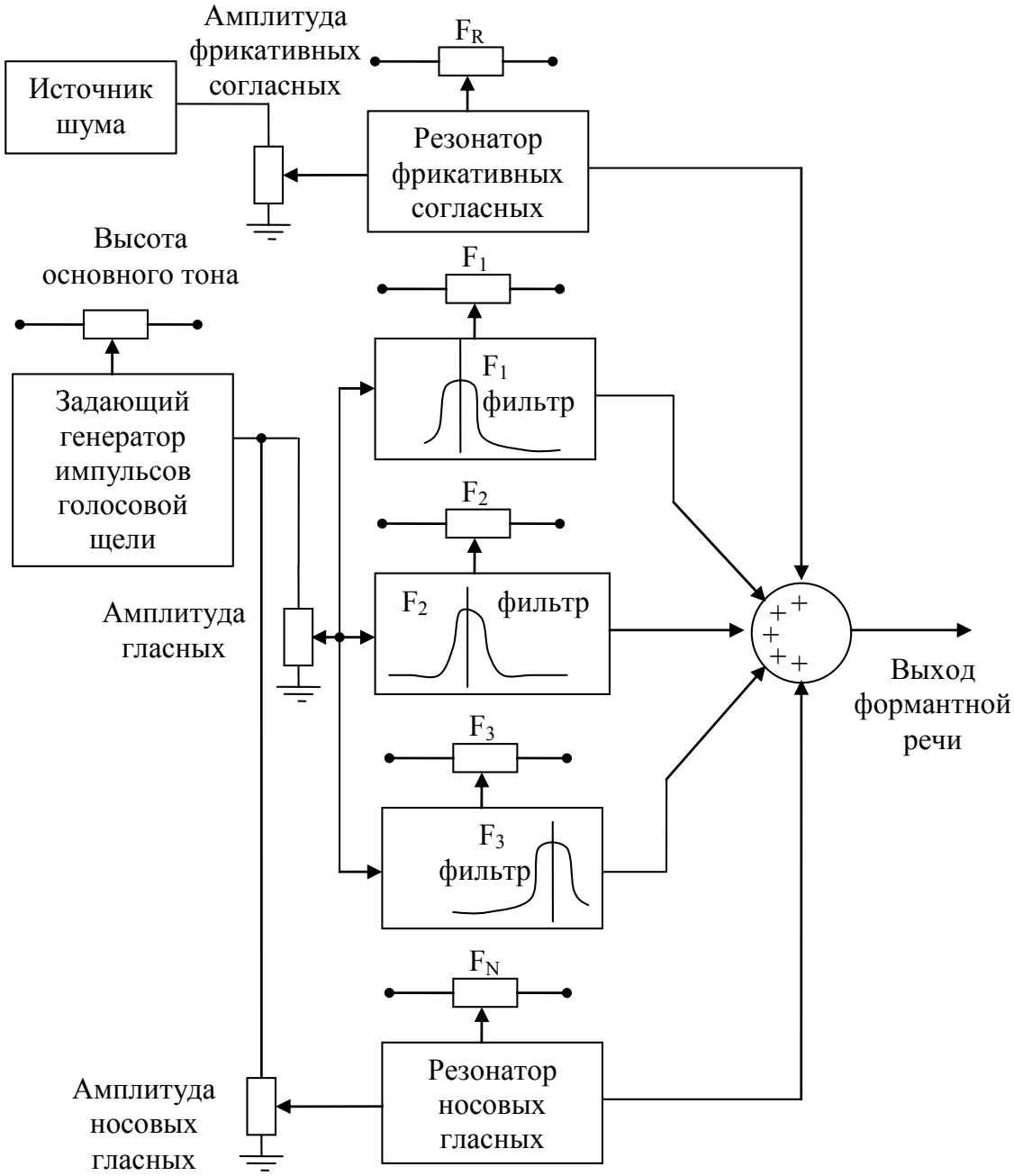


Рисунок 6.10 — Полная схема формантного синтеза речи

Синтез речи по методу линейного предиктивного кодирования (ЛПК)

Этот метод наиболее труден для понимания, т.к. регенерация речи при его использовании выполняется математически с использованием уравнений преобразования закодированной речи в спектры ее исходных частот. На рисунке показана упрощенная схема генерации речи по методу ЛПК.

Рисунок 6.11 — Принцип линейного предиктивного кодирования

7 ТЕХНОЛОГИИ ПОВЫШЕНИЯ НАДЕЖНОСТИ СИСТЕМ

Достижение высокой надёжности электронных систем требует использования специальных компонентов и методов проектирования, соответствующих конечным условиям эксплуатации оборудования.

К сожалению, истинную причину отказа выяснить удаётся далеко не всегда. Выход из строя компонента подразумевает, что он не функционирует должным образом или его параметры больше не соответствуют изначальным техническим требованиям. Это может произойти по целому ряду причин, в том числе из-за перегрузок по току или напряжению, чрезмерного нагревания, воздействия агрессивных химических веществ или повышенной влажности, а также некоторых других условий, встречающихся в процессе производства и использования оборудования.

Постоянно увеличивающаяся сложность компонентов электронного оборудования и потребность в портативных маломощных устройствах, способных работать в жёстких условиях эксплуатации, — вот основные причины, заставляющие разработчиков создавать новые высоконадёжные

устройства. Ключом к достижению повышенной надёжности даже в чрезвычайно жёстких условиях эксплуатации является использование такого оборудования и программного обеспечения, в основе которых лежит специальная, защищённая от ошибок философия проектирования. Всесторонний обзор различных типов отказов и профилактических методов проектирования, позволяющих их избежать, поможет разработчикам значительно повысить надёжность оборудования.

7.1 Источники бесперебойного питания (ИБП)

ИБП — устройство, предназначенное для поддержания бесперебойного питания приборов и оборудования, а также обеспечения качества подаваемой энергии.

ИБП выбирается исходя из видов неисправности электросети, исключаемых с помощью ИБП; суммарной мощности техники, подключаемой к ИБП; наличия дополнительных функциональных возможностей ИБП. Источники бесперебойного питания подразделяются на несколько типов.

Один из самых распространенных типов бесперебойников — так называемый ИБП Line Interactive. Работают они по следующему принципу: когда с напряжением в сети никаких проблем нет, бесперебойник, по сути, бездействует. Системный блок питается напрямую от розетки, но стоит ситуации измениться — например, скачок напряжения или, наоборот, резкое его падение — как в действие вступает стоящая на входе ИБП триада из инвертора, выпрямителя и автотрансформатора. Выпрямитель передает ток из сети на трансформатор, который, в свою очередь, преобразовывает его из постоянного в переменный. Автотрансформатор же все время отслеживает скачки напряжения, чтобы вовремя подключить к работе пару «инвертор/выпрямитель».

Когда перепадов напряжения нет и компьютер работает напрямую от сети, выпрямитель работает еще и как зарядник батареи аккумулятора, обеспечивая достаточный заряд для обеспечения работы компьютера, если вдруг напряжение в сети исчезнет. Все в данном типе бесперебойников было бы хорошо, если бы не одно «но»... Дело в том, что, когда ИБП переключает компьютер с работы от сети на работу от аккумулятора или наоборот, неизбежно возникает скачок напряжения. Нельзя быть уверенным, что в момент переключения компьютер не перезагрузится. После этого вы сможете запитать системник от ИБП, даже если электричество еще не включилось, но данные будут потеряны. Вероятность такого исхода очень мала, чаще Line Interactive ИБП справляются со своей работой нормально, но шанс все же есть.

Еще один тип источников бесперебойного питания — так называемые онлайновые феррорезонансные (Ferroresonant). В данные приборы встроен трансформатор большой индуктивности, который не обязательно постоянно держать подключенным к сети питания. Во

включенном состоянии он накапливает в себе большое количество энергии, которая в случае сбоев в сети идет на поддержание работоспособности системного блока. Такие ИБП хороши тем, что в момент переключения с питания от сети на питание от бесперебойника нет скачка напряжения. Переход происходит плавно и незаметно для компьютера. Единственный недостаток такого вида приборов — необходимость их периодически переключать в различные режимы работы: накопления заряда и поддерживающий.

Но при всей надежности феррорезонансных ИБП они все же уступают самому продвинутому типу бесперебойников — онлайновым с двойным преобразованием (Online Double Conversion). Данный тип ИБП, когда скачков напряжения в сети нет, питает компьютер переменным током, который получается в результате действия инвертора. Постоянный ток от выпрямителя обеспечивает зарядку батареи аккумулятора, если она разряжена. Во время нормального напряжения в сети подобный ИБП участвует в преобразовании тока для питания компьютера, когда же в сети происходит сбой, то подключается аккумулятор, который и питает системник до стабилизации напряжения.

Данный тип бесперебойников — самый надежный. Будучи подключенными к сети, они постоянно поддерживают максимальный заряд аккумулятора, который в случае неполадок в электропитании обеспечивает стабильную работу компьютера на протяжении нескольких минут или даже часов.

Еще один распространенный тип ИБП — Stand By. Работают они по тому же принципу, что и Line Interactive: когда в сети все спокойно, занимаются только тем, что подзаряжают батарею, когда же напряжение начинает шалить, начинают питать компьютер от встроенного аккумулятора. В схеме также есть выпрямитель и инвертор, преобразующий постоянный ток в переменный. В отличие от бесперебойников типа LI у Stand By нет встроенного автотрансформатора, так что в работе они еще менее надежны, чем Line Interactive. Зато дешевы.

7.2 Метод резервирования

Мажоритарное резервирование используют для повышения надежности устройств. Устройство имеет три или более идентичных канала. Результаты вычислений в каждом канале сравниваются между собой. С помощью мажоритарного логического элемента путем голосования выбирается наиболее достоверный результат. В простейшем случае этот результат представлен единственным сигналом, который используется, например, для выключения-включения электродвигателя [8].

Проблема заключается в том, что мажоритарный элемент и выходной транзистор с точки зрения надежности составляют наиболее слабое звено схемы — их отказы приводят к потере работоспособности всей системы. Вероятнее всего, при таком отказе электродвигатель будет

либо постоянно выключен, либо постоянно включен. В зависимости от требований к системе наиболее нежелательно либо ложное (несанкционированное) включение электродвигателя, либо его ложное выключение. В первой ситуации можно применить дублированный выходной каскад с последовательно соединенными выходными транзисторами, во второй — с параллельно включенными транзисторами.

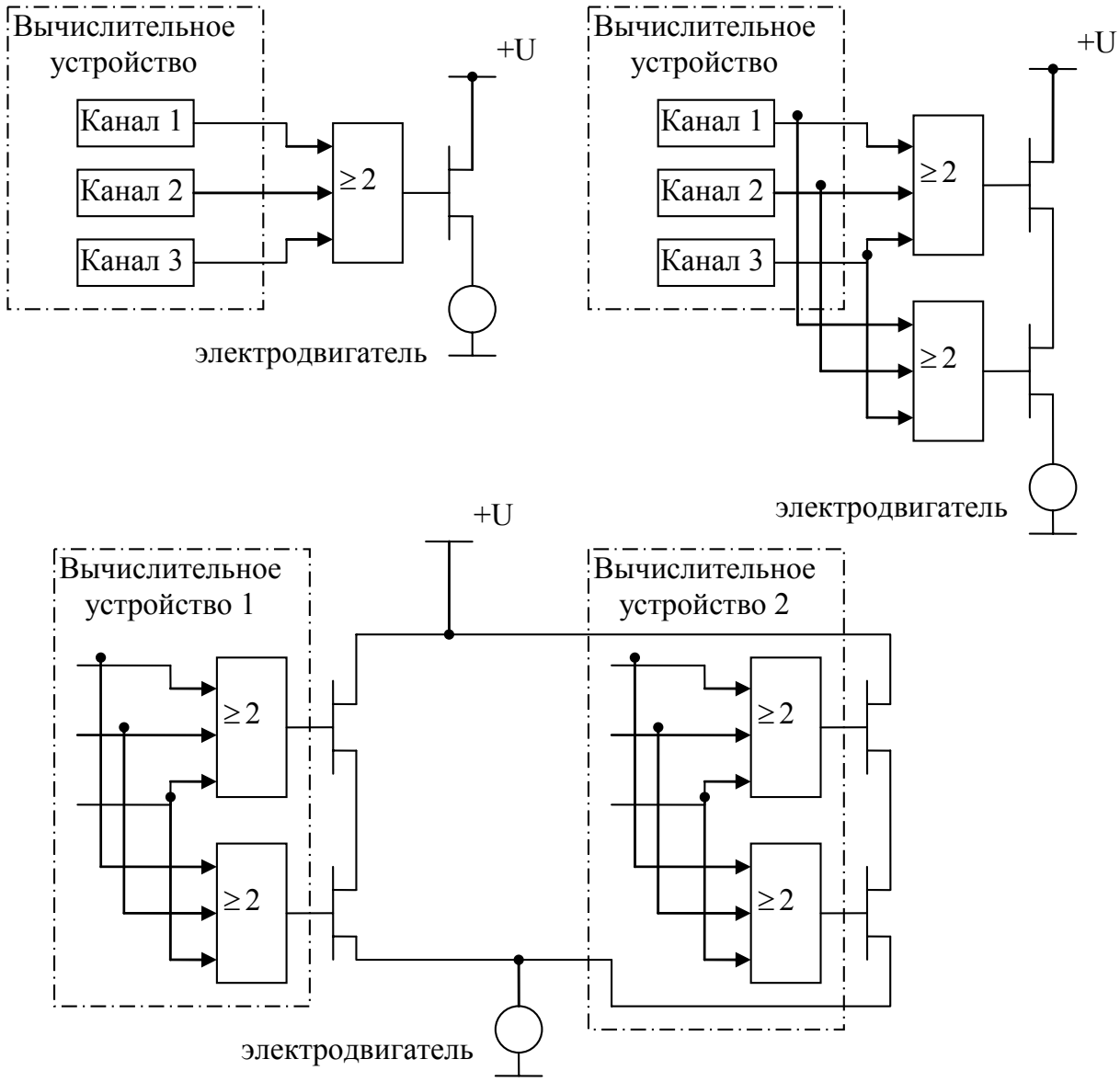


Рисунок 7.1 — Резервированная система

Самоконтролируемый выходной каскад

Мажоритарно-резервированное устройство содержит три идентичных канала, к которым подключен выходной каскад на элементах D1—D7 (рис. 7.2) [8]. Мажоритарный элемент D1 формирует на выходе сигнал V = 1, если на его входах две или три единицы, и сигнал V = 0, если на входах два или три нуля (голосование по большинству). Компаратор D2

вырабатывает сигнал U = 1 при взаимном равенстве входных сигналов, т.е. при Q1 = = Q2 = Q3 = 0 или Q1 = Q2 = Q3 = 1.

Сумматор по модулю два D3 формирует сигнал W = 1, если число единиц, поданных на его входы, нечетно, и W = 0 при четном числе единиц. Последовательная цепь из элементов Исключающее ИЛИ D4—D6 суммирует по модулю два сигналы V, U, W и тестовый сигнал T, элемент НЕ D7 инвертирует выходной сигнал. Результирующие сигналы T’ и U отображают состояние устройства (нет ошибки — есть ошибка), причем обнаруживаются неисправности элементов D1—D7. Трехканальное резервирование позволяет исправлять все одиночные и некоторые двойные ошибки.

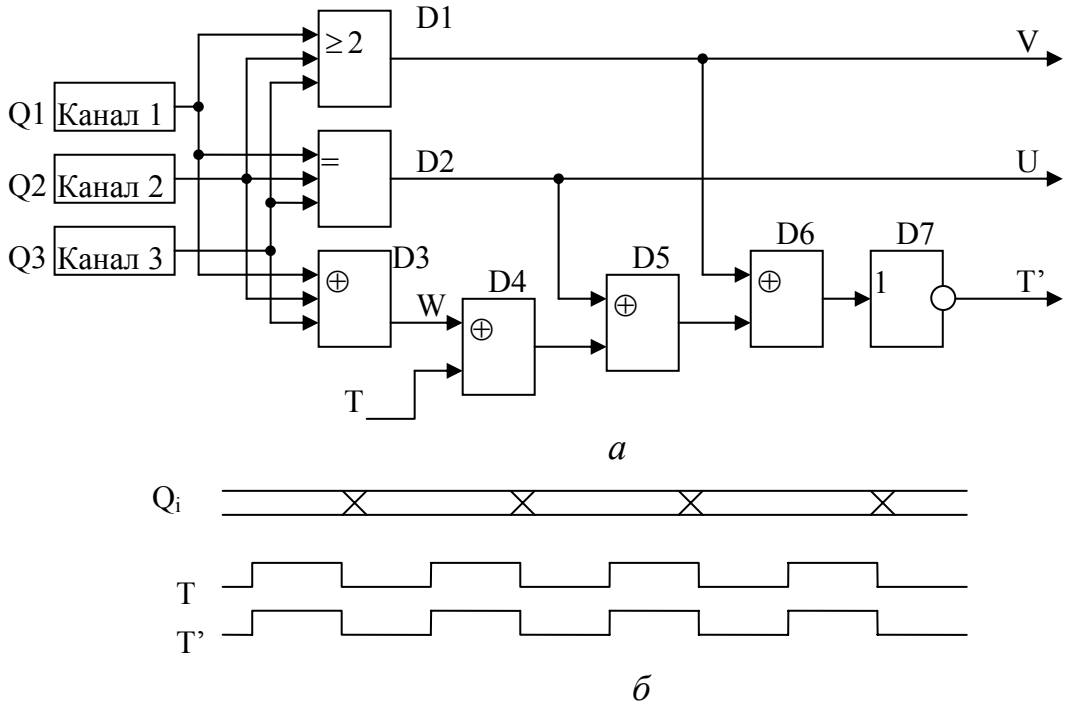


Рисунок 7.2 — Самоконтролируемый выходной каскад (*а*) мажоритарно-резервированного устройства и временные диаграммы (*б*) его работы в отсутствие ошибок

При отсутствии неисправностей аппаратуры и безошибочных вычислениях сигналы Q1—Q3 одинаковы (см. строки табл. 7.1) [8]. Тестовый сигнал T, подаваемый на контрольный вход устройства, суммируется по модулю два с нечетным числом единиц, поэтому с учетом инвертирования результата получим, что T’ = T, U = 1. Внешняя система контроля (на рисунке не показана) в каждом такте работы (при T = 0 и T = 1, см. рис. 7.2, *б*) принимает к сведению, что ошибок на выходах каналов 1—3 нет (U = 1), а сам выходной каскад (элементы D1—D7) исправен (T’ = T).

Таблица 7.1

Номер строки	Входные сигналы				Выходные сигналы				
	Q1	Q2	Q3	T	V	U	W	T	
1	0	0	0	0	0	1	0	0	Ошибок нет
2	0	0	0	1	0	1	0	1	
3	1	1	1	0	1	1	1	0	
4	1	1	1	1	1	1	1	1	
5	0	0	1	0	0	0	1	0	Ошибка типа ложная 1
6	0	0	1	1	0	0	1	1	
7	0	1	0	0	0	0	1	0	
8	0	1	0	1	0	0	1	1	
9	1	0	0	0	0	0	1	0	
10	1	0	0	1	0	0	1	1	
11	1	1	0	0	1	0	0	0	Ошибка типа ложный 0
12	1	1	0	1	1	0	0	1	
13	1	0	1	0	1	0	0	0	
14	1	0	1	1	1	0	0	1	
15	0	1	1	0	1	0	0	0	
16	0	1	1	1	1	0	0	1	

Таким образом, исключается неконтролируемое накопление ошибок в устройстве. Если бы не было цепи D3—D7, то, например, при U = 1 (из-за отказа элемента D2) внешняя система контроля оставалась бы в неведении при отказе канала 1 и т.п.

7.3 Однокристальная (ИЛИ одноплатная) микроЭВМ, «создающая себя» при включении напряжения питания из имеющихся исправных блоков [9]

МикроЭВМ (рис. 7.3) содержит несколько одинаковых процессоров ЦП, ПЗУ и ОЗУ.

Процессор ЦП включает «собственный процессор» ЦП* и схему распространения стартового сигнала $S_{вх}$ (рис. 7.4). При подаче стартового сигнала $S_{вх}$ триггер D1 устанавливается в 1 и ЦП* начинает процедуру тестирования. Условие E = 1 соответствует подключению данного процессора к общей магистрали. Если за приемлемое время получен положительный результат тестирования, то сигнал Финиш равен 1, триггер D2 устанавливается в 1, элемент И закрывается, так что сигнал с выхода элемента задержки не сможет пройти к следующему процессору. Если же тестирование не завершилось успехом к моменту прохождения сигнала $S_{вх}$ через элемент задержки, то срабатывает элемент И, триггер D1 устанавливается в 0, отключая данный процессор от магистрали, а стартовый сигнал передается в следующий процессор.

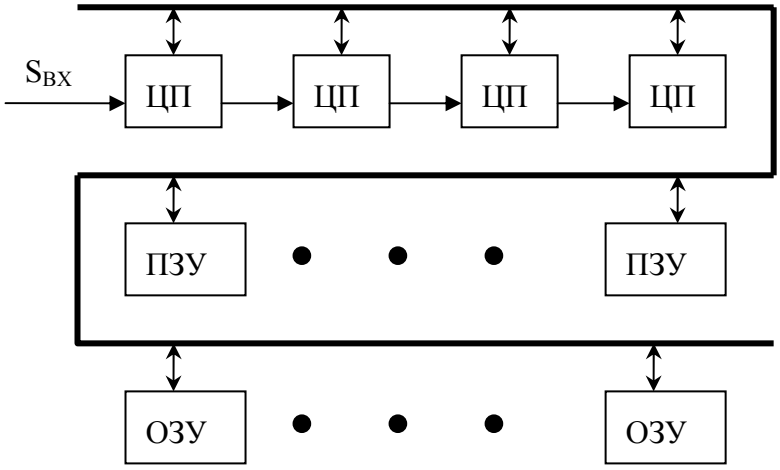


Рисунок 7.3 — Резервированная микроЭВМ

Порядок тестирования.

1. При включении напряжения питания микроЭВМ вырабатывается сигнал $S_{вх}$, поступающий на первый процессор. Этот процессор считывает содержимое первого ПЗУ и сравнивает полученную контрольную сумму с кодом, записанным в последней ячейке этого ПЗУ (действия по п. 1 выполняются под управлением микропрограммы, хранящейся в процессоре). Если коды не совпали, процессор переходит к проверке следующего ПЗУ, и т.д. Кодировка всех ПЗУ одинакова. Если работающее ПЗУ не найдено данным процессором, то запускается следующий процессор.

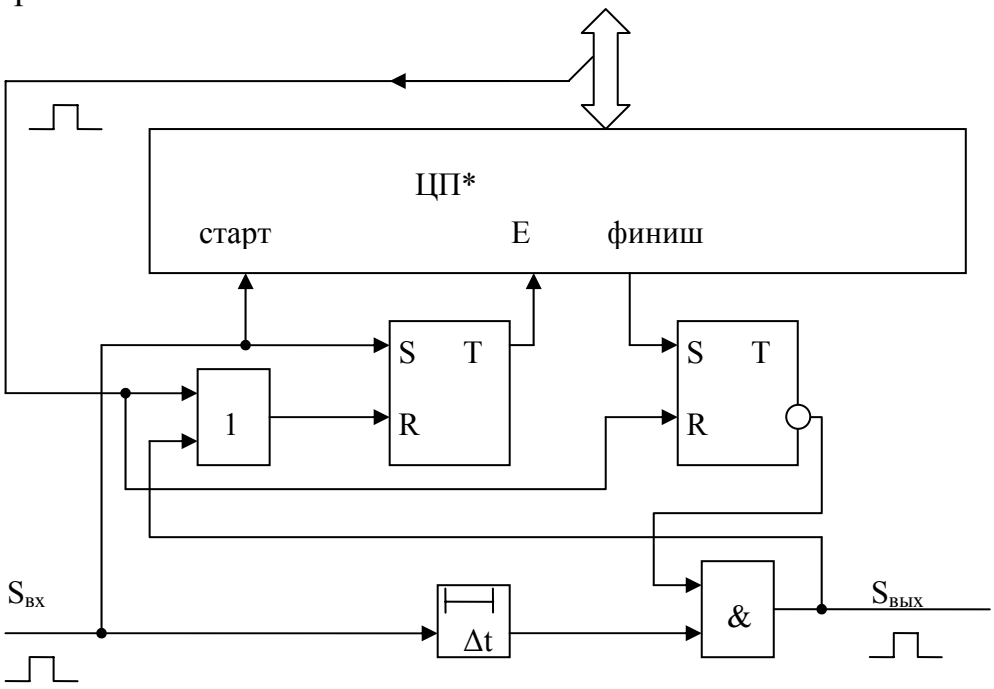


Рисунок 7.4 — Структура процессора ЦП резервированной микроЭВМ

2. Когда работающее ПЗУ найдено, процессор, работая по программе, хранящейся в этом ПЗУ, проверяет себя. Если тестирование процессора произведено удачно, то вырабатывается единичный сигнал Финиш и распространение стартового сигнала запрещается. В противном случае данный процессор отключается, и стартовый сигнал поступает на следующий процессор. Этот процесс продолжается до тех пор, пока не будет найдена работающая пара процессор — ПЗУ. Если этого не произошло, то микроЭВМ признается неисправной (в течение, например, 10 с на экране дисплея, подключенного к микроЭВМ, не появилось сообщение о готовности микроЭВМ к работе).

3. После того как найдена пара процессор — ПЗУ, производится тестирование блоков ОЗУ с запоминанием таблицы годности либо на регистрах процессора, либо в первом обнаруженном исправном ОЗУ. Тест ОЗУ хранится в ПЗУ.

В микроЭВМ используется принцип выделения страниц ОЗУ по требованию пользовательских программ. Каждой странице соответствует отдельный блок ОЗУ. Программам выделяются только исправные страницы, а неисправные рассматриваются операционной системой как недоступные для обращения. В микроЭВМ могут быть введены и другие устройства, например контроллеры, представленные в «множественном числе».

Таким образом, всякий раз при включении напряжения питания микроЭВМ «собирает себя» из отдельных блоков. Достаточным условием обеспечения работоспособности является наличие единичных исправных экземпляров всех входящих в нее блоков и исправность общих коммуникационных элементов.

7.4 Метод следящего самоконтроля микроЭВМ на основе предварительного прогнозирования вариантов ее «поведения» [9]

При использовании микроЭВМ в системе, работающей в реальном масштабе времени, не исключена возможность того, что в результате сбоя или отказа она начнет беспорядочное «блуждание» по памяти, рассматривая числовые массивы как цепочки команд, команды — как данные или адреса и т.д. Поскольку неправильное управление недопустимо для многих объектов, возникает задача быстрого обнаружения подобных «срывов».

Любую программу, записанную в машинных кодах в память микроЭВМ, можно представить в виде конечного числа линейных участков, связанных между собой командами переходов (рис. 7.5, *а*). В отсутствие внешних прерываний «поведение» микроЭВМ на каждом линейном участке строго детерминировано в том смысле, что при правильном его прохождении на управляющих линиях общей магистрали

будет сформирована строго определенная временная диаграмма, которую можно «вычислить» еще до начала прогона самой программы (рис. 7.5, *б*).

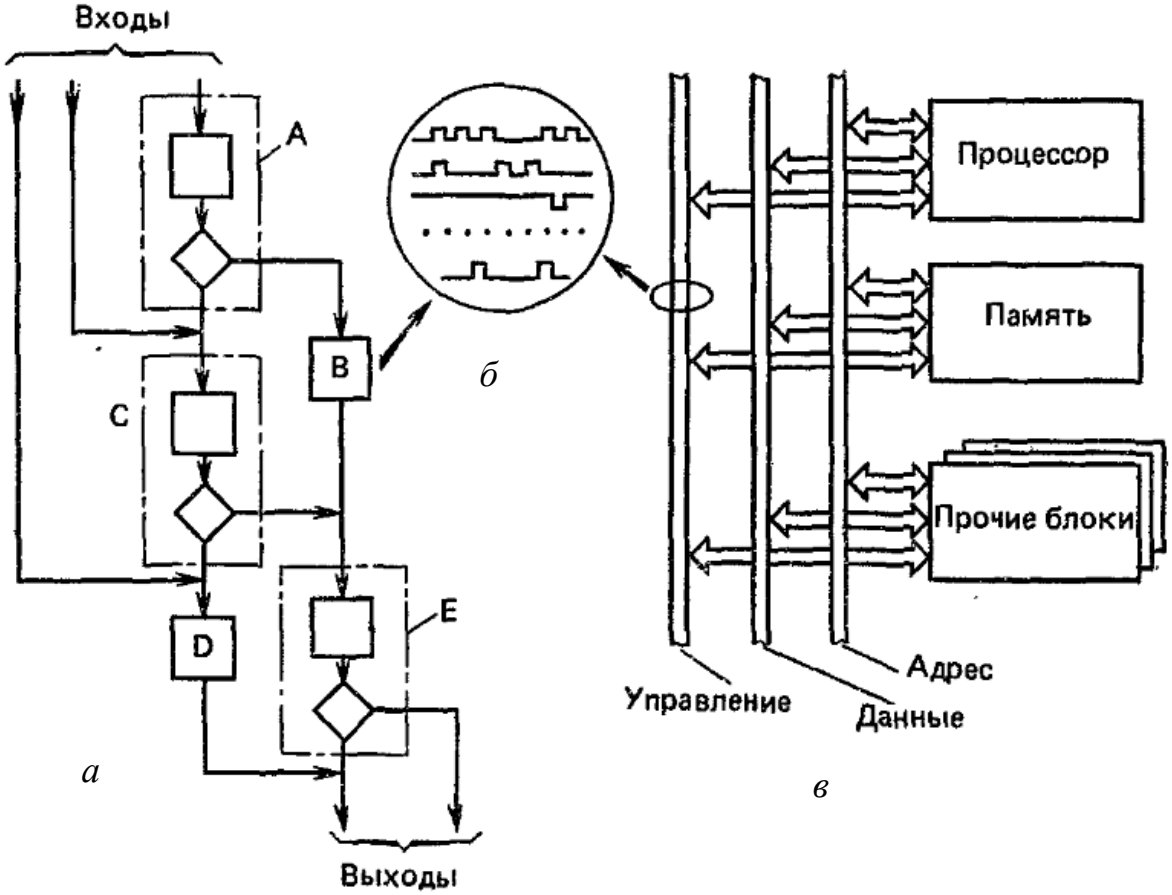


Рисунок 7.5 — Предсказание «поведения» микроЭВМ при работе по произвольной программе: *а* — фрагмент программы; *б* — «поведение» микроЭВМ на линейном участка В; *в* — структурная схема микроЭВМ

В простейшем случае подсчитывается ожидаемое число импульсов на каждой управляющей линии. Вычисления сводятся к суммированию констант, определяющих число импульсов на выбранной управляющей линии при выполнении каждой команды данного линейного участка. В результате предварительных подсчетов выясняется, например, что при правильном прохождении участка А программы на первой управляющей линии должно быть зарегистрировано 27 импульсов, на второй — 19 и т.д.

Предполагается также, что совокупность управляющих сигналов, которыми сопровождается выполнение команд условных переходов, не зависит от того, выполнено проверяемое условие или нет. Это предположение является общепринятым для многих микроЭВМ и в данном методе позволяет рассматривать команду условного перехода как завершающую команду линейного участка (см. участок А на рис. 7.5, *а*). Процесс прогнозирования поддается автоматизации, поскольку выделение в произвольной программе всех линейных участков производится путем анализа адресной информации, содержащейся в программе, без прогона самой программы.

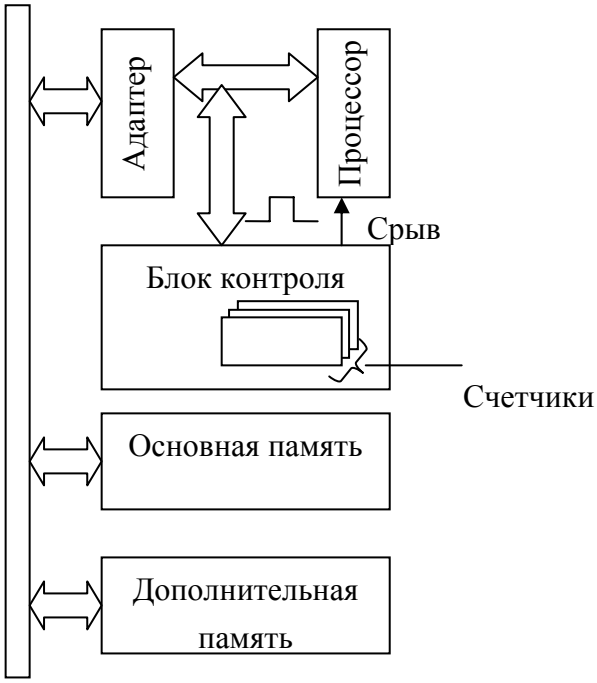


Рисунок 7.6 — Структурная схема микроЭВМ с системой самоконтроля

Необходимо знать лишь адреса внешних входов в эту программу. В результате прогнозирования формируется массив вспомогательной информации о программе. Формирование этого массива может производиться как на кроссистеме программирования, т.е. «вдали» от исполнительной микроЭВМ, так и на самой исполнительной микроЭВМ (после трансляции программы).

При работе системы самоконтроля (рис. 7.6) одновременно с выбором из основной памяти начальной команды некоторого линейного участка из дополнительной памяти (по отдельным шинам) извлекается служебная информация, которая помещается в аппаратные счетчики, содержимое их уменьшается по мере приближения к концу данного линейного участка. (Разрядность счетчиков может быть небольшой; в этом случае ведется счет по модулю.) При выходе на начало другого линейного участка счетчики проверяются на равенство нулю их содержимого, затем они вновь загружаются и т.д. Таким образом, ход выполнения любой, даже сильно разветвленной, программы постоянно контролируется. «Срыв» не будет зарегистрирован только в том случае, если он произошел точно в конце какого-либо линейного участка и микроЭВМ «ушла» точно на начало «чужого» линейного участка. Вероятность именно таких «точных» срывов мала. Время между возникновением сбоя или отказа и его обнаружением не превышает времени прохождения линейного участка максимальной длины и, поскольку последний может быть искусственно разбит на достаточно короткие звенья, может составлять, например, 100 мкс.

При обработке прерываний в число параметров возврата, «упрятываемых» в стек, входит содержимое счетчиков, которое

восстанавливается при возврате с учетом необходимой поправки на «лишние» импульсы, сформированные на управляющих линиях в связи с выполнением стандартных процедур упрятывания — восстановления.

Отметим, что пользователь «не подозревает» о существовании такого механизма самоконтроля, так как вся необходимая информация генерируется без его участия.

8 ЗАЩИТА ИНФОРМАЦИИ

Под защитой информации чаще всего понимают принятие мер для контроля за доступом к этой информации. Многие полагают, что достаточно закрыть паролями и сетевыми экранами те или иные части информационной системы и задача защиты информации уже решена. Действительно, криптографические и программные способы закрытия данных от постороннего вмешательства в наши дни уже достаточно совершенны, и злоумышленнику требуется потратить слишком много времени для взлома паролей и обхода проверочных алгоритмов сетевых экранов.

Однако в наши дни наиболее распространённым и удачным способом атаки систем информационной безопасности являются, как это ни странно, вовсе не компьютерные сети как таковые, а обычные телефонные линии. Методы социальной инженерии и знание психологических особенностей наёмных работников, умение находить подход и повод для беседы — всё это позволяет вскрыть многие защитные меры обеспечения информационной безопасности. Не последнее место занимают и средства телефонного прослушивания, приборы записи человеческой речи, жучки и устройства считывания по электросети.

8.1 Информация как объект защиты

Согласно законодательству РФ, информация — это сведения о лицах, предметах, фактах, событиях, явлениях и процессах независимо от формы их представления. Защите подлежит внутренняя, конфиденциальная и секретная информация. Внутренняя информация — информация о компании, которая еще не была опубликована (использование внутренней информации при заключении биржевых сделок считается незаконным). Информация конфиденциальная — служебная, профессиональная, промышленная, коммерческая или иная информация, правовой режим которой устанавливается ее собственником на основе законов о коммерческой, профессиональной тайне, государственной службе и других законодательных актов. Под коммерческой тайной предприятия понимаются не являющиеся государственным секретом сведения, связанные с производством, технологической информацией, управлением, финансами и другой деятельностью предприятия, разглашение (передача, утечка информации) которых может нанести вред

его интересам. К секретной информации относится информация, содержащая государственную тайну. Государственной тайной является информация, несанкционированное распространение которой может нанести ущерб интересам государственных органов, организациям, субъектам и стране в целом.

Информация как объект познания имеет следующие характеристики:

– нематериальность в смысле неизмеримости таких параметров, как масса, размеры, энергия, известными физическими методами и приборами;

– записанная на материальный носитель информация может храниться, обрабатываться, передаваться по различным каналам связи;

– любой материальный объект содержит информацию о себе или о другом объекте.

Без информации не может существовать жизнь в любой форме и не могут функционировать созданные человеком любые искусственные системы, без неё сама жизнь и всё созданное человеком представляет собой груду химических/физических элементов. Опыты по изоляции органов чувств человека, затрудняющей его информационный обмен с окружающей средой, показали, что информационный голод (дефицит информации) по своим последствиям не менее разрушителен, чем голод физический.

Защита информации определяется рядом свойств информации, основными из которых являются следующие:

1. Информация доступна человеку, если она содержится на материальном носителе. С помощью материальных средств можно защищать только материальный объект, т.о., объектом защиты информации являются материальные носители информации.

Носителями информации бывают:

– источники информации (чертёж — это источник, а бумага, на которой он нарисован, — носитель, однако бумага, без нанесённого на ней текста или рисунка является источником информации о её физических и химических характеристиках);

– переносчики информации;

– получатели информации.

Передача информации путём перемещения её носителей в пространстве связана с затратами энергии, причём величина затрат зависит от длины пути, параметров среды и типа носителя.

2. Ценность информации оценивается степенью полезности её для пользователя (собственника, владельца, получателя). Полезность информации всегда конкретна — нет ценной информации вообще — информация полезна или вредна для конкретного её пользователя, поэтому при защите информации прежде всего определяют круг субъектов (государств, фирм, групп лиц, людей), заинтересованных в защищаемой информации, так как вероятно, что среди них окажутся злоумышленники.

В интересах защиты информации её владелец наносит на носитель информации условный знак полезности содержащейся на нём информации

— гриф секретности или конфиденциальности. В качестве критерия для определения грифа конфиденциальности информации могут служить результаты прогноза последствий попадания информации к конкуренту или злоумышленнику:

– величина экономического и морального ущерба, наносимого организации;

– реальность создания предпосылок для катастрофических последствий в деятельности организации (банкротства и т.п.).

3. Так как информация для получателя может быть полезной или вредной, то информацию можно рассматривать как товар.

Цена информации как любого товара складывается из себестоимости и прибавочной стоимости (прибыли).

Себестоимость определяется расходами владельца информации на её получение путём:

– исследований в лабораториях, аналитических центрах, группах и т.п.;

– покупки информации;

– добычи информации противоправными действиями.

Прибыль от информации может быть получена в результате следующих действий:

– продажи информации на рынке;

– материализации информации в продукции с новыми свойствами или в технологиях, приносящих прибыль;

– использования информации для принятия эффективных решений (экономия средств, ресурсов и т.п.).

4. Ценность информации изменяется во времени. Распространение информации и её использование приводят к изменению её ценности и цены. Характер изменения ценности от времени зависит от вида информации.

5. Невозможно объективно (без учёта полезности её для потребителя, владельца, собственника) оценить количество информации.

Иногда полезность информации связывают с её качеством, но понятие «качество» применительно к информации не имеет самостоятельного значения, т.к. оно поглощается понятием «количество». Количество информации зависит от её качества: чем более качественная фотография, тем больше оттенков и полутонов она содержит, тем меньше на ней помех. Под качеством информации подразумевают качество отображения её на носителе или её достоверность (соответсвие оригиналу).

6. При копировании, не изменяющем информационные параметры носителя, количество информации не меняется, а цена снижается.

Все перечисленные свойства информации являются важными составляющими для формирования политики информационной безопасности организации, государства, группы лиц — деятельности

любых субъектов информационного пространства и информационных систем.

8.2 Потенциальные угрозы безопасности

В литературе, посвященной вопросам защиты информации, можно найти различные варианты моделей угроз безопасности информации. Это объясняется стремлением более точно описать многообразные ситуации воздействия на информацию и определить наиболее адекватные меры парирования. В принципе, можно пользоваться любой понравившейся моделью, необходимо только убедиться, что она описывает максимально большое число факторов, влияющих на безопасность информации. Но, прежде всего, надо помнить, что пользователю, то есть потребителю информации и информационных услуг, оказываемых корпоративной сетью, глубоко без разницы, получит ли он информацию не вовремя, получит ее в искаженном виде или вообще потеряет по вине неправильной работы технических средств, пожара в серверном зале или за счет действий злоумышленника. Итог для него во всех случаях одинаков — понесенные убытки (моральные или материальные).

Что же такое угроза безопасности информации? Это — действие, направленное против объекта защиты, проявляющееся в опасности искажений и потерь информации. Надо оговориться, что речь идет не о всей информации, а только о той ее части, которая, по мнению ее собственника (пользователя), имеет коммерческую ценность (информация как товар) или подлежит защите в силу закона (конфиденциальная информация).

Необходимо также учитывать, что источники угроз безопасности могут находиться как внутри фирмы — внутренние источники, так и вне ее — внешние источники. Такое деление оправдано потому, что для одной и той же угрозы (например, кража) методы парирования для внешних и внутренних источников будут разными.

При составлении модели угроз авторы использовали различные широко применяемые в настоящее время варианты моделей, разработанные специалистами в области защиты информации государственных и негосударственных научных учреждений, и собственный опыт работы в этой области. Исходя из проведенного анализа все источники угроз безопасности информации, циркулирующей в корпоративной сети, можно разделить на три основные группы:

I. Угрозы, обусловленные действиями субъекта (антропогенные угрозы).

II. Угрозы, обусловленные техническими средствами (техногенные угрозы).

III. Угрозы, обусловленные стихийными источниками.

Первая группа наиболее обширна и представляет наибольший интерес с точки зрения организации парирования этим угрозам, так как

действия субъекта всегда можно оценить, спрогнозировать и принять адекватные меры. Методы противодействия этим угрозам управляемы и напрямую зависят от воли организаторов защиты информации.

Субъекты, действия которых могут привести к нарушению безопасности информации, могут быть как внешние:

1) криминальные структуры;
2) рецидивисты и потенциальные преступники:
   - недобросовестные партнеры;
   - конкуренты;
   - политические противники;

так и внутренние:

- персонал учреждения;
- персонал филиалов;
- лица с нарушенной психикой;
- специально внедренные агенты.

Основываясь на результатах международного и российского опыта, действия субъектов могут привести к ряду нежелательных последствий, среди которых применительно к корпоративной сети, можно выделить следующие.

1. Кража
   а) технических средств (винчестеров, ноутбуков, системных блоков);
   б) носителей информации (бумажных, магнитных, оптических и пр.);
   в) информации (чтение и несанкционированное копирование);
   г) средств доступа (ключи, пароли, ключевая документация и пр.).
2. Подмена (модификация)
   а) операционных систем;
   б) систем управления базами данных;
   в) прикладных программ;
   г) информации (данных), отрицание факта отправки сообщений;
   д) паролей и правил доступа.
3. Уничтожение (разрушение)
   а) технических средств (винчестеров, ноутбуков, системных блоков);
   б) носителей информации (бумажных, магнитных, оптических и пр.);
   в) программного обеспечения (ОС, СУБД, прикладного ПО);
   г) информации (файлов, данных);
   д) паролей и ключевой информации.
4. Нарушение нормальной работы (прерывание)
   а) скорости обработки информации;
   б) пропускной способности каналов связи;
   в) объемов свободной оперативной памяти;

г) объемов свободного дискового пространства;

д) электропитания технических средств;

5. Ошибки

а) при инсталляции ПО, ОС, СУБД;

б) при написании прикладного ПО;

в) при эксплуатации ПО;

г) при эксплуатации технических средств.

6. Перехват информации (несанкционированный)

а) за счет ПЭМИ от технических средств;

б) за счет наводок по линиям электропитания;

в) за счет наводок по посторонним проводникам;

г) по акустическому каналу от средств вывода;

д) по акустическому каналу при обсуждении вопросов;

е) при подключении к каналам передачи информации;

ж) за счет нарушения установленных правил доступа (взлом).

Вторая группа содержит угрозы менее прогнозируемые, напрямую зависящие от свойств техники и поэтому требующие особого внимания. Технические средства, содержащие потенциальные угрозы безопасности информации, так же могут быть внутренними:

- некачественные технические средства обработки информации;
- некачественные программные средства обработки информации;
- вспомогательные средства (охраны, сигнализации, телефонии);
- другие технические средства, применяемые в учреждении;

и внешними:

- средства связи;
- близко расположенные опасные производства;
- сети инженерных коммуникаций (энерго-, водоснабжения, канализации);
- транспорт.

Последствиями применения таких технических средств, напрямую влияющими на безопасность информации, могут быть:

1. Нарушение нормальной работы:

а) нарушение работоспособности системы обработки информации;

б) нарушение работоспособности связи и телекоммуникаций;

в) старение носителей информации и средств ее обработки;

г) нарушение установленных правил доступа;

д) электромагнитное воздействие на технические средства.

2. Уничтожение (разрушение)

а) программного обеспечения (ОС, СУБД);

б) средств обработки информации (броски напряжений, протечки);

в) помещений

г) информации (размагничивание, радиация, протечки и пр.);

д) персонала.

3. Модификация (изменение)

а) программного обеспечения. ОС, СУБД;

б) информации при передаче по каналам связи и телекоммуникациям.

Третью группу составляют угрозы, которые совершенно не поддаются прогнозированию, и поэтому меры их парирования должны применяться всегда. Стихийные источники, составляющие потенциальные угрозы информационной безопасности, как правило, являются внешними по отношению к рассматриваемому объекту, и под ними понимаются, прежде всего, природные катаклизмы:

- пожары;
- землетрясения;
- наводнения;
- ураганы;
- другие форс-мажорные обстоятельства;
- различные непредвиденные обстоятельства;
- необъяснимые явления.

Эти природные и необъяснимые явления так же влияют на информационную безопасность, опасны для всех элементов корпоративной сети и могут привести к следующим последствиям:

1. Уничтожению (разрушению)

а) технических средств обработки информации;

б) носителей информации;

в) программного обеспечения (ОС, СУБД, прикладного ПО);

г) информации (файлов, данных);

д) помещений;

е) персонала.

2. Исчезновению (пропаже)

а) информации в средствах обработки;

б) информации при передаче по телекоммуникационным каналам;

в) носителей информации;

г) персонала.

Даже первичный анализ приведенного перечня угроз безопасности информации показывает, что для обеспечения комплексной безопасности необходимо принятие как организационных, так и технических решений парирования. Такой подход позволяет дифференцировано подойти к распределению материальных ресурсов, выделенных на обеспечение информационной безопасности.

Необходимо отметить, что оценить весовые коэффициенты каждой угрозы достаточно затруднительно из-за высокой латентности их проявлений и отсутствия вразумительной статистики по этому вопросу. Поэтому в современной литературе можно найти различные шкалы оценок. Вместе с тем на основе анализа, проводимого различными специалистами в области компьютерных преступлений, и собственных

наблюдений по частоте проявления угрозы безопасности можно расставить так:

- кража (копирование) программного обеспечения;
- подмена (несанкционированный ввод) информации;
- уничтожение (разрушение) данных на носителях информации;
- нарушение нормальной работы (прерывание) в результате вирусных атак;
- модификация (изменение) данных на носителях информации;
- перехват (несанкционированный съем) информации;
- кража (несанкционированное копирование) ресурсов;
- нарушение нормальной работы (перегрузка) каналов связи;
- непредсказуемые потери.

8.3 Методы защиты информации

На сегодняшний день нормой обработки информации в корпоративных приложениях становится работа пользователя на одном и том же компьютере как с открытой, так и с конфиденциальной информацией, требующей для обработки различной номенклатуры ресурсов. Это существенно расширяет потенциальную возможность хищения (нарушения конфиденциальности) данных санкционированным пользователем (пользователем, допущенным к работе на вычислительном средстве в рамках выполнения своих служебных обязанностей — инсайдером) и несанкционированной модификацией (нарушением доступности и целостности) конфиденциальных данных в результате обработки на том же компьютере открытой информации.

Из теории защиты информации известно, что эффективная защита может быть построена только на основе реализации разграничительной политики доступа к ресурсам (механизмы контроля, в частности контентного контроля, по очевидным причинам могут использоваться лишь как вспомогательные). Однако как в существующих требованиях к средству защиты, так в известных практических реализациях — в ОС и в приложениях применение разграничительной политики предполагается для разграничения доступа различных пользователей, допущенных к обработке информации на компьютере, к ресурсам. При реализации же защиты информации от инсайдерских атак задача реализации разграничительной политики доступа к ресурсам уже иная, собственно, в своей постановке — необходимо разграничивать режимы обработки различных категорий информации на одном компьютере для одного и того же пользователя (а не доступ к ресурсам между различными пользователями). Правильнее здесь уже говорить не о разграничительной, а о разделительной политике доступа к ресурсам. Как следствие, необходимы совершенно иные подходы к решению задачи защиты, а механизмами защиты должны выполняться совсем иные требования.

Рассмотрим в данной работе возможный апробированный подход к решению задачи защиты и сформулируем требования, реализация которых необходима для эффективного решения задачи защиты от инсайдерских атак.

Поскольку на одном компьютере обрабатывается информация различных уровней конфиденциальности, и при этом обработка информации различных уровней конфиденциальности требует различных ресурсов (различные приложения, файловые объекты, устройства, сетевые ресурсы, и т.д., и т.п.), причем, как правило, чем ниже уровень конфиденциальности обрабатываемой информации, тем шире номенклатура ресурсов может использоваться при ее обработке (что и составляет потенциальную угрозу хищения конфиденциальной информации), задача защиты информации состоит в формировании и изолировании режимов обработки информации различных уровней конфиденциальности.

Формирование режимов обработки категорированной информации состоит в подключении соответствующего набора ресурсов при обработке информации каждого уровня конфиденциальности.

Изолирование режимов обработки категорированной информации состоит в противодействии любой возможности изменения режима обработки (санкционированного набора ресурсов) информации каждого уровня конфиденциальности.

При реализации подобного подхода уже нечего контролировать, т.к. предотвращается сама возможность хищения конфиденциальной информации за счет использования несанкционированных (не используемых для ее обработки) ресурсов.

Решая задачу защиты информации, необходимо учитывать, что в общем случае защита состоит не только в противодействии хищению конфиденциальной информации (нарушению конфиденциальности информации), но и в обеспечении ее доступности и целостности.

Рассматриваемый в работе подход к реализации защиты от инсайдерских атак, состоящий в реализации разделительной политики доступа к ресурсам, проиллюстрирован на рис. 8.1.

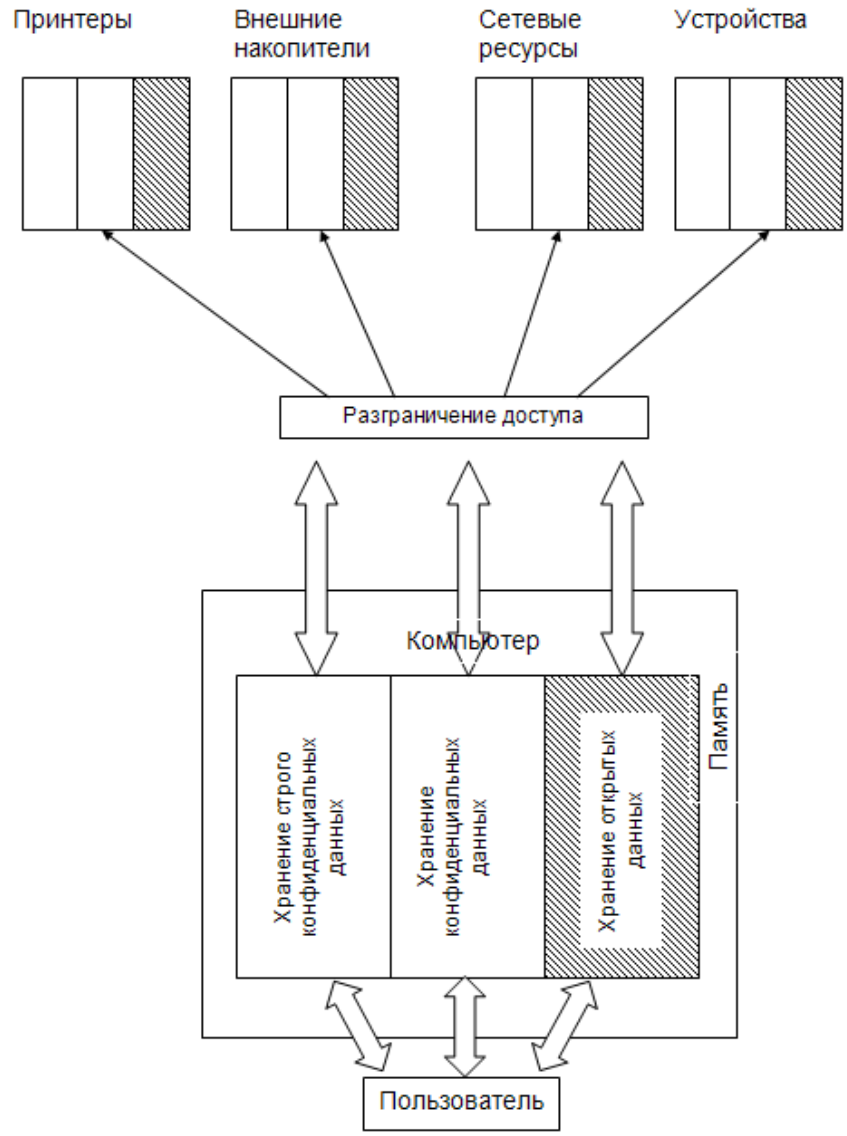


Рисунок 8.1 — Иллюстрация рассматриваемого подхода к реализации защиты от инсайдерских атак

Принципы криптографической защиты информации

Криптография представляет собой совокупность методов преобразования данных, направленных на то, чтобы сделать эти данные бесполезными для противника. Такие преобразования позволяют решить две главные проблемы защиты данных: проблему конфиденциальности (путем лишения противника возможности извлечь информацию из канала связи) и проблему целостности (путем лишения противника возможности изменить сообщение так, чтобы изменился его смысл, или ввести ложную информацию в канал связи).

Проблемы конфиденциальности и целостности информации тесно связаны между собой, поэтому методы решения одной из них часто применимы для решения другой.

Обобщенная схема криптографической системы, обеспечивающей шифрование передаваемой информации, показана на рис. 8.2. Отправитель генерирует открытый текст исходного сообщения M, которое должно быть

передано законному получателю по незащищенному каналу. За каналом следит перехватчик с целью перехватить и раскрыть передаваемое сообщение. Для того чтобы перехватчик не смог узнать содержание сообщения M, отправитель шифрует его с помощью обратимого преобразования $E_к$ и получает шифртекст (или криптограмму) $C = E_к(M)$, который отправляет получателю.

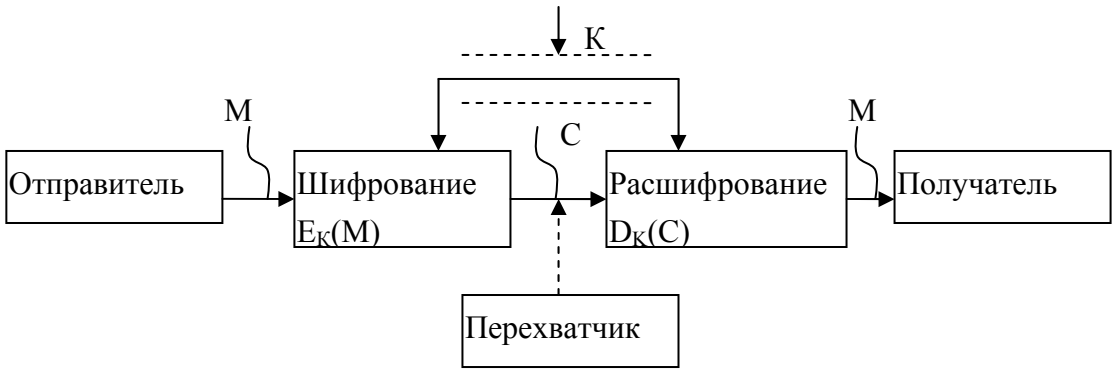


Рисунок 8.2 — Обобщенная схема криптосистемы

Законный получатель, приняв шифртекст C, расшифровывает его с помощью обратного преобразования $D = E_K^{-1}$ и получает исходное сообщение в виде открытого текста M:

$$D_k(C) = E_k^{-1}(E_k(M)) = M.$$

Преобразование $E_к$ выбирается из семейства криптографических преобразований, называемых криптоалгоритмами. Параметр, с помощью которого выбирается отдельное используемое преобразование, называется криптографическим ключом K. Криптосистема имеет разные варианты реализации: набор инструкций, аппаратные средства, комплекс программ компьютера, которые позволяют зашифровать открытый текст и расшифровать шифртекст различными способами, один из которых выбирается с помощью конкретного ключа K.

Говоря более формально, криптографическая система — это однопараметрическое семейство $(E_к)_{к \in \overline{к}}$ обратимых преобразований

$$E_K: \overline{M} \to C$$

из пространства $\overline{M}$ сообщений открытого текста в пространство $\overline{C}$ шифрованных текстов. Параметр K (ключ) выбирается из конечного множества $\overline{K}$, называемого пространством ключей.

Вообще говоря, преобразование шифрования может быть симметричным или асимметричным относительно преобразования расшифрования. Это важное свойство функции преобразования определяет два класса криптосистем;

– симметричные (одноключевые) криптосистемы;

– асимметричные (двухключевые) криптосистемы (с открытым ключом).

Схема симметричной криптосистемы с одним секретным ключом была показана на рис. 8.2. В ней используются одинаковые секретные ключи в блоке шифрования и блоке расшифрования.

Обобщенная схема асимметричной криптосистемы с двумя разными ключами $К_1$, и $К_2$ показана на рис.8.3. В этой криптосистеме один из ключей является открытым, а другой — секретным.

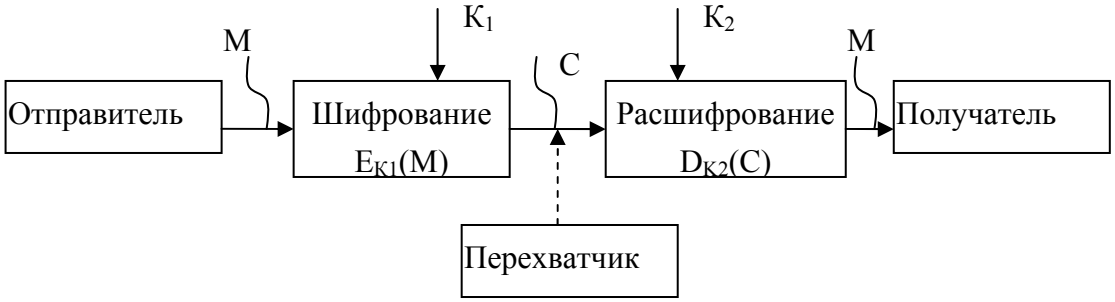


Рисунок 8.3 — Обобщенная схема асимметричной криптосистемы с открытым ключом

В симметричной криптосистеме секретный ключ надо передавать отправителю и получателю по защищенному каналу распространения ключей, например такому, как курьерская служба. На рис. 8.2 этот канал показан «экранированной» линией. Существуют и другие способы распределения секретных ключей. В асимметричной криптосистеме передают по незащищенному каналу только открытый ключ, а секретный ключ сохраняют на месте его генерации.

На рис. 8.4 показан поток информации в криптосистеме в случае активных действий перехватчика. Активный перехватчик не только считывает все шифртексты, передаваемые по каналу, но может также пытаться изменять их по своему усмотрению.

Любая попытка со стороны перехватчика расшифровать шифртекст С для получения открытого текста М или зашифровать свой собственный текст М' для получения правдоподобного шифртекста С, не имея подлинного ключа, называется криптоаналитической атакой.

Если предпринятые криптоаналитические атаки не достигают поставленной цели и криптоаналитик не может, не имея подлинного ключа, вывести М из С или С из М', то полагают, что такая криптосистема является криптостойкой.

Криптоанализ — это наука о раскрытии исходного текста зашифрованного сообщения без доступа к ключу. Успешный анализ может раскрыть исходный текст или ключ. Он позволяет также обнаружить слабые места в криптосистеме, что, в конечном счете, ведет к тем же результатам.

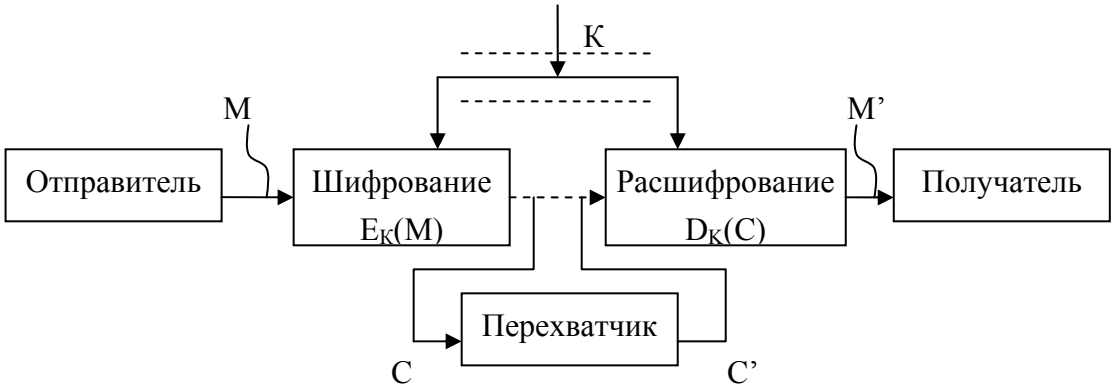


Рисунок 8.4 — Поток информации в криптосистеме при активном перехвате сообщений

Фундаментальное правило криптоанализа, впервые сформулированное голландцем А. Керкхоффом еще в XIX веке, заключается в том, что стойкость шифра (криптосистемы) должна определяться только секретностью ключа. Иными словами, правило Керкхоффа состоит в том, что весь алгоритм шифрования, кроме значения секретного ключа, известен криптоаналитику противника. Это обусловлено тем, что криптосистема, реализующая семейство криптографических преобразований, обычно рассматривается как открытая система. Такой подход отражает очень важный принцип технологии защиты информации: защищенность системы не должна зависеть от секретности чего-либо такого, что невозможно быстро изменить в случае утечки секретной информации. Обычно криптосистема представляет собой совокупность аппаратных и программных средств, которую можно изменить только при значительных затратах времени и средств, тогда как ключ является легко изменяемым объектом. Именно поэтому стойкость криптосистемы определяется только секретностью ключа.

Другое почти общепринятое допущение в криптоанализе состоит в том, что криптоаналитик имеет в своем распоряжении шифртексты сообщений.

Существует четыре основных типа криптоаналитических атак. Конечно, все они формулируются в предположении, что криптоаналитику известны применяемый алгоритм шифрования и шифртексты сообщений. Перечислим эти криптоаналитические атаки.

1. Криптоаналитическая атака при наличии только известного шифртекста. Криптоаналитик имеет только шифртексты C_1, C_2,...,C_i нескольких сообщений, причем все они зашифрованы с использованием одного и того же алгоритма шифрования $E_к$. Работа криптоаналитика заключается в том, чтобы раскрыть исходные тексты M_1, M_2,..., M_i по возможности большинства сообщений или, еще лучше, вычислить ключ К,

использованный для шифрования этих сообщений, с тем чтобы расшифровать и другие сообщения, зашифрованные этим ключом.

2. Криптоаналитическая атака при наличии известного открытого текста. Криптоаналитик имеет доступ не только к шифртекстам C_1, C_2,...,C_i нескольких сообщений, но также к открытым текстам M_1, M_2,..., M_i этих сообщений. Его работа заключается в нахождении ключа К, используемого при шифровании этих сообщений, или алгоритма расшифрования D_K любых новых сообщений, зашифрованных тем же самым ключом.

3. Криптоаналитическая атака при возможности выбора открытого текста. Криптоаналитик не только имеет доступ к шифртекстам C_1, C_2,...,C_i и связанным с ними открытым текстам M_1, M_2,..., M_i нескольких сообщений, но и может по желанию выбирать открытые тексты, которые затем получает в зашифрованном виде. Такой криптоанализ получается более мощным по сравнению с криптоанализом с известным открытым текстом, потому что криптоаналитик может выбрать для шифрования такие блоки открытого текста, которые дадут больше информации о ключе. Работа криптоаналитика состоит в поиске ключа К, использованного для шифрования сообщений, или алгоритма расшифрования D_K новых сообщений, зашифрованных тем же ключом.

4. Криптоаналитическая атака с адаптивным выбором открытого текста. Это особый вариант атаки с выбором открытого текста. Криптоаналитик может не только выбирать открытый текст, который затем шифруется, но и изменять свой выбор в зависимости от результатов предыдущего шифрования. При криптоанализе с простым выбором открытого текста криптоаналитик обычно может выбирать несколько крупных блоков открытого текста для их шифрования; при криптоанализе с адаптивным выбором открытого текста он имеет возможность выбрать сначала более мелкий пробный блок открытого текста, затем выбрать следующий блок в зависимости от результатов первого выбора и т.д. Эта атака предоставляет криптоаналитику еще больше возможностей, чем предыдущие типы атак.

Кроме перечисленных основных типов криптоаналитических атак, можно отметить, по крайней мере, еще два типа.

5. Криптоаналитическая атака с использованием выбранного шифртекста. Криптоаналитик может выбирать для расшифрования различные шифртексты C_1, C_2,...,C_i и имеет доступ к расшифрованным открытым текстам M_1, M_2,..., M_i. Например, криптоаналитик получил доступ к защищенному от несанкционированного вскрытия блоку, который выполняет автоматическое расшифрование. Работа криптоаналитика заключается в нахождении ключа. Этот тип криптоанализа представляет особый интерес для раскрытия алгоритмов с открытым ключом.

6. Криптоаналитическая атака методом полного перебора всех возможных ключей. Эта атака предполагает использование криптоаналитиком известного шифртекста и осуществляется посредством

полного перебора всех возможных ключей с проверкой, является ли осмысленным получающийся открытый текст. Такой подход требует привлечения предельных вычислительных ресурсов и иногда называется силовой атакой.

Существуют и другие, менее распространенные, криптоаналитические атаки.

Организационные мероприятия по защите информации

Инженерно-техническая защита информации на объекте достигается выполнением комплекса организационно-техничес-ких и технических мероприятий с применением (при необходимости) средств защиты информации от утечки информации или несанкционированного воздействия на нее по техническим каналам, за счет несанкционированного доступа и неконтролируемого распространения информации, по предупреждению преднамеренных программно-технических воздействий с целью нарушения целостности (модификации, уничтожения) информации в процессе ее обработки, передачи и хранения, нарушения ее доступности и работоспособности технических средств и носителей информации и т.п.

Мероприятия по защите конфиденциальной информации и противодействию техническим средствам разведки подразделяются на организационно-технические и технические (рис. 8.5).

Защита информации техническими способами и средствами

Защита информации может осуществляться инженерно-техническими и криптографическими способами.

Техническая защита конфиденциальной информации — защита информации некриптографическими методами, направленными на предотвращение утечки защищаемой информации по техническим каналам, от несанкционированного доступа к ней и от специальных воздействий на информацию в целях ее уничтожения, искажения или блокирования.

Порядок защиты некриптографическими способами и средствами определен руководящими документами Гостехкомиссии России.

Рисунок 8.5 — Возможные мероприятия по защите информации и противодействию техническим средствам разведки

Порядок разработки, производства, реализации и использования средств криптографической защиты информации с ограниченным доступом, не содержащей сведений, составляющих государственную тайну, определяется Положением ПКЗ-99, утвержденным приказом ФАПСИ от 23 сентября 1999 г., а также Инструкцией об организации и обеспечении безопасности хранения, обработки и передачи по каналам связи с использованием средств криптографической защиты информации с ограниченным доступом, не содержащей сведений, составляющих государственную тайну, утвержденной приказом ФАПСИ от 13 июня 2001 г. № 152.

Технические мероприятия по защите информации и противодействию техническим средствам разведки (ТСР) основаны на применении средств защиты и реализации проектных и конструкторских решений, направленных на защиту объекта.

Способы противодействия техническим средствам разведки определяют как преднамеренное воздействие на технический канал утечки информации для достижения цели по противодействию техническим разведкам, а способы защиты объекта — преднамеренное воздействие на объект защиты для достижения поставленных целей противодействия.

Для защиты объекта возможно использование таких способов, как скрытие и техническая дезинформация.

Скрытие — это способ защиты объекта от технической разведки путем устранения или ослабления технического демаскирующего признака объекта защиты.

Техническая дезинформация — способ защиты объекта путем искажения технических демаскирующих признаков объекта защиты или имитации технических демаскирующих признаков объекта, не являющегося объектом защиты.

Это, как правило, объект, на который в соответствии с легендой должен походить настоящий объект защиты.

Легендирование — это один из способов противодействия техническим разведкам, заключающийся в преднамеренном распространении ложной информации о предназначении объекта и характере выполняемых на нем работ.

В качестве средств защиты и противодействия рассматривается аппаратура и технические устройства (в том числе на различных носителях — автомобиль, корабль, стационарное помещение и т.п.), используемые для защиты объекта.

Средства защиты и противодействия могут быть активными и пассивными.

Активное средство противодействия — это средство, обеспечивающее создание маскирующих или имитирующих активных помех средством технической разведки или средство, приводящее к нарушению нормального функционирования этих средств разведки.

Пассивное же средство противодействия — это средство инженерно-технической защиты информации (ИТЗИ), обеспечивающее скрытие объекта защиты от технических разведок путем поглощения, отражения или рассеивания его излучений.

Комплексное противодействие обеспечивается при комплексном использовании средств защиты и организационно-технических способов и методов в целях защиты охраняемых сведений об объекте и осуществляемое согласованно с целями и задачами защиты информации и противодействия ТСР, этапами жизненного цикла объекта и способами противодействия.

Защита должна производиться активно, убедительно, разнообразно, непрерывно, комплексно, планово.

Активность противодействия состоит в настойчивом осуществлении эффективных мер противодействия.

Разнообразие противодействия исключает шаблон в организации и проведении мероприятий по противодействию.

Комплексность предусматривает системный подход, т.е. равнозначное закрытие всех возможных каналов утечки информации об объекте. Недопустимо применять отдельные технические средства или методы, направленные только на защиту отдельных из общего числа возможных каналов утечки информации.

Непрерывность противодействия предусматривает проведение подобных мероприятий на всех этапах жизненного цикла разработки и существования специальной продукции или обеспечения производственной деятельности объекта защиты.

Важно также, чтобы мероприятия по защите и противодействию выглядели правдоподобно и отвечали условиям обстановки, выполнялись в соответствии с планами защиты информации объекта. В связи с этим разрабатываются и осуществляются практические меры защиты. При этом особое внимание обращается на выбор замысла защиты информации объекта, замысла противодействия. Замысел защиты — общая идея и основное содержание организационных, организационно-технических и технических мероприятий, обеспечивающих устранение или ослабление (искажение) демаскирующих признаков и закрытие технических каналов утечки охраняемых сведений и несанкционированного воздействия на них.

Защита информации от утечки за счет побочного электромагнитного излучения и наводок

Одним из наиболее вероятных каналов утечки информации в компьютерных сетях считается побочное электромагнитное излучение и наводки (ПЭМИН), создаваемые техническими средствами, то есть персональными компьютерами и линиями связи, так как ПЭМИН-канал способен переносить сигнал на расстояния в десятки и сотни метров. Наиболее мощными источниками ПЭМИН в ПК являются дисплеи, дисководы и принтеры. Протяженные проводные линии связи также создают сильные электромагнитные поля в окружающем пространстве, в особенности современные высокоскоростные компьютерные сети. Несмотря на многочисленные исследования по разработке и усовершенствованию оборудования для уменьшения ПЭМИН от линий связи, защита от перехвата данных в проводах LAN (Local area Network) остается большой проблемой.

Наличие сигналов ПЭМИН за пределами территории учреждения (ТУ) дает возможность перехватить их, тем самым создает канал утечки данных. ТУ в данном случае называется территория, контролируемая учреждением, на которой исключено нахождение посторонних лиц и

специальной подслушивающей аппаратуры. Средой распространения сигналов являются не только воздух и телефонные линии связи, но и линии электропитания, пожарной сигнализации, а также коммуникации (в том числе трубы водоснабжения и отопления), выходящие за ТУ.

Вероятность утечки информации по ПЭМИН-каналу зависит от таких факторов:

– как мощность источника излучения;

– характеристики среды распространения (ослабление и искажение сигнала в среде);

– характеристики принимающей аппаратуры.

Основной характеристикой канала, зависящей от вышеперечисленных факторов, является отношение информативный сигнал/шум на входе приемника. Чем ближе приемник расположен к источнику сигнала, тем больше это отношение, тем более вероятно правильное распознавание сигнала.

При разработке системы мер по защите информации следует исходить из предположения, что принимающая аппаратура может быть установлена в любой точке вне ТУ, вплоть до ее границ.

Следовательно, если ПЭМИН-сигнал может быть принят вне ТУ, то считают, что данный канал утечки информации существует.

Проблема ПЭМИН для LAN чрезвычайно важна. В современных компьютерных сетях наиболее распространены кабели на основе витых пар. Они разработаны для передачи сигнала в «симметричном режиме», при этом предполагается, что токи, текущие по проводам витой пары навстречу друг другу, равны по величине. В таком случае побочное излучение отсутствует. Из-за недостижимости идеального режима на практике, ввиду невозможности произвести «идеальный» кабель и нагрузок, создаваемых при работе активного оборудования, в кабеле всегда присутствует неравенство токов. То есть имеет место «несимметричный режим». Излучения обоих проводов складываются, в результате происходит значительное излучение от витой пары.

Качество экранирования и соединения экрана кабеля с оборудованием сильно влияют на мощность излучения передаваемых данных. Также не менее важным является качество заземления активного и пассивного оборудования, поскольку даже очень хорошо экранированный кабель не обеспечивает безопасность передачи, если экран не соединен должным образом. Таким образом, наряду с качеством экранирования кабелей важным является правильное проведение монтажных работ с учетом всех требований, иначе нельзя гарантировать надежность линии.

Безусловно, наиболее защищенными от несанкционированного съема информации через ПЭМИН-канал являются волоконно-оптические сети. Для LAN, использующих медные кабели, следует отдавать предпочтение экранированным системам. Так как согласно результатам исследования швейцарской компании Reichle & De-Massari AG неэкранированные кабельные системы обеспечивают необходимый

уровень безопасности только для низкоскоростных приложений. Кроме того, чрезвычайно важным является качество монтажных работ по соединению экранов и заземлению оборудования.

Существуют также устройства для защиты помещений от утечки информации и предотвращения съема информации с персональных компьютеров и локальных вычислительных сетей на базе ПК, маскировки средств вычислительной техники.

Защита информации от случайных воздействий

Рассмотрим некоторые приемы и методы, связанные с защитой информации от случайных ошибок или некомпетентности пользователей, а также от сбоев аппаратуры, в частности из-за помех в электросети, то есть причин возможной потери информации, не связанных с несанкционированным доступом и происками злоумышленников. Потеря файлов, а также крах системы вполне возможны и без внешних, корыстных помыслов. В связи с этим во всех операционных системах предусматриваются простейшие средства профилактики. Так, при удалении файлов, как правило, требуется дополнительное подтверждение, а удаленный файл, как правило, при необходимости может быть восстановлен, поскольку определенное время он хранится в специальном буфере («корзина для мусора»).

Для того чтобы обезопасить себя от неприятных последствий (связанных с вышеперечисленными инцидентами), приводящих к потере данных на сервере или рабочих станциях, которые могут представлять большую ценность, так как являются результатом больших трудовых затрат, необходимо выполнение определенных мероприятий. Существует три основных способа защиты от таких воздействий — резервное копирование данных, избыточное дублирование и установка специализированных устройств защиты от нарушений в системе электропитания.

1. Резервное копирование данных.

Методы, используемые для резервного копирования, зависят от объема информации, важности информации и динамики ее изменения. Если говорить о носителях, применимых для хранения резервных копий, то дискеты годятся лишь в частных случаях для небольших объемов информации и личных архивов пользователей. В большинстве случаев используются либо накопители на магнитной ленте (стримеры), либо магнитно-оптические устройства, либо оптические типа WORM или WARM. Независимо от типа устройства для резервного копирования необходимо систематически проводить копирование данных во избежание их потери. Выбор конкретного способа зависит от того, как часто изменяются данные, какую ценность они представляют и как много времени потребуется для этой процедуры. Существуют следующие способы резервного копирования:

– Случайный. При таком подходе производится случайное

копирование отдельных файлов. Метод является наименее надежным, так как если обнаружится, что копия не самая новая, приходится проделать весь объем работы от момента изготовления этой резервной копии. Еще хуже, если носитель, на котором находится резервная копия, окажется поврежденным. Однако это лучше, чем ничего;

– Серьезный. Резервные копии производятся регулярно, и для их изготовления используются два набора носителей;

– Профессиональный. Этот метод используют вычислительные центры с дорогостоящим оборудованием и большими компьютерами. В нем используются три копии данных на трех наборах носителей (для надежности иногда используются по два экземпляра для каждого из наборов). При работе поочередно используется каждый из наборов. Этот метод иногда называют схемой «сын — отец — дед».

2. Избыточность данных.

Резервирование также подразумевает избыточность данных. С точки зрения подлинности лучше иметь два средних размера файловых сервера в локальной сети, чем один большой. Тогда в случае выхода из строя одного из них можно временно продолжать работать с другим. Конечно же, при этом на втором сервере должны находиться резервные копии рабочих файлов.

Несмотря на то, что системы хранения данных, основанные на магнитных дисках, производятся уже 40 лет, массовое производство отказоустойчивых систем началось совсем недавно. К ним относятся дисковые массивы с избыточностью данных, которые принято называть RAID.

Производители файловых серверов, учитывая необходимость избыточности данных, предлагают модели с дисковыми массивами — системами носителей на жестких магнитных дисках (НЖМД), в которых информация зеркально дублирована на различных дисководах. Естественно, что избыточность данных ни в коей мере не заменяет необходимость резервного копирования.

3. Защита от помех в электросети.

Сбои электропитания всегда происходят неожиданно. В момент сбоя электросети практически любая программа может в какой-то степени испортить файл, с которым она работала. Для защиты от таких ситуаций необходимо использовать источники бесперебойного питания (см. гл. 7) файл-серверов.

Защита информации от аварийных ситуаций

Защита информации от аварийных ситуаций заключается в создании и поддержании в работоспособном состоянии различных средств предупреждения, организации контроля и мероприятий по исключению несанкционированного доступа на средствах вычислительной техники при возникновении стихийных бедствий и аварийных ситуаций.

СПИСОК ЛИТЕРАТУРЫ

1. Гриценко В.И., Паньшин В.Н. Информационная технология: вопросы развития и применения. — Киев: Наукова думка, 1988. — 255 с.

2. Данилевский Ю.Г., Петухов И.А., Шибанов В.С. Информационная технология в промышленности. — Л.: Машиностроение, 1988. — 238 с.

3. Жуковский О.И. Информационные технологии: Учеб. пособие. — Томск: ТУСУР, 2003. — 167 с. ISBN 5-86889-122-8.

4. Огнев И.В., Сарычев К.Ф. Надежность запоминающих устройств. — М.: Радио и связь, 1988. — 224 с.

5. Блейхут Р. Теория и практика кодов, контролирующих ошибки: Пер.с англ. — М.: Мир, 1986. — 576 с.

6. Питерсон У., Уэлдон Э, Коды, исправляющие ошибки: Пер. с англ. — М.: Мир, 1976. — 600 с.

7. Кейтер Дж. Компьютеры — синтезаторы речи. — М.: Мир, 1985. — 237 с.

8. Шевкопляс Б.В. Микропроцессорные структуры. Инженерные решения. Дополнение первое: Справочник. — М.: Радио и связь, 1993. — 256 с.

9. Шевкопляс Б.В. Микропроцессорные структуры. Инженерные решения. Справочник — 2-е изд. перераб. и доп. — М.: Радио и связь, 1990. — 512 с.

10. Мельников В.П., Клейменов С.А., Петраков А.М. Информационная безопасность: Учеб. пособие для сред. проф. образования. — М.: Издательский центр «Академия», 2005. — 336 с.

Колегов А.А.

Информационные технологии в электронике
Учебное пособие

Усл. печ. л. Препринт
Томский государственный университет
систем управления и радиоэлектроники
634050, г.Томск, пр.Ленина, 40