Министерство науки и высшего образования РФ
Федеральное государственное бюджетное образовательное учреждение
высшего образования
АМУРСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
(ФГБОУ ВО «АмГУ»)

ТЕОРИЯ ИНФОРМАЦИИ
сборник учебно-методических материалов

для направления подготовки 10.03.01 Информационная безопасность

Благовещенск

2019

Печатается по решению
редакционно-издательского совета
факультета математики и информатики
Амурского государственного
университета

Составитель: Самохвалова С.Г.

Теория информации: сборник учебно-методических материалов для направлений подготовки 10.03.01 Информационная безопсность – Благовещенск: Амурский гос. ун-т, 2019

КРАТКОЕ ИЗЛОЖЕНИЕ ЛЕКЦИОННОГО МАТЕРИАЛА

Понятие и виды информации

Термин "информация" происходит от латинского слова "Information"- разъяснение, изложение, осведомленность. Можно считать, что этот термин в начальном представлении является общим понятием, означающим некоторые сведения, совокупность данных, знаний и т.д. Понятие информации должно быть с определенным объектом, свойства которого она отражает. Кроме того, наблюдается относительная независимость информации от ее носителя, поскольку возможны ее преобразование и передача по различным физическим средам с помощью разнообразных физических сигналов безотносительно к ее содержанию, т.е. к семантике, что и явилось центральным вопросом многих исследований, в том числе и в философской науке. Информация о любом материальном объекте может быть получена путем наблюдения, натурного либо вычислительного эксперимента, а также на основе логического вывода.

Поэтому говорят о доопытной (или априорной) информации и послеопытной (т.е. апостериорной) полученной, в итоге эксперимента.

Для человека любое восприятие реальных объектов окружающей действительности происходит через ощущения. Органы чувств человека и высшая нервная система позволяют ему воспринимать объекты. При обмене информацией имеют место источник в виде объекта материального мира и приемник - человек либо какой-то материальный объект. Информация возникает за счет отражения, которое является свойством всей материи, любой материальной системы. Свойство отражения совершенствуется по мере развития материи от элементарного отражения до высшей его формы - сознания. Процесс отражения означает взаимодействие объектов материального мира. Этот процесс наиболее прост в неорганической природе. Здесь преобладают механические, химические и физические взаимодействия. При таком отражении объекты пассивны. Новые формы отражения (физиологическое и психологическое) возникают в органической природе. В живом организме на основе отражения формируется способность приспосабливаться к изменяющимся окружающим условиям. У человека получают развитие более сложные формы отражения: познавательная и творческая. Эти формы носят сознательный характер и позволяют человеку активно воздействовать на окружающий мир.

Выделяют следующие аспекты информации:

- прагматический,
- семантический,
- синтаксический.

Прагматический аспект связан с возможностью достижения поставленной цели с использованием получаемой информации. Этот аспект информации влияет на поведение потребителя. Если информация была эффективной, то поведение потребителя меняется в желаемом направлении, т. е. информация имеет прагматическое содержание. Таким образом, этот аспект характеризует поведенческую сторону проблемы.

Семантический аспект позволяет оценить смысл передаваемой информации, соотнося ее с информацией, хранящейся до появления данной. Семантические связи между словами или другими смысловыми элементами языка отражают словарь-тезаурус. Он состоит из двух частей: списка слов и устойчивых словосочетаний, которые сгруппированы по смыслу, и некоторого ключа, т е. алфавитного словаря, позволяющего расположить слова и словосочетания в определенном порядке. Тезаурус имеет особое значение в системах хранения информации, в которые могут вводиться семантические отношения, в основном подчинения, что позволяет на логическом уровне осуществлять организацию информации в виде отдельных записей, массивов и их комплексов. Существуют развитые тезаурусы, в которые включаются сложные высказывания и семантические связи между ними. Это позволяет хранить более сложную информацию и детально оценивать семантическое содержание вновь поступающей информации. Наличие тезауруса позволяет переводить поступающую семантическую информацию на некоторый стандартизированный семантический язык в соответствии с выбранным тезаурусом. Таким образом, при возникновении информации можно изменить исходный тезаурус. Степень изменения тезауруса может быть принята как характеристика количества информации.

Синтаксический аспект информации связан со способом ее представления. В зависимости от реального процесса, в котором участвует информация, т.е. осуществляется ее сбор, передача, преобразование, отображение, представление, ввод или вывод, она представляется в виде специальных знаков, символов. Характерным носителем информации является **сообщение,** под которым обычно понимают все то, что подлежит передаче. Сообщения представляют в виде электрического сигнала, передаваемого по выбранной физической среде. Для этого сообщение подвергают преобразованию, т. е. придают ему электрический характер, далее кодированию, при котором сообщение превращается в некоторую. Последовательность символов, однозначно его отображающих, и модуляции, при которой каждый элемент кода (либо код в целом) переводится в электрический сигнал, способный передаваться на заданное расстояние по выбранному каналу связи. Процессы преобразования, кодирования и модуляции исключительно многообразны, а синтаксический аспект информации при ее передаче в настоящее время хорошо развит. Иной характер синтаксический аспект имеет, например, при хранении информации. В этом случае могут быть предложены такие формы, при которых удается осуществить быстрый поиск, введение новой информации, вывод требуемой информации из информационной базы и в целом обновления базы данных. Требуемому представлению информации при ее хранении отвечают разработанные к настоящему времени типовые структуры баз и банков данных, которые позволяют наилучшим образом реализовать информационное обслуживание пользователей в системе управления. Таким образом, развитие общества привело к тому, что оказалось необходимым хранить, обрабатывать, передавать, преобразовывать огромные объемы данных.

Виды информации

Все виды деятельности человека по преобразованию природы и общества сопровождались получением новой информации. Логическая, адекватно отображающая объективные закономерности природы, общества и мышления получила название *научной информации.* Ее делят по областям получения или пользования на следующие виды: политическую, техническую, биологическую, химическую, физическую и т.д.; по назначению - на массовую и специальную. Часть информации. Которая занесена на бумажный носитель, получила название *документальной информации.* Любое производство при функционировании требует перемещения документов, т.е. возникает документооборот. Для автоматизированных систем управления информация в документах составляет внешнее информационное обеспечение. В то же время большая часть информации хранится в памяти ЭВМ на магнитных лентах, дисках и т.д. Она определяется как внутримашинное информационное обеспечение.

Наряду с научной информацией в сфере техники при решении производственных задач используется *техническая информация.* Она сопровождает разработку новых изделий, материалов, конструкций агрегатов, технологических процессов. Научную и техническую информацию объединяют термином научно-техническая *информация:* в сфере материального производства может циркулировать технологическая информация, закрепленная в конструкторско-технологической документации. В плановых расчетах существует планово-экономическая информация, которая содержит интегральные сведения о ходе производства, значения различных экономических показателей.

Информация с точки зрения ее возникновения и совершенствования проходит следующий путь: человек наблюдает некоторый факт окружающей действительности, это факт отражается в виде совокупности данных, при последующем структурировании в соответствии с конкретной предметной областью данные превращаются в знания. Таким образом, верхним уровнем информации как результата отражения окружающей действительности (результата мышления) являются знания. Знания возникают как итог теоретической и практической деятельности. Информация в виде знаний отличается высокой структуризацией. Это позволяет выделить полезную информацию при анализе окружающих нас физических, химических и прочих процессов и явлений. На основе структуризации информации формируется информационная модель объекта. По мере развития общества информация как совокупность научно-технических данных и знаний превращается в базу системы информационного обслуживания научно-технической деятельности общества.

В настоящее время информация используется всеми отраслями народного хозяйства и наряду с энергией, полезными ископаемыми является ресурсом общества. С развитием общества возникает необходимость целесообразной организации информационного ресурса, т.е. конкретизации имеющихся фактов, данных и знаний по направлениям науки и техники.

В теории информации – понятие информации определяется исходя из понятия количества информацию при этом пользуются чисто математическим методом.

Ценность информации и ее количественная оценка

Исследованием методов передачи, хранения, приема информации занимается теория информации, инструментами, которой служат теория случайных процессов, теория кодирования, теория вероятностей. Внимание к проблеме передачи информации и ее количественной оценке было привлечено фундаментальными работами Норберта Винера и Клода Шеннона, положившими начало теории информации.

Важнейшим этапом в развитии теории информации явилась количественная оценка информации. Понятно, что эта оценка не должна быть связана с содержательной стороной информации. Только в этом случае появится возможность оценки информационных потоков в таких разных по своей природе объектах, как система связи, вычислительные машины, процессы управления, нервная система живого организма.

Основываясь на идее, что информация устраняет некоторую неопределенность, т.е. незнание, описание любого события или объекта может рассматривать как указание на то, в каком из возможных состояний находятся описываемый объект. Тогда протекание событий во времени есть не что иное, как смена состояний, выбранных с некоторой вероятностью из числа возможных. Чем больше неопределенности выбора, тем больше информации, так как результат выбора имеет большую степень неожиданности. Вот почему в теории информации количество информации является мерой снятия неопределенности одной случайной величины в результате наблюдения за другой. Если величины независимы, то количество информации равно нулю.

Количественный метод - одно из направлений в теории информации - наиболее распространенный и наиболее разработанный.

Измерение только количества информации не отвечает насущным потребностям современного общества – необходима мера ценности информации.

Количество информации в двух сообщениях может быть совершенно одинаковым, а смысл совершенно разным. Два слова, например «Мир» и «Рим», содержат одинаковое количество информации, состоят из одних и тех же букв, но смысл слов различен.

Вот еще один пример и еще одно направление (прагматическое - деловое) в этой науке.

Пассажиры едут в автобусе. Водитель объявляет остановку. Кое-кто выходит, остальные не обращают внимания на слова водителя - переданную им информацию. Почему? Потому что информация здесь имеет *разную ценность* для получателей, в роли которых в этом примере выступают пассажиры. Вышел тот, для кого информация была ценна. Значит, **ценность можно определить как свойство информации, влияющей на поведение ее получателя**.

Проблема определения ценности информации исключительно актуальна в настоящее время. Если в 60-е годы эта научная задача могла показаться несколько надуманной, то сейчас, когда уже трудно даже с помощью компьютеров обрабатывать информационный поток, разработанные методы определения ценности информации сыграли бы существенную роль в получении человеком необходимой информации.

Одной из первых работ советских авторов по проблеме ценности информации являлась статья А.А. Харкевича, в которой предлагалось принять за меру ценности информации количество информации, необходимое для достижения поставленной цели. Так, если до получения информации вероятность достижения цели равнялась p0, а после ее получения p1, то ценность информации определяется отношением p1/p0, ценность информации при этом измеряется в битах.

Пользоваться критерием оценки значимости информации следует избирательно. Дело в том, что информация относительна: полученные сведения могут не иметь отношения к решаемой задаче. Но, тем не менее, быть информативными.

Статья в специальном журнале для специалиста может иметь большую ценность, в то время как для читателя журнала <<За рулем>> - никакой.

Вообще, оценка значимости информации производится человеком часто интуитивно, опираясь на собственный интеллект и опыт. Сложность проблемы оценки информации хорошо иллюстрируется известной шуткой из сборника *<<Физики продолжают шутить>>: <<Альберт Эйнштейн любил фильмы Чарли Чаплина и относился с большой симпатией к созданному им герою. Однажды он написал в письме к Чаплину: Ваш фильм Золотая лихорадка понятен во всем мире, и вы непременно станете великим человеком. На что, Чаплин ответил так: Я вами восхищаюсь еще больше. Вашу теорию относительности никто в мире не понимает, а вы все-таки стали великим человеком.>>*

В других работах говорится о том, что максимальной ценностью обладает то количество информации, которое уменьшает потери до нуля при достижении поставленной цели.

В том варианте теории ценности информации, который развивает М.М. Бонгард, предполагается, что получаемое количество информации может не иметь ценности или ее ценность может быть отрицательной. При этом можно рассматривать случай, когда передается ложная информация и неопределенность не сокращается, а возрастает.

Информация называется полезной, если она уменьшает неопределенность решаемого алгоритма. Бонгард отмечает, что не имеет смысла говорить о полезной информации, если не указана задача, которая решается, начальное состояние решающего алгоритма.

Названными вариантами теория ценности информации отнюдь не исчерпывается. Нельзя в связи с этим не отметить попытки построить семантическую теорию информации. Действительно, полное незнание предмета не позволяет извлечь существенной научной информации, извлекаемой из сообщения.

Таким образом, различные подходы к решению проблемы ценности информации имеют принципиально общие черты: они предлагают измерять ценность информации через ее количество, связывают ценность информации с поставленной задачей.

Количественная мера информации

Для того чтобы оценить и измерять количество информации в соответствии с вышеизложенными аспектами, применяются различные подходы и методы. Среди них выделяются статистический, семантический, прагматический и структурный. Исторически наибольшее развитие получил статистический подход.

Статистический подход. Он изучается в обширном разделе кибернетики, называемой теорией информации. Основоположником этого подхода считается К.Шеннон, опубликовавший в 1948 г. Свою математическую теорию связи. Большой вклад в теорию информации до него внесли ученые Найквист и Хартли, которые соответственно в 1924 и 1928 гг. напечатали работы по теории телеграфии и передачи информации. Признаны во всем мире исследования по теории информации российских ученых А.Н.Колмогорова, А.Я.Хинчина, В.А. Котельникова, А.А.Харкевича и т.д.

К.Шенноном было введено понятие количества информации как меры неопределенности состояния системы, снимаемой при получении информации. Количественно выраженная неопределенность состояния получила название энтропии по аналогии с подобным понятием в статистической механике. При получении информации уменьшается неопределенность, т.е. энтропия системы. Очевидно, что, чем больше информации получает наблюдатель, тем больше снижается неопределенность, и энтропия системы уменьшается. При энтропии, равной нулю, о системе имеется полная информация, и наблюдателю она представляется целиком упорядоченной. Таким образом, получение информации связано с изменением степени неосведомленности получателя о состоянии этой системы.

До получения информации её получатель мог иметь некоторые предварительные (априорные) сведения о системе X. Оставшаяся неосведомленность и является для него мерой неопределенности состояния (энтропии) системы. Обозначим априорную энтропию системы X через $H(X)$. после получения некоторого сообщения наблюдатель приобрел дополнительную информацию $I(X)$, уменьшившую его начальную неопределенность так, что апостериорная (после получения

информации) неопределенность состояния системы стала *H'(X)*. Тогда количество информации I может быть определено как

I(X)=H(X)-H'(X)

Другими словами, количество информации измеряется уменьшением (изменением) неопределенности состояния системы.

Если апостериорная энтропия системы обратилась в нуль, то первоначально неполное значение заменится полным значением и количество информации, полученной в этом случае наблюдателем, будет

I(X)=H(X)

т.е. энтропия системы может рассматриваться как мера недостающей информации.

Если система X обладает дискретным состояниями (т.е. переходит из состояния в состояния скачком), их количество равно N, а вероятность нахождения системы в каждом из состояний – $p_1, p_2, p_3, ..., p_N$ (причем $\sum_{i=1}^{N} p_i = 1$ и $p_i \le 1$), то согласно теореме Шеннона энтропия системы равна

$$H(X) = -K_0 \sum_{i=1}^{N} p_i \log_a p_i$$

Здесь коэффициент K_0 и основание логарифма a определяют систему единиц измерения количества информации. Логарифмическая мера информации была предложена Хартли для представления технических параметров систем связи как более удобная и более близкая к восприятию человеком, привыкшим к линейным сравнениям с принятыми эталонами. Например, каждый чувствует, что две однотипные дискеты должны обладать вдвое большей емкостью, чем одна, а два идентичных канала связи должны иметь удвоенную пропускную способность.

Знак минус поставлен для того, чтобы значение энтропии было положительным, так как $p_i \le 1$ и логарифм в этом случае отрицательный.

Если все состояния системы равновероятны, т.е. $p_i = \frac{1}{N}$, её энтропия

Энтропия H обладает рядом интересных свойств. Вот некоторые из них.

Энтропия H равна нулю только тогда, когда все вероятности p_i , кроме одной, равны нулю, а эта единственная вероятность равна единице. Таким образом, *H=0* только в случае полной определенности состояния системы.

При заданном числе состояний системы N величина H максимальна и равна $K_0 \log_a N$, когда все p_i равны.

Определим единицы измерения количества информации с помощью выражения для энтропии системы с равновероятными состояниями.

Пусть система имеет два равновероятных состояния, т.е. N=2. Будем считать, что снятие неопределенности о состоянии такой системы дает одну единицу информации, так как при полном снятии неопределенности энтропия количественно равна информации *H=I*. Тогда

$$1 = K_0 \log_a 2$$

Очевидно, что правая часть равенства будет тождественно равна единице информации, если принять $K_0 = 1$ и основание логарифма a=2. В общем случае при N равновероятных состояний количество информации будет

$$I = \log_2 N$$

Эта формула получила название формулы Хартли и показывает, что количество информации, необходимое для снятия неопределенности о системе с равновероятными состояниями, зависит лишь от количества этих состояний.

Информация о состояниях системы передается получателю в виде сообщений, которые могут быть представлены в различной синтаксической форме, например в виде кодовых комбинаций,

использующих m различных символов и n разрядов, в каждой из которых может находиться любой из символов. Если код не избыточен, то каждая кодовая комбинация отображает одно из состояний системы. Количество кодовых комбинаций будет

$$N = m^n$$

Подставив это выражение в формулу для I .

$$I = n \log_2 m$$

Если код двоичный, т.е. используются лишь два символа (0 или 1), то m=2 и I=n. В этом случае количество информации в сообщении составит n двоичных единиц, называемых битами (binary digit (bit) – двоичная цифра).

При использовании в качестве основания логарифма числа десять единиц измерения информации могут быть десятичными, или дитами. Так как $\log_2 N = \log_{10} \frac{N}{\log_{10} 2} = 3{,}321 \log_{10} N$, то десятичная единица составляет примерно 3,33 бита.

Иногда удобно применять натуральное основание логарифма ℓ . В этом случае получающие единицы информации называются натуральными или *натами*. Переход от основания a к основанию b требует лишь умножения на $\log_b a$.

Введенная количественная статистическая мера информации широко используется в теории информации для оценки собственной, взаимной, условной и других видов информации. Рассмотрим в качестве примера собственную информацию. Под *собственной информацией* будем понимать информацию, содержащуюся в данном конкретном сообщении. А конкретное сообщение, как указывалось, дает получателю информацию о возможном существовании конкретного состояния системы. Тогда количество собственной информации, содержащееся в сообщении X_i , определяется как

$$I(X_i) = -\log_2 p(X)$$

Собственная информация имеет следующие свойства:

- собственная информация неотрицательна
- чем меньше вероятность возникновения сообщения, тем больше информации оно содержит. Именно поэтому неожиданные сообщения так воздействуют на психику человека, что содержащаяся в них большое количество информации создает информационный психологический удар, иногда приводящий к трагическим последствиям.
- Если сообщение имеет вероятность возникновения, равную единице, то информация, содержащаяся в нем, равна нулю, так как заранее известно, что может прийти только это сообщение, а значит, ничего нового потребитель информации не получает.
- Собственная информация обладает свойством аддитивности, т.е. количество собственной информации нескольких независимых сообщений равно их сумме. Например, для собственной информации двух сообщений X_i и Y_i может быть записано:

$$I(X_i, Y_i) = -\log_2 P(X_i) - \log_2 P(Y_i) = I(X_i) + I(Y_i).$$

Следует еще раз отметить, что статистический подход к количественной оценке информации был рассмотрен для дискретных систем, случайным образом переходящих из состояния в состояние, и, следовательно, сообщение об этих состояниях также возникает случайным образом.

Кроме того, статистический метод определения количества информации практически не учитывает семантического и прагматического аспектов информации.

Семантический подход. Этот подход является наиболее трудно формализуемым и до сих пор окончательно неопределившимся.

Наибольшее признание для измерения смыслового содержания информации получила тезаурусная мера, предложенная Ю.И. Шнейдером. Идеи тезаурусного метода были сформулированы

ещё основоположником кибернетики Н. Винером. Для понимания и использования информации её получатель должен обладать определенным запасом знаний.

Если индивидуальный тезаурус потребителя $S_П$ отражает его знания о данном предмете, то количество смысловой информации I_C, содержащаяся в некотором сообщении, можно оценить степенью изменения этого тезауруса, произошедшего под воздействием данного сообщения. Очевидно, что количество информации I_C нелинейно зависит от состояния индивидуального тезауруса пользователь, и хотя смысловое содержание сообщения S постоянно, пользователи, имеющие отличающиеся тезаурусы, будут получать *неодинаковое* количество информации.

В самом деле, если индивидуальный тезаурус получателя информации близок к нулю, $S_П \approx 0$, то в этом случае и количество воспринятой информации равно нулю: $I_C = 0$.

Иными словами, получатель не понимает принятого сообщения, и, как следствие, для него количество воспринятой информации равно нулю. Такая ситуация эквивалентна прослушиванию сообщения на неизвестном иностранном языке. Несомненно, сообщение не лишено смысла, однако оно непонятно, а значит, не имеет информативности.

Количество семантической информации I_C в сообщении также будет равно нулю, если пользователь информации абсолютно все знает о предмете, т.е. его тезаурус $S_П$, и сообщение не дает ему ничего нового.

Интуитивно мы чувствуем, что между этими полярными значениями тезауруса пользователя существует некоторое оптимальное значение , $S_{П.ОПТ}$, при котором количество информации I_C, извлекаемое из сообщения, становиться для получателя максимальным.

Тезаурусный метод подтверждает тезис о том, что информация обладает свойством относительности и имеет, таким образом, относительную, субъективную ценность. Для того чтобы объективно оценить научную информации, появилось понятие общечеловеческого тезауруса, степень изменения которого и определяет значительность получаемых человечеством новых знаний.

Прагматический подход. Он определяет количество информации как меру, способствующую достижению поставленной цели. Одной из первых работ, реализующих этот подход, явилась статья А.А. Харкевича. В ней он предлагал принять за меру ценности информации количество информации, необходимое для достижения поставленной цели. Этот подход базируется на статической теории Шеннона и рассматривает количество информации как приращение вероятности достижения цели. Так, если принять вероятность достижение цели до получения информации равной p_0, а после её получения - p_1, то прагматическое количество информации $I_П$ определяется как

$$I_П = \log \frac{p_1}{p_0}.$$

Если основание логарифма сделать равным двум, то $I_П$ будет измеряться в битах, как и при статистическом подходе.

При оценке количества информации в семантическом и прагматическом аспектах необходимо учитывать и временную зависимость информации. Дело в том, что информация, особенно в системах управления экономическими объектами, имеет свойство стареть, т.е. её ценность со временем падает, и важно использовать её в момент наибольшей ценности.

Структурный подход. Он связан с проблемами хранения, реорганизации и извлечения информации и по мере увеличения объемов накапливаемой в компьютерах информации приобретает все большее значение.

При структурном подходе абстрагируются от субъективности, относительно ценности информации и рассматривают логические и физические структуры организации информации. С

изобретением компьютеров появилось возможность хранить на машинных носителях громадные объемы информации. Но для её эффективного использования необходимо определить такие структуры организации информации, чтобы существовала возможность быстрого поиска, извлечения, записи, модификации информационной базы.

При машинном хранении структурной единицей информации является один байт, содержащий восемь бит (двоичных единиц информации). Менее определенной, но также переводимой в байты является неделимая единица экономической информации – реквизит.

Реквизиты объединяют в показатели, показатели – в записи, записи – в массивы, из массивов создают комплексы массивов, а из комплексов – информационные базы. Структурная теория позволяет на логическом уровне построить оптимальную структуру информационной базы, которая затем с помощью определенных средств реализуется на физическом уровне – уровне технических устройств хранения информации. От выбранной структуры хранения зависит такой важный параметр, как время доступа к данным, т.е. структура влияет на время записи и считывания информации, а значит, и на время создания и реорганизации информационной базы.

Информационная база совместно с системой управления базой данных (СУБД) формирует автоматизированный банк данных.

Значение структурной теории информации растет при переходе от банков данных к банкам знаний, в которых информация подвергается ещё более высокой степени структуризации.

После преобразования информации в машинную форму (рис.2.) её аналитический и прагматический аспекты как бы уходят в тень, и дальнейшая обработка информации происходит по «машинным законам», одинаковым для информации любого смыслового содержания. Информация в машинном виде, т.е. в форме электрических, магнитных и тому подобных сигналов и состояний, носит название данных. Для того, чтобы понять их смысловое содержание, необходимо данные снова преобразовать в информацию.

Преобразование “информация – данные ” производятся в устройствах ввода-вывода ЭВМ.

Базисным понятием всей теории информации является понятие энтропии. Энтропия – мера неопределенности некоторой ситуации.

Энтропия как мера неопределенности

Подойдем к описанию случайных событий несколько с иной стороны. То, что событие случайно, означает отсутствие полной уверенности в его наступлении, что, в свою очередь, создает ***неопределенность*** в исходах опытов, связанных с данным событием. Безусловно, степень неопределенности различна для разных ситуаций. Например, если опыт состоит в определении возраста случайно выбранного студента 1-го курса дневного отделения вуза, то с большой долей уверенности можно утверждать, что он окажется менее 30 лет; хотя по положению на дневном отделении могут обучаться лица в возрасте до 35 лет, чаще всего очно учатся выпускники школ ближайших нескольких выпусков. Гораздо меньшую определенность имеет аналогичный опыт, если проверяется, будет ли возраст произвольно выбранного студента меньше 20 лет. Для практики важно иметь возможность произвести численную оценку неопределенности разных опытов. Попробуем ввести такую количественную меру неопределенности.

Начнем с простой ситуации, когда опыт имеет n равновероятных исходов. Очевидно, что неопределенность каждого из них зависит от n, т.е.

неопределенность = f(n)

Можно указать некоторые свойства этой функции:

(1) f(1)=0, поскольку при n=1 исход опыта не является случайным и, следовательно, неопределенность отсутствует;

(2) f(n) возрастает с ростом n, т.к. ввиду большого числа возможных исходов предсказание результата опыта становится весьма затруднительным.

Для определения явного вида функции f(n) рассмотрим два независимых опыта A и B, с количествами равновероятных исходов, соответственно n_A и n_B. Рассмотрим сложный опыт C, который состоит в одновременном выполнении опытов A и B. Число возможных исходов опыта C равно $n_A \cdot n_B$, причем, все они равновероятны. Очевидно, неопределенность исхода такого опыта будет больше неопределенности опыта A, поскольку к ней добавляется неопределенность B. Естествен-

но допустить, что мера неопределенности С равна сумме неопределенностей опытов А и В, т.е. неопределенность аддитивна:

$$f(n_A n_B) = f(n_A) + f(n_B) \qquad (1.1)$$

Теперь можно задуматься о том, каким может быть явный вид функции f(n), чтобы он удовлетворял свойствам (1) и (2). Легко увидеть, что такому набору свойств удовлетворяет функция log(n), причем, можно показать, что она единственная из всех возможных классов функций. Таким образом:

за меру неопределенности опыта с n равновероятными исходами можно принять число log(n).

Следует заметить, что выбор основания логарифма в данном случае значения не имеет, поскольку в силу известной формулы перехода от одного основания логарифма к другому

$$log_b n = log_b a \cdot log_a n \quad ,$$

переход к другому основанию состоит во введении одинакового для обеих частей выражения постоянного множителя $\log_b a$, что равносильно изменению масштаба (т.е. размера единицы) измерения неопределенности. Поскольку это так, мы имеет возможность выбрать удобное для нас (из каких-то дополнительных соображений) основание логарифма. Таким удобным основанием оказывается 2, поскольку в этом случае за единицу измерения принимается неопределенность, содержащаяся в опыте, имеющем лишь два равновероятных исхода, которые можно обозначить, например, ИСТИНА (True) и ЛОЖЬ (False) и использовать для анализа таких событий аппарат математической логики.

*(а) Единица измерения неопределенности при двух возможных исходах опыта называется **бит**.*

(Название **бит** происходит от английского **binary digit**, что в дословном переводе означает «двоичный разряд» или «двоичная единица».)

Таким образом, нами установлен явный вид функции, описывающей неопределенность опыта, имеющего n равновероятных исхода:

$$f(n) = \log_2 n \qquad (1.2)$$

На основании формул ($p = \frac{1}{n}$) несложно найти неопределенность, вносимую каждым отдельным исходом в общую. Поскольку исходов n и все они равновероятны (и, следовательно, равнозначны), а общая неопределенность равна $\log_2 n$, из свойства аддитивности неопределенности следует, что неопределенность, вносимая одним исходом составляет

$$\frac{1}{n}\log_2 n = -\frac{1}{n}\log_2 \frac{1}{n} = -p \cdot \log_2 p \quad ,$$

где $p = \frac{1}{n}$ – вероятность любого из отдельных исходов.

Таким образом, неопределенность (обозначим, наконец, ее H), вносимая каждым из равновероятных исходов, равна:

$$H = -p \log_2 p = \log n \qquad (1.3)$$

Данную формулу в 1926 г предложил Хартли.

Теперь попробуем обобщить формулу на ситуацию, когда исходы опытов не равновероятны, например, $p(A_1)$ и $p(A_2)$. Тогда:

$$H_1 = -p(A_1) \cdot \log_2 p(A_1) \quad \text{и} \quad H_2 = -p(A_2) \cdot \log_2 p(A_2)$$

$$H_0 = H_1 + H_2 = -p(A_1) \cdot \log_2 p(A_1) - p(A_2) \cdot \log_2 p(A_2)$$

Обобщая это выражение на *n* неравновероятных исходов, получим:

$$H(A) = -\sum_{i=1}^{n} p(A_i) \log_2 p(A_i) \qquad (1.4)$$

Введенная таким образом величина получила название ***энтропия***.

Впервые мера (1.4) была предложена Клодом Шенноном в его фундаментальной работе "Математические основы теории связи" опубликованной в 1948г в которой были заложены основы современной ТИ. Предполагающая мера была названа энтропией не случайно. Дело в том, что вид формулы (1.4) совпадает с полученным ранее результатом Больцманом выражением для энтропии термодинамической системы.

Рассмотрим взаимосвязь меры Шеннона с мерой Хартли если в источнике может быть реализовано h равновероятных состояний, то вероятность каждого из них, с учетом этого меру неопределенности источника Хартли можно трактовать, как количество информации приходящей на одно дискретное сообщение (поскольку все сообщения источника равновероятные количества информации в каждом из них равны) в тоже время энтропия по Шеннону это среднее количество информации содержащееся в одном из не равновероятных состояний. Она позволяет учесть статистические свойства источника информации.

Наряду с рассмотренными мерами Хартли и Шеннона существуют и другие подходы к определению количества информации. Наиболее интересной, наиболее новой явилась информационная концепция Колмогорова, ее основным тезисом является то, что на основании определения энтропии (1.4) количество информации связывается с вероятностью наступления Pi, т.к. понятие вероятности имеет смысл лишь в связи с массовыми явлениями количества единиц информации в единичном акте и представляющих интерес в связи с данным исходом, оказывается выраженным через вероятности массовых явлений. Шенноновская мера интересна не сама по себе, а как основание встроенной теории позволяющей изменить и расширить существующие предположения о возможностях в технике связи, которая и подлежит в рассмотрении ТИ.

*(b) энтропия является **мерой неопределенности** опыта, в котором проявляются случайные события, и равна **средней неопределенности** всех возможных его исходов.*

Впервые понятие энтропии было введено в 1865 г. немецким физиком Рудольфом Клаузиусом как функции состояния термодинамической системы, определяющей направленность самопроизвольных процессов в системе. Клаузиус сформулировал II начало термодинамики. В частности, он показал, что максимума энтропия достигает при полной раз упорядоченности в системе, чему соответствует состояние равновесия. Другими словами, в физике энтропия оказывается мерой беспорядка в системе. Позднее (в 1872 г.) Людвиг Больцман, развивая статистическую теорию, связал энтропию системы с вероятностью ее состояния, дал статистическое (вероятностное) толкование II-му началу термодинамики и, в частности, показал, что вероятность максимальна у полностью разупорядоченной (равновесной) системы, причем, энтропия и термодинамическая вероятность оказались связанными логарифмической зависимостью! Сходство понятий и соотношений между ними в теории информации и статистической термодинамике, как оказалось позднее, имеет глубокий смысл.

Что дает понятие энтропии для решения практических задач? Рассмотрим одну из них. Пусть имеются два ящика, в каждом из которых по 12 шаров. В первом – 3 белых, 3 черных и 6 красных; во втором – каждого цвета по 4. Опыты состоят в вытаскивании по одному шару из каждого ящика. Что можно сказать относительно неопределенностей этих опытов? Согласно находим энтропии обоих опытов:

$$H_1 = -\frac{3}{12}\log_2\frac{3}{12} - \frac{3}{12}\log_2\frac{3}{12} - \frac{6}{12}\log_2\frac{6}{12} = 1{,}5$$

$$H_2 = -\frac{4}{12}\log_2\frac{4}{12} - \frac{4}{12}\log_2\frac{4}{12} - \frac{4}{12}\log_2\frac{4}{12} = \log_2 3 = 1{,}58$$

Ясно, что $H_2 > H_1$, т.е. во втором опыте неопределенность исхода выше, что, кстати, иллюстрирует справедливость формулы.

Чем больше энтропия источника, тем больше степень неожиданности выдаваемых им сообщений в среднем, т.е. тем более неопределенным является ожидание сообщений.

Вернемся к понятию энтропии как меры неопределенности некоторого опыта, исход которо-

го зависит от выбора одного элемента из множества исходных. Множество исходных элементов называется выборочным пространством. Вероятности нахождения элементов исходного множества в том или ином состоянии есть числа положительные, сумма их равна 1.

Выборочное пространство и его вероятностные характеристики представляют собой ансамбль сообщений. Для дискретного ансамбля вероятность события равна сумме вероятностей элементов выборочного пространства, содержащихся в этом событии.

Ансамбль сообщений на выходе источника будем называть ансамблем источника сообщений и обозначать буквой А. Абстрактный алфавит, при помощи которого мы представляем исходное множество элементов источника сообщений, обозначается $\{a_1, a_2, \ldots, a_i, \ldots, a_m\}$. Вероятности появления буквы на выходе источника сообщений обозначают $p(a_1)$, $p(a_2)$, …, $p(a_i)$,… $p(a_m)$. $\sum_{i=1}^{m} p(a_i) = 1$. В этом случае энтропия источника сообщений

$$H(A) = -\sum_{i=1}^{M} p(a_i) \log p(a_i)$$

и представляет собой неопределенность появления на выходе источника сообщений буквы первичного алфавита.

Ансамбль сообщений на выходе приемника будем называть ансамблем приемника сообщений и обозначать буквой В. Для того чтобы отличить переданные и принятые сигналы, абстрактный алфавит в котором представлен ансамбль приемника сообщений, обозначается $\{b_1, b_2, \ldots, b_j, \ldots, b_n\}$, а соответствующие вероятности - $p(b_1)$, $p(b_2)$, …, $p(b_i)$,… $p(b_n)$.

Энтропия приемника сообщений

$$H(B) = -\sum_{j=1}^{N} p(b_j) \log p(b_j)$$

и представляет собой неопределенность появления на входе приемника буквы после ее появления на выходе источника сообщений. Если в канале связи не происходит потерь информации, то всегда буква a_1 соответствует букве b_1, $a_2 - b_2$ и т.д. При этом $H(A)=H(B)$.

Понятие энтропии используется не только при передаче сообщений. Энтропия широко применяется для описания состояния механических и термодинамических систем, для изучения свойств алфавитов различных языков, при исследовании экономических систем.

Свойства энтропии.

1. Энтропия является вещественной и неотрицательной величиной, так как для любого i ($1<i<N$) p_i изменяется в интервале от 0 до 1, $\log p_i$ отрицателен и, следовательно, $- p_i \log p_i$ положительна.

2. Энтропия – величина ограниченная.

3. Энтропия обращается в нуль лишь в том случае, если вероятность одного из состояний равна единице; тогда вероятности всех остальных состояний, равны нулю.

4. Энтропия максимальна, когда все состояния источника равновероятны.

$H_{\max} = \log N$

5. Энтропия объединения нескольких статистически независимых источников информации равна сумме энтропий исходных источников.

Рассмотрим объединение, включающее два источника информации *а* и *в*. Под объединением двух источников *а* и *в* понимают обобщенный источник информации (*а,в*), характеризующийся вероятностями *p(а,в)* всех возможных комбинаций состояний *а* источника *а* и *в* источника *в*. Аналогично трактуется и объединение ансамблей.

В случае статистической независимости источников информации *а* и *в* запишем $p(a_i,в_j)=p(a_i)p(в_j)$

тогда

$$H(a,b) = -\sum_{i=1}^{N}\sum_{j=1}^{K} p(a_i)p(b_j)\log p(a_i)p(b_j) =$$

$$= -\sum_{i=1}^{N} p(a_i)\log p(a_i)\sum_{j=1}^{K} p(b_j) - \sum_{j=1}^{K} p(b_j)\log p(b_j)\sum_{i=1}^{N} p(a_i)$$

Учитывая, что

$\sum_{i=1}^{N} p(a_i) = 1$ и $\sum_{i=1}^{N} p(b_j) = 1$

получим

H(A,B)=H(A)+H(B)=H(B,A)

Условная энтропия

Если состояния элементов системы не зависят друг от друга, если состояние одной системы не зависит от состояния другой системы, то неопределенность того, что некоторый элемент системы будет находиться в одном из *k* возможных состояний полностью определялась бы вероятностными характеристиками отдельных элементов системы, либо вероятностными характеристиками состояний самих систем. При этом подразумевается, что символы сообщения взаимонезависимы, т.е. с приходом одного символа распределение вероятностей последующих символов не изменяется. На практике же чаще всего встречаются взаимозависимые символы и сообщения. Если передавать не просто отдельные буквы алфавита, а смысловые сообщения, то можно убедиться, что существует взаимозависимость передаваемых символов. Одни буквы встречаются чаще. другие реже, одни буквы и слова часто следуют за другими, другие редко.

Понятие условной энтропии широко используется для определения информационных потерь при передаче информации.

Если элементы источника сообщений принимают состояния $a_1, a_2, \ldots a_n$ с вероятностями соответственно $p(a_1), p(a_2), \ldots p(a_n)$, а элементы адресата – состояния $в_1, в_2, \ldots в_m$, с вероятностями соответственно $p(в_1), p(в_2), \ldots p(в_m)$, то понятие условной энтропии $H(в_j/a_i)$ выражает неопределенность того, что отправив a_i, мы получим $в_j$. Если в канале связи присутствуют помехи, то с различной степенью вероятности может быть принят любой из сигналов b_j, и наоборот, принятый сигнал $в_j$ может появится в результате отправления любого из сигналов a_i. Если в канале связи помехи отсутствуют, то всегда посланному сигналу а соответствует принятый сигнал в и т.д. При этом энтропия источника $H(A)$ равна энтропии приемника $H(B)$. Если в канале связи присутствуют помехи, то они уничтожают часть передаваемой информации.

Информационные потери полностью описываются через частную и общую условную энтропию. Вычисление частных и общей условной энтропии удобно производить при помощи канальных матриц. **Если канал связи описывается со стороны источника сообщений (т.е. известен посланный сигнал)**, то вероятность того, что при передаче сигнала а по каналу связи с помехами мы получим сигнал в, обозначается как условная вероятность $p(в_j/a_i)$, а канальная матрица имеет вид:

B / A	b_1	b_2 …	b_j ……..	b_m
a_1	$p(b_1/a_2)$,	$p(b_2/a_1)$, ...,	$p(b_j/a_1)$,	
a_2	$p(b_1/a_2)$,	$p(b_2/a_2)$, ...,	$p(b_j/a_2)$,	
	..., $p(b_m/a_1)$			
a_i	..., $p(b_m/a_2)$			
	……			
	…..			
a_m				

	$p(b_1/a_i),\quad p(b_2/a_i),\ ...,\quad p(b_j/a_i),\quad ...,\quad p(b_m/a_2)$ …… …… $p(b_1/a_m),\quad p(b_2/a_m),\quad ...,\quad p(b_j/a_m),$ $...,\quad p(b_m/a_m)$

Вероятности, которые расположены по диагонали, определяют вероятности правильного приема, остальные – ложного. Значения цифр, заполняющих колонки канальной матрицы, обычно уменьшаются по мере удаления от главной диагонали и при полном отсутствии помех все, кроме цифр, расположенных на главной диагонали, равны нулю.

Прохождение данного вида сигнала со стороны источника сообщений в данном канале связи описывается распределением условных вероятностей вида $p(в_j/a_i)$. Например, для сигнала a_1 распределением вида

$$p(b_1/a_1)+p(b_2/a_1)+\ldots+p(b_j/a_i)+\ldots+p(b_m/a_1)=1$$

Потери информации, приходящиеся на долю сигнала a_i описываются при помощи частной условной энтропии. Например, для сигнала a_1

$$H(b_j/a_1)=-\sum p(b_j/a_1)\log p(b_j/a_1)$$

Суммирование производится по j, так как i-е состояние остается постоянным.

Потери при передаче всех сигналов по данному каналу связи описываются при помощи общей условной энтропии. Для ее вычисления следует просуммировать все частные условные энтропии, т.е. произвести двойное суммирование по i и по j. При этом, в случае равновероятных появлений сигналов на выходе источника сообщений

$$H(B/A)=-\frac{1}{N}\sum_j\ \sum_i p(b_j/a_i)\log p(b_j/a_i)$$

В случае неравновероятного появления символов источника сообщений следует учесть вероятность появления каждого символа, умножив на нее соответствующую частную условную энтропию. При этом общая условная энтропия

$$H(B/A)=-\sum_i\sum_j p(a_i)p(b_j/a_i)\log p(b_j/a_i)$$

B A	b_1 b_2 … b_j …….. b_m

a_1	$p(a_1/b_2),\quad p(a_2/b_1),\ ...,\quad p(a_1/b_j),$
a_2	$p(a_1/b_2),\quad p(a_2/b_2),\ ...,\quad p(a_2/b_j),$
	$...,\quad p(a_1/b_m)$
a_i	$...,\quad p(a_1/b_m)$
	……
	……..
a_m	
	$p(a_i/b_1),\quad p(a_i/b_2),\ ...,\quad p(a_i/b_j),\quad ...,\quad p(a_i/b_m)$
	……
	……
	$p(a_m/b_1),\quad p(a_m/b_2),\quad ...,\quad p(a_m/b_j),\ ...,\quad p(a_m/b_m)$

Если исследовать канал связи ***со стороны приемника сообщений*** (т.е. известен принятый сигнал), то с получением сигнала $в_j$ предполагаем, что был послан какой-то из сигналов a_i. При этом канальная матрица имет вид:

В этом случае единице должны равняться суммы условных вероятностей не по строкам, а по столбцам канальной матрицы

$$p(a_1/b_j)+p(a_2/b_j)+\ldots+p(a_i/b_j)+\ldots+p(a_m/b_j)=1$$

Частная условная энтропия

$$H(a_i/b_j)=-\sum_{i=1}^{m}p(a_i/b_j)\log p(a_i/b_j)$$

Общая условная энтропия

$$H(A/B)=-\sum_j\sum_i p(b_j)p(a_i/b_j)\log p(a_i/b_j).$$

Понятие условной энтропии в теории информации используется при определении взаимозависимости между символами кодируемого алфавита, для определения потерь при передаче информации по каналам связи, при вычислении энтропии объединения.

Во всех случаях при вычислении условной энтропии в том или ином виде используются условные вероятности.

Если в канале связи помехи отсутствуют, то все элементы канальной матрицы, кроме элементов, расположенных на главной диагонали, равны нулю. Вероятность получения правильного сигнала станет безусловной, а условная энтропия будет равна нулю. Канальная матрица будет иметь вид

$$p(a/b)=\begin{vmatrix}1\,0\,0....0\\0\,1\,0....0\\.....\\0\,0\,0.....1\end{vmatrix}$$

В этом случае условная энтропия будет равна нулю.

Энтропия объединения

Взаимная энтропия, или как ее часто называют, *энтропия объединения* используется для вычисления энтропии совместного появления статистических зависимых сообщений.

Пусть $(a_1,a_2,\ldots,a_i,\ldots,a_n)$ есть выборочное пространство *A*, характеризующее источник сообщений, а $(b_1,b_2,\ldots,b_j,\ldots,b_m)$ есть выборочное пространство *B*, характеризующее приемник сообщений. При этом *а* есть сигнал на входе шумящего канала, а *в* – сигнал на его выходе. В этом случае взаимная энтропия представляет собой информацию о переданном сигнале a_i, содержащегося в принятом сигнале b_j. Взаимосвязь переданных и принятых сигналов описывается вероятностями совместных событий вида $p(a_i,b_j)$, а взаимосвязь выборочных пространств *A* и *B* описывается матрицей объединения вида:

$$p(a_i,b_j)=\begin{vmatrix} p(a_1,b_1) & p(a_1,b_2) & \ldots & p(a_1,b_m) \\ p(a_2,b_1) & p(a_2,b_2) & \ldots & p(a_2,b_m) \\ \ldots\ldots\ldots & \ldots\ldots\ldots & \ldots & \ldots\ldots\ldots \\ p(a_m,b_1) & p(a_m,b_2) & \ldots & p(a_m,b_m) \end{vmatrix}$$

Если матрица описывает канал связи, то число строк матрицы равно числу столбцов, *m=n*, и пределы суммирования по *i* и по *j* одинаковы.

Независимо от равенства или неравенства числа строк числу столбцов матрица объединения обладает следующими свойствами

1) $\sum_i p(a_i,b_j)=p(b_j)$

Сумма вероятностей по столбцам равна вероятности приёмника.

2) $\sum_j p(a_i,b_j)=p(a_i)$

Сумма вероятностей по строкам рана вероятности источника.

3) $\sum_i p(a_i)=\sum_j p(b_j)=1$, т. е. $\sum_i\sum_j p(a_i,b_j)=1$

Сумма всех элементов равна 1.

Условные вероятности при помощи матрицы объединения находятся следующим образом

$$p(a_i/b_j)=\frac{p(a_i,b_j)}{\sum_i p(a_i,b_j)}=\frac{p(a_i,b_j)}{p(b_j)}.$$

$$p(b_j/a_i)=\frac{p(a_i,b_j)}{\sum_j p(a_i,b_j)}=\frac{p(a_i,b_j)}{p(a_i)}.$$

Взаимная энтропия ансамблей А и В про помощи матрицы объединения вычисляется путем последовательного суммирования по строкам или по столбцам всех вероятностей вида *p(a,b)*, умноженных на логарифм этих же вероятностей

$$H(A,B)=-\sum_i\sum_j p(a_i,b_j)\log p(a_i,b_j)$$ бит/два символа

Размерность «бит/два символа» объясняется тем, что взаимная энтропия представляет собой неопределенность возникновения пары символов, то есть неопределенность на два символа.

Взаимная энтропия передаваемого ансамбля A и принимаемого ансамбля B равна сумме безусловной энтропии H(A) и условной энтропии H(B/A)

$$H(A,B) = H(A) + H(B/A)$$

H(B/A) в данном случае представляет ту добавочную информацию, которую дает сообщение B после того как стала известна информация, содержащаяся в сообщении A.

Таким образом, условная энтропия представляет собой неопределенность того, что при приеме *в* было послано *а*, а взаимная энтропия отражает неопределенность возникновения пары вида *ав*.

Так как взаимная энтропия есть неопределенность относительно пары символов, сигналов, состояний, в общем случае, относительно пары элементов взаимосвязанных выборочных пространств A и B, то не имеет значения имеет ли эта пара вид *ав* или *ва,* так как неопределенность возникновения такого сочетания – одинакова. *Взаимная энтропия обладает свойством симметрии.*

$$H(A,B) = H(B,A)$$

Если построена матрица вероятностей p(a,в), описывающая взаимосвязь двух произвольных выборочных пространств, в частности взаимосвязь входа и выхода шумящего канала связи, то остальные информационные характеристики могут не задаваться, так как матрица объединения обладает информационной полнотой.

Определение. Набор информационных характеристик произвольного канала связи считается информационно полным, если с помощью этого набора, путем алгебраических преобразований, можно получить любую другую информационную характеристику того же канала связи.

Сигналы и помехи в системах.

Теория информации представляет собой ветвь статистической теории связи. Информация, выраженная в определенной форме, представляют собой сообщение.

Под сообщением понимают совокупность знаков или первичных сигналов содержащих информацию, т.е. сообщение - это информация представленная в какой-либо форме. Пример сообщений: текст телеграммы, данные на выходе ЭВМ, речь, музыка и т.д. Иначе говоря, сообщение – это, то что подлежит передаче.

Для того чтобы сообщение можно было передать получателю, необходимо воспользоваться некоторым физическим процессом, способным с той или иной скоростью распространяться от источника к получателю сообщения. Изменяющийся во времени физический процесс, отражающий передаваемое сообщение *называется сигналом*.

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.

Сигнал - ест материальный носитель информации, средство перенесения информации в пространстве и времени.

Сообщения могут быть функциями времени (когда информация представлена в виде первичных сигналов: речь, музыка) и не является ими (когда информация представлена в виде совокупности знаков).

Поскольку сигналы служат для переноса информации в пространстве и времени, для образования сигналов могут использоваться только объекты, состояния которых достаточно устойчивы по отношению к течению времени или к изменению положения в пространстве. С этой точки зрения сигналы делятся на два типа.

К первому типу относятся сигналы, являющиеся стабильными состояниями физических объектов (книга, фотография, магнитная запись и т.д.), такие сигналы называются *статическими.*

Ко второму типу относятся сигналы, в качестве которых используются динамические состояния силовых полей. Примерами таких сигналов могут служить звуки, световые и радиосигналы. Сигналы указанного типа называются *динамическими.*

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.

С дискретными сигналами мы встречаемся в цифровой вычислительной технике, в телеграфии. Так при передаче обычной телеграммы сообщением является текст телеграммы, элементами сообщения – буквы, сигналами – кодовые комбинации, соответствующие этим буквам.

Непрерывное сообщение – это некоторая физическая величина, принимающая любые значения в заданном интервале.

Сигнал всегда является функцией времени. В зависимости от того, какие значения могут принимать аргумент (время *t*) и уровни сигналов их делят на 4 типа.

1) *Непрерывный или аналоговый сигналы* (случайные сигналы этого типа называются непрерывными случайными процессами). Они определены для всех моментов времени и могут принимать все значения из заданного диапазона. Чаще всего физические процессы, порождающие сигналы являются непрерывными. Этим и объясняется второе название сигналов данного типа аналоговый т.е. аналогичные порождающим процессам.

2) *Дискретизированный или дискретно непрерывные сигналы* (случайные сигналы этого типа называют процессами с дискретным временем или непрерывными случайными последовательностями). Они определены лишь в отдельные моменты времени и могут принимать любые значения уровня. Временной интервал t между соседними отсчетами называется шагом дискретизации. Часто такие сигналы называют дискретными по времени.

3) *Дискретные по уровню или квантованные сигналы* (случайные сигналы этого типа называют дискретными случайными процессами). Они определены для всех моментов времени и принимают лишь разрешенные значения уровней отделенные от друг друга на величину шага квантования $x=x_{k+1}+x_k$

4) *Дискретные по уровню и по времени сигналы* (случайные сигналы этого типа называют дискретными случайными последовательностями). Они определены лишь в отдельные разрешенные моменты времени и могут принимать лишь разрешенные значения уровней.

Помехи и искажения.

Передача сигналов по реальным каналам всегда сопровождается изменениями (преобразованиями) этих сигналов, с точки зрения передачи информации по каналу важно подразделение преобразований сигнала на обратимые и необратимые. Обратимые преобразования не влекут за собой потери информации. При необратимых преобразованиях потери информации неизбежны. Поэтому для обратимых преобразований сигнала также часто используется термин ***"искажение"*** а необратимые преобразования называют ***"помехами".***

Линейные искажения сигналов появляются в линейном инерционном четырехполюснике с постоянными параметрами из-за наличия в нем реактивных элементов. При линейных искажениях нарушаются существующие частотные и фазовые соотношения между отдельными составляющими сигнала и форма сигналов. Для отсутствия искажений необходимо, чтобы модуль коэффициента передачи и времени запаздывания для всех составляющих были одинаковы. Нелинейными называют искажения сигналов, которые возникают в нелинейных безынерционных четырехполюсниках с постоянными параметрами *из-за* нелинейности характеристик активных элементов: транзисторов, диодов и др. В результате нелинейных искажений спектр сигналов расширяется, в них появляются дополнительные компоненты, растут уровни взаимных помех в каналах.

Как линейные, так и нелинейные искажения обусловлены известными характеристиками каналами поэтому, в принципе, могут быть устранены или уменьшены путем надлежащей коррекции.

Следует четко отделить искажения то помех, имеющих случайный характер. Помехи заранее неизвестны и поэтому не могут быть полностью устранены.

Под *помехой* понимается любое воздействие, накладывающееся на полезный сигнал и затрудняющее его прием. Помехи весьма разнообразны как по своему происхождению, так и по физическим свойствам.

Помехи - это электрические возмущения, возникающие в самой аппаратуре или попадающие в нее извне. Наиболее распространенными являются флуктуационные, или случайные помехи (например, тепловые шумы, возникающие в оборудовании). Они представляют собой последовательность импульсов, имеющих случайную амплитуду и следующих друг за другом через различные промежутки времени.

В радиоканалах наиболее распространенными являются атмосферные помехи, обусловленные электрическими процессами в атмосфере и, прежде всего, грозовыми разрядами. Энергия этих помех сосредоточена главным образом в области длинных и средних волн. Сильные помехи создаются также промышленными установками. Это так называемые индустриальные помехи, возникающие из-за резких изменений тока в электрических цепях всевозможных электроустройств. Сюда относятся помехи от электротранспорта, электрических моторов, систем зажигания двигателей и т.п.

Распространенным видом помех являются помехи от посторонних радиостанций и каналов. Этот вид помех обусловлен нарушением регламента распределения рабочих частот, недостаточной стабильностью частот и плохой фильтрацией гармоник сигнала.

В проводных каналах связи основным видом помех являются импульсные шумы и прерывания связи. Прерывание связи есть явление, при котором сигнал в линии резко затухает или совсем исчезает. Такие прерывания могут быть вызваны различными причинами, из которых наиболее частыми являются нарушения контактов в реле, разъемах и т.п. Практически в любом диапазоне частот имеют место внутренние шумы аппаратуры, обусловленные хаотическим движением носителей заряда в усилительных приборах, сопротивлениях и других элементах аппаратуры.

По своей электрической структуре помехи - это колебания, сходные с сигналами, но беспорядочные и, конечно, ненужные. В приемнике помехи могут подавить информационный сигнал, то есть ослабить настолько, что приемник или не обнаружит его, или воспримет как ложный. В частности, в двоичном канале "единица" может перейти в "ноль" и наоборот. При равнозначной вероятности появления таких переходов канал связи считается симметричным, в противном случае - несимметричным. В реальных условиях каналы связи обычно бывают несимметричными.

Наличие помех в системе связи приводит к большому числу неверно выполняемых вычислений неправильному чтению командных и управляющих посылок, снижению эффективности сети.

Трудности борьбы с помехами заключаются в беспорядочности, нерегулярности и в структурном сходстве помех с информационными сигналами. Поэтому защита информации от ошибок и вредного влияния помех имеет огромное практическое значение и является одной из важнейших проблем современной теории и техники связи.

Существует несколько источников возникновения помех. Например, атмосферные помехи возникают вследствие электрических возмущений в земной атмосфере. Космические помехи могут прийти с Солнца или других звезд, которые излучают электромагнитную энергию в очень широком частотном спектре. Помехи можно также обнаружить в проволоке-проводнике или коаксиальном проводнике вследствие того, что случайное движение электронов в проводнике приводит к образованию тепловой энергии.

Чтобы успешно бороться с тепловым шумом (а также с другими видами шумов, например разрядными помехами флуктуациями мощности и так далее), приемники в системах связи должны проверять данные и в случаях обнаружения "нарушений" запрашивать повторную передачу. "Нарушения" или ошибки можно широко классифицировать как случайные, импульсные и смешанные. В каналах со случайными ошибками для каждого бита данных существует вероятность P

неправильного приема и Р-1 правильного приема. Ошибки происходят случайно в блоках принятых данных.

Большинство каналов с вещественными носителями (а также спутниковые каналы) подвержены случайным ошибкам. Каналы с импульсными ошибками демонстрируют состояние, свободное от ошибок, большую часть времени, но иногда появляются групповые или разовые ошибки. Объектом таких ошибок являются радиосигналы, так же как кабели и провода, например телефонные каналы из витых проводных пар. Проблема канального шума обусловлена свойствами самого канала и никогда не может быть устранена полностью.

Для рассмотрения помех в непрерывных каналах выходной сигнал представляют в виде:

$$x(t,\tau) = \mu(t)S[t-\tau(t)] + \xi(t)$$

где *S(t)* — входной сигнал; $\mu(t)$ и $\xi(t)$ — соответственно мультипликативная и аддитивная помехи; τ *(t)* — задержка сигнала в канале.

Мультипликативные помехи обусловлены случайными изменениями коэффициента передачи канала из-за изменения характеристик среды, в которой распространяются сигналы, и коэффициентом усиления схем при изменении питающих напряжений, из-за замираний сигналов в результате интерференции и различного затухания сигналов при многолучевом распространении радиоволн. К мультипликативным помехам следует отнести и "квантовый шум" лазеров, применяемых в оптических системах передачи и обработки информации. "Квантовый шум" лазера вызван дискретной природой светового излучения и зависит от интенсивности излучения, т.е. от самого полезного сигнала.

Аддитивные помехи обусловлены флуктуационными явлениями (случайными колебаниями тока и напряжения), связанными с тепловыми процессами в проводах, резисторах, транзисторах и других элементах схем, наводками под действием атмосферных явлений (грозовые разряды и т. д.) и индустриальных процессов (работа промышленных установок, других линий связи и т. д.).

Математически аддитивную помеху можно записать в виде:

$$x(t) = S(t) + \xi(t)$$

Аддитивные помехи делят на: сосредоточенные и флуктуационные. Сосредоточенные аддитивные помехи отличаются сосредоточенностью энергии помех в полосе частот (узкополосные помехи) или на отрезке времени (импульсные помехи).

Узкополосные помехи в основном обусловлены действием посторонних источников — ширина спектра этих помех сравнима или значительно меньше ширины спектра полезных сигналов. Узкополосные помехи как помехи от соседних станций характерны для передачи информации по радиоканалам. Борьба с узкополосными аддитивными помехами ведется методами улучшения технических характеристик устройств приема, и обработки сигналов.

Импульсные помехи — это случайные последовательности импульсов, создаваемые промышленными установками и атмосферными источниками сигналов. Эти помехи характеризуются широким энергетическим спектром. Ширина их спектра, как известно, обратно пропорциональна длительности импульсов. Энергия спектральных составляющих импульсных помех падает в области сверхнизких и сверхвысоких частот.

Флуктуационная аддитивная помеха характеризуется "размытостью" энергии спектра в широком диапазоне частот Она обусловлена главным образом внутренними шумами элементов аппаратуры (тепловой шум, дробовой эффект и т. д.)

Флуктуационную помеху из-за "внутренней" природы невозможно устранить, можно лишь учесть ее характеристики при синтезе такой оптимальной системы, в которой наличие флуктуационной помехи меньше всего сказывается на качестве передачи информации.

Математическими моделями сосредоточенных аддитивных помех являются узкополосные случайные сигналы и случайные последовательности импульсов. Математической моделью флуктуационной аддитивной помехи служит гауссовский "белый шум".

Передача информации

Передача сообщения от источника к приемнику всегда связывается с некоторым процессом, происходящим в материальной среде – это условие является обязательным, поскольку сама информация материальным объектом или формой существования материи не является.

Сообщения передаются от объекта к адресату при помощи совокупности технических средств, которые образуют систему передачи информации.

Способов передачи информации существует множество: почта, телефон, радио, телевидение, компьютерные сети и пр. Однако при всем разнообразии конкретной реализации способов связи в них можно выделить общие элементы, представленные на рис.3.

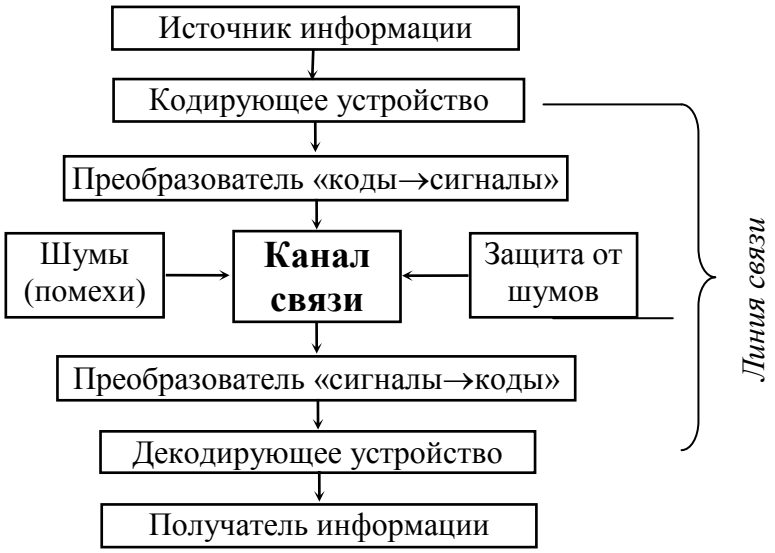


Рис. 3. Общая схема передачи информации

Понимать схему нужно следующим образом. Источник, порождающий информацию, для передачи должен представить ее в виде сообщения, т.е. последовательности сигналов. При этом для представления информации он должен использовать некоторую систему кодирования. Устройство, выполняющее операцию кодирования информации, может являться подсистемой источника (например, наш мозг порождает информацию и он же кодирует эту информацию с помощью языка, а затем представляет в виде речевого сообщения посредством органов речи; компьютер обрабатывает и хранит информацию в двоичном представлении, но при выводе ее на экран монитора производит ее перекодировку к виду, удобному пользователю). Возможна ситуация, когда кодирующее устройство оказывается внешним по отношению к источнику информации, например, телеграфный аппарат или компьютер по отношению к работающему на нем оператору. Далее коды должны быть переведены в последовательность материальных сигналов, т.е. помещены на материальный носитель – эту операцию выполняет преобразователь. Преобразователь может быть совмещен с кодирующим устройством (например, телеграфный аппарат), но может быть и самостоятельным элементом линии связи (например, модем, преобразующий электрические дискретные сигналы с частотой компьютера в аналоговые сигналы с частотой, на которой их затухание в телефонных линиях будет наименьшим). К преобразователям относят также устройства, которые переводят сообщение с одного носителя на другой, например, мегафон или телефонный аппарат, преобразующие голосовые сигналы в электрические; радиопередатчик, преобразующие голосовые сигналы в радиоволны; телекамера, преобразующая изображение в последовательность электрических импульсов. В общем случае при преобразовании выходные сигналы не полностью воспроизводят все особенности сообщения на входе, а лишь его существенные стороны. Например, полоса пропускания частот при телефонной связи от 300 до 3400 Гц, в то время, как частоты человеческого голоса лежат в интервале 16–20000 Гц (т.е. телефонные линии «обрезают» высокие частоты голоса, что приводит к его искажениям); в черно-белом телевидении при преобразовании терялся цвет изображения. Именно в связи с этим встает задача выработки такого способа кодирования сообщения, который обеспечивал бы возможно более полное представление исходной информации и, в то же время, был согласован со скоростью передачи информации по данной линии связи.

После преобразователя сигналы поступают и распространяются по каналу связи. Понятие канала связи включает в себя ***материальную среду***, а также ***физический*** или иной ***процесс***, посредством которого осуществляется передача сообщения, т.е. распространение сигналов в пространстве с течением времени.

Любой реальный канал связи подвержен внешним воздействиям, а также в нем могут происходить внутренние процессы, в результате которых искажаются передаваемые сигналы и, следовательно, связанное с ними сообщение. Такие воздействия называются шумами (помехами). Источники помех могут быть внешними, например, так называемые «наводки» от мощных потребителей электричества или атмосферных явлений, приводящие к появлению помех в радиосвязи; одновременное действие нескольких близкорасположенных однотипных источников (одновременный разговор нескольких человек). К помехам могут приводить и внутренние особенности данного канала, например, физические неоднородности носителя; паразитные явления в шинах; процессы затухания сигнала в линии связи из-за большой удаленности. Если уровень помех оказывается соизмерим с интенсивностью несущего сигнала, то передача информации по данному каналу оказывается вообще невозможной. Однако и при относительно низких уровнях шумов они могут приводить к искажениям передаваемого сигнала. Существуют и применяются методы защиты от помех, например, экранирование электрических линий связей; улучшение избирательности приемного устройства и т.д. Другим способом защиты от помех является использование специальных методов кодирования информации, о чем речь пойдет ниже.

После прохождения сообщения по каналу связи сигналы с помощью приемного преобразователя переводятся в последовательность кодов, которые декодирующим устройством представляются в форме, необходимой приемнику информации. На этапе приема, как и при передаче, преобразователь может быть совмещен с декодирующим устройством (например, радиоприемник или телевизор) или существовать самостоятельно (например, модем).

Не следует путать канал связи и линию связи. Канал связи - совокупность технических средств, предназначенных для передачи информации от объекта к адресату; линия связи - среда, в которой распространяются сигналы, несущие информацию.

В зависимости от линий связи каналы связи делятся на проводные (металл), радио (воздух), оптические (световой луч), гидроакустические (вода).

Кабельные и воздушные линии связи на основе металлических проводников

Передача информации при помощи проводов – наиболее древнее и по сей день наиболее распространенное средство связи объекта с адресатом. Проводные каналы связи бывают одностороннего (симплексный) и двустороннего (дуплексный) действий.

Проводные каналы связи используются обычно в диапазоне от долей герца до 12 кГц. Частотный диапазон проводного канала связи ограничен по той причине, что с увеличением частоты возрастает активное сопротивление провода под влиянием поверхностного эффекта. Максимальная протяженность проводного канала связи определяется затуханиями в нем, которые, в свою очередь, зависят от параметров линии связи: активного сопротивления, индуктивности, емкости и проводимости изоляции проводов. Эти параметры меняются в зависимости от времени года (активное сопротивление зимой минимальное, летом - максимальное), расстояния между проводами (чем больше расстояние, тем больше емкость), диаметра проводов (чем больше диаметр, тем больше индуктивность), влажности воздуха (чем больше влажность, тем больше проводимость изоляции проводов). Поэтому при передаче информации на большие расстояния возникает необходимость в промежуточной аппаратуре, которая осуществляла бы усиление и частичную регенерацию импульсов, а также коррекцию их частотных искажений.

Для передачи информации применяют телефонные и телеграфные каналы, в которых дополнительная аппаратура установлена лишь на передающем и приемном концах. При этом чаще стараются использовать подземные (кабельные) каналы связи, так как они меньше зависят от внешних условий и имеют стабильные параметры. Кроме того, у кабелей значительно лучше частотные характеристики. Разработаны специальные коаксиальные кабели, которые могут использоваться в диапазоне от 60 до 12000 кГц, что позволяет передавать по ним даже телевизионные программы.

Существующие типы линий связи (ЛС) в зависимости от используемой среды распространения сигналов принято делить на проводные и линии в атмосфере (радиолинии).

К линиям связи предъявляются следующие основные требования: осуществление связи на практически требуемые расстояния; широкополосность и пригодность для передачи различных видов сообщений; защищенность цепей от взаимных влияний и внешних помех, а также от физических воздействий (атмосферных явлений, коррозии и пр.); стабильность параметров линии, устойчивость и надежность связи; экономичность системы связи в целом.

В простейшем случае проводная ЛС - физическая цепь, образуемая парой металлических проводников. Кабельные ЛС (кабели связи) образованы проводами с изоляционными покрытиями, помещенными в защитные оболочки. По конструкции и взаимному расположению проводников различают *симметричные* (СК) и *коаксиальные* (КК) кабели связи.

Симметричная цепь состоит из двух совершенно одинаковых в электрическом и конструктивном отношениях изолированных проводников. В зарубежных источниках СК часто называют "витая пара" (TP - twisted pair). Различают экранированные (shielded) и неэкранированные (unshielded) СК.

Коаксиальная цепь представляет собой два цилиндра с совмещенной осью, причем один цилиндр - сплошной внутренний проводник, концентрически расположен внутри другого полого цилиндра (внешнего проводника). Проводники изолированы друг от друга диэлектрическим материалом.

В настоящее время выпускается широкая номенклатура кабелей, отличающихся в зависимости от назначения, области применения, условий прокладки и эксплуатации и пр.

Воздушные ЛС (ВЛС) не имеют изолирующего покрытия между проводниками, роль изолятора играет слой воздуха. Проводники выполняются, в основном, из биметаллической сталемедной (сталеалюминевой) проволоки. Внутренний диаметр стальной проволоки обычно составляет 1.2...4 мм, толщина внешнего слоя меди (алюминия) - 0.04...0.2 мм. Проволока подвешивается на деревянных или железобетонных опорах с помощью фарфоровых изоляторов. Используемый частотный диапазон ВЛС не превышает 150 кГц.

Кабельные системы

В настоящее время проводные линии связи широко используются при построении локальных сетей. Данные линии связи стандартизированы и обычно называются *структурированной кабельной проводкой* или *кабельной системой*. Известны кабельные системы категорий 3, 4, 5 стандартов EIA/TIA-568, TSB-36, TSB-40 специального подкомитета TR41.8.1.

Длина горизонтальных кабелей - не более 90 м независимо от типа кабеля.

К применению допускаются кабели четырех типов: 4-парный из неэкранированных витых пар с волновым сопротивлением 100 Ом; 2-парный из экранированных витых пар с волновым сопротивлением 150 Ом; коаксиальный с волновым сопротивлением 50 Ом; волоконно-оптический с волокнами диаметром 62,5/125 мкм;

Типы соединителей: модульный 8-контактный RJ-45; 4-контактный по стандарту IEEE 802.5; коаксиальный BNC; оптический не определен.

Радиолинии

В радиолиниях связи средой распространения электромагнитных волн в подавляющем большинстве случаев (за исключением случая связи между космическими аппаратами) является атмосфера Земли.

Строение атмосферы более сложно и приведенное деление на тропосферу, стратосферу и ионосферу достаточно условно. Высота слоев приведена приблизительно и различна для разных географических точек Земли. В тропосфере сосредоточено около 80% массы атмосферы и около 20% - в стратосфере. Плотность атмосферы в ионосфере крайне мала, граница между ионосферой и космическим пространством является условным понятием, так как следы атмосферы встречаются даже на высотах более 400 км. Считается, что плотные слои атмосферы заканчиваются на высоте около 120 км.

Типичным примером радиолиний являются линии сетей передачи сообщений массового характера (сети телевизионного и радиовещания). Радиолиния может содержать несколько промежуточных переприемных станций. Так строятся линии радиорелейных систем передачи.

Радиоволны, излучаемые передающей антенной, прежде чем попасть в приемную антенну, проходят в общем случае сложный путь. На величину напряженности поля в точке приема оказывает влияние множество факторов. Основные из них:

- отражение электромагнитных волн от поверхности Земли;
- преломление (отражение) в ионизированных слоях атмосферы (ионосфере);
- рассеяние на диэлектрических неоднородностях нижних слоев атмосферы (тропосфере);
- дифракция на сферической выпуклости Земли;

Также напряженность поля в точке приема зависит от длины волны, освещенности земной атмосферы Солнцем и ряда других факторов.

Волоконно-оптические линии связи

Волоконно-оптические линии связи (ВОЛС) имеют ряд существенных преимуществ по сравнению с линиями связи на основе металлических кабелей. К ним относятся: большая пропускная способность, малое затухание, малые масса и габариты, высокая помехозащищенность, надежная техника безопасности, практически отсутствующие взаимные влияния, малая стоимость из-за отсутствия в конструкции цветных металлов.

В ВОЛС применяют электромагнитные волны оптического диапазона. Напомним, что видимое оптическое излучение лежит в диапазоне длин волн 380...760 нм. Практическое применение в ВОЛС получил *инфракрасный* диапазон, т.е. излучение с длиной волны более 760 нм.

Принцип распространения оптического излучения вдоль оптического волокна (ОВ) основан на отражении от границы сред с разными показателями преломления. Оптическое волокно изготавливается из кварцевого стекла в виде цилиндров с совмещенными осями и различными коэффициентами преломления. Внутренний цилиндр называется *сердцевиной* ОВ, а внешний слой - *оболочкой* ОВ.

В зависимости от вида профиля показателя преломления сердцевины различают *ступенчатые* и *градиентные* ОВ. У ступенчатых ОВ показатель преломления сердцевины постоянен, а у градиентных ОВ показатель преломления сердцевины плавно меняется вдоль радиуса от максимального значения на оси до значения показателя преломления оболочки.

В ОВ может одновременно существовать несколько типов волн (мод). В зависимости от модовых характеристик ОВ со ступенчатым профилем преломления делятся на два вида: *многомодовые* и *одномодовые*.

Одномодовый режим реализуется при $V < 2.405$. Заранее определенными и сравнительно малыми величинами являются рабочая длина волны и разность показателей преломления. Диаметр сердцевины одномодовых волокон также является малой величиной и составляет 5...15 (обычно 9 или 10) мкм.

Для многомодовых волокон диаметр сердцевины составляет около 50 (обычно 50 или 62,5) мкм. Диаметр оболочки у всех типов ОВ 125 мкм. Диаметр защитного покрытия - 500 мкм. Наружный диаметр кабеля с числом ОВ от 2..32 с учетом всех защитных оболочек и элементов обычно составляет 5..17 мм.

Затухание ОВ неоднородно для разных длин волн. Коэффициент затухания ОВ зависит от рабочей длины волны. Данная зависимость имеет три минимума, называемые *окнами прозрачности*. Исторически первым было освоено первое окно прозрачности на рабочей длине волны 0.85 мкм.

Первые полупроводниковые излучатели (лазеры и светодиоды) и фотоприемники были разработаны именно для данной длины волны. Коэффициент затухания в первом окне значителен и составляет единицы дБ/км. Позднее были созданы излучатели и фотоприемники, способные работать на больших длинах волн (1,3 и 1,55 мкм). Современные системы связи обычно используют второе или третье окно с малыми коэффициентами затухания. Современная технология позволяет получить ОВ с коэффициентом затухания порядка сотых долей дБ/км.

Гидроакустический канал связи стоит несколько обособлено от перечисленных выше каналов, так как передача информации по нему связана не столько с привычной передачей электрических сигналов или электромагнитных волн, сколько с передачей упругих колебаний водной среды.

Особенностью гидроакустического канала является неоднородность среды, образующей линию связи. Присутствие в морской воде солей обуславливает существование в ней свободных и связанных ионов, число которых изменяется при распространении акустической волны под влиянием сжатия и разрежения, вследствие чего она теряет часть энергии. На свойство воды как звукопроводящей среды существенно влияет степень ее нагретости и солености. Поэтому в различных слоях моря условия распространения акустической волны не одинаковы. Кроме того, при распространении звука в воде происходит отражение волн от поверхности и дна моря. Отраженные волны искажают информационные посылки, а также вызывают реверберацию (послезвучание).

Перечисленные факторы обусловливают специфические требования к помехоустойчивости и надежности кода, передаваемых по гидроакустическому каналу.

Согласование физических характеристик сигнала и канала.

Конкретный канал связи обладает определенными физическими параметрами, от которых зависит возможность передачи по нему тех или иных сигналов. Независимо от назначения непрерывного канала его можно характеризовать тремя основными параметрами: время, в течении которого он представляется для передачи сигнала Tk, шириной полосы пропускания сигнала Fk и допустимым превышением сигнала над помехами Hk . Превышение допустимого превышения сигнала Hk характеризуется разностью максимально допустимого сигнала в канале Pu max и уровня помех Ps. Для проводных каналов превышение в основном определяется пробивным напряжением и уровнем перекрестных помех, для радиоканалов - возможностями выявления сигнала на соответствующих расстояниях.

Произведение указанных основных параметров канала связи принято называть объемом канала и обозначать Vk

Vk = Tk Fk Hk

При оценке возможностей передачи сигнала по каналу с заданными физическими характеристиками также ограничиваются рассмотрением трех основных параметров сигнала: его длительности Tc ширины спектра Fc и превышением над помехой Hc причем

Hc =log(Pu /P s)

где Ps - средняя мощность помехи в канале.

Превышение Hc связано с возможностями передатчика и дальностью передачи. Чем больше превышение Hc, тем меньше вероятность ошибочного приема. Аналогично объему канала вводится понятие объема V 4c 0 передаваемого сигнала:

Vc=TcFcHc

Как объем сигнала, так и объем канала могут быть представлены в трехмерном пространстве с соответствующими координатами T, F, H

Необходимым условием принципиальной возможности неискаженной передачи сигнала выполнения соотношения

$$V_c \le V_k$$

При этом, однако могут потребоваться преобразования для обеспечения достаточных условий передачи, а именно:

$$T_c \le T_k;\ F_c \le F_k;\ H_c \le H_k$$

Когда канал имеет меньшую полосу пропускания, чем практическая ширина спектра, подлежащего передаче сигнала, последнюю можно уменьшить за счет увеличения длительности сигнала. Объем сигнала при этом сохраняется неизменным. Практически также преобразование можно осуществить, например, посредством записи сигнала на магнитную ленту с высокой скоростью и последующего воспроизведения со скоростью, при которой ширина его спектра равна полосе пропускания канала.

Если, наоборот, широкополосный канал представляется на время меньшее длительности сигнала, то согласование осуществляется за счет расширения спектра сигнала. Для реализации

также может использоваться накопитель на магнитной ленте, однако в данном случае скорость воспроизведения должна быть выше скорости записи.

При низком допустимом уровне превышения сигнала в канале преобразование заключается в уменьшении уровня превышения передаваемого сигнала с одновременным увеличением его длительности путем многократного повторения передачи. Возможны и другие виды преобразования.

Информационные характеристики источника дискретных сообщений и канала связи

1. Производительность ИС – величина, определяемая КИ, выдаваемой ИС в единицу времени:

$$\Pi(X) \equiv \frac{I_0(X)}{T} \qquad \left[\frac{бит}{с}\right].$$

Если ИС с Э $H(X)$ за время Т выдает n сообщений каждая длительностью τ_c, то общее КИ: $I_0(X) = n * H(X).$

Т= $n * \tau_c$, следовательно $\Pi(X) = \frac{n * H(X)}{n * \tau_c} = \frac{H(X)}{\tau_c} \qquad \left[\frac{бит}{с}\right].$

При $\tau_c = \text{const}$ $\Pi(X)$ зависит только от Э ИС, т. е. от статистической структуры источника:

$$\max \Pi(X) = \frac{\max H(X)}{\tau_c} = \frac{\log_2 m}{\tau_c}.$$

Таким образом, максимальной производительностью обладают источники, сообщения которых независимы и равновероятны. Это оптимальные источники. Чем больше отличается распределение в источнике от независимого и равновероятного, тем меньше производительность ИС, т. е. источник содержит сообщения с меньшей информацией.

2. Избыточность ИС – величина, показывающая, насколько сообщение от источника с Э $H(X)$ отличается по количеству содержащейся в них информации от оптимального источника:

$$\mathrm{R(X)} \equiv 1 - \frac{H(X)}{H\max(X)} = 1 - \frac{H(X)}{\log_2 m}.$$

Избыточность заключена в $0 \le \mathrm{R(X)} \le 1$.

При оптимальном источнике R(X) = 0, при $H(X) = 0$ R(X) = 1.

Наличие в источнике избыточности показывает, что сообщение от этого источника не является оптимальным, и производительность источника и скорость передачи информации меньше максимально возможных.

Для русского языка избыточность $\max H(X) = \log_2 32 = 5 \quad \left[\frac{бит}{символ}\right]$. Численное значение Э с учетом неравномерного появления букв:

$$H(X) = 4.42 \quad \left[\frac{бит}{символ}\right],$$

$$R(X) = 1 - \frac{4.42}{5} = 0.116.$$

Вероятности появления отдельных букв с учетом предшествующих букв уменьшает Э.

С учетом двухбуквенных сочетаний $H(X)$=3.52, R(X)=0.3.

Общим способом уменьшения избыточности ИС является переход от первичного ИС (алфавита) к искусственному вторичному источнику (алфавиту) с R(X) = 0, статистическая структура сообщений в котором оптимальна.

3. Техническая скорость передачи C_m определяется числом посылок сигнала, проходящих по каналу в единицу времени (с):

$$C_m = \frac{n}{T} = \frac{n}{n * \tau_c} = \frac{1}{\tau_c} \qquad \text{[baud] или [бод].}$$

1 бод – 1 посылка в секунду.

Техническая скорость выбирается из условия согласования ширины спектра сигнала и полосы пропускания канала: $\frac{1}{\tau_c} = \Delta F_c = \Delta F_k$.

Техническая скорость не зависит от информационных характеристик ИС.

4. Информационная скорость передачи C_u определяется количеством информации, проходящей по каналу передачи в единицу времени:

$$C_u \equiv \frac{I_0(X)}{T} = \frac{n * H(X)}{T} = C_m * H(X) \qquad \left[\frac{бит}{с}\right].$$

Информационная скорость является более полной характеристикой канала сообщений, чем техническая скорость.

5. Пропускная способность канала C_0. Пропускная способность определяется как предельное значение информационной скорости передачи, достижимое в определенных условиях передачи.

$$C_0 = \max C_u = \max[C_m * H(X)] \qquad \left[\frac{бит}{с}\right].$$

Пропускная способность наиболее полно характеризует возможности передачи по каналу и зависит как от производительности ИС, так и от мощности помех, действующих в канале связи.

а) При отсутствии помех, действующих в канале связи (телефонная линия), КИ, проходящей по каналу, определяется понятием собственной информации, и информационная скорость определяется как отношение общего КИ ко времени передачи:

$$C_r = \frac{I_0(X)}{T} = \frac{nH(X)}{n\tau_c} = \frac{H(X)}{\tau_c} = \Pi(X).$$

Информационная скорость численно совпадает с производительностью ИС:

$$C_0 = \max C_r = \max \Pi(X) = C_m \log_2 m.$$

Пропускная способность в этом случае реализуется тогда, когда производительность источника максимальна, т. е. в источниках независимых и равновероятных сообщений.

Если ИС является двоичным m = 2, то $C_0 = C_m \log_2 2 = C_m$.

Формально в этом случае пропускная способность равна технической скорости передачи.

В большинстве реальных случаев ИС обладают избыточностью, т. е. R(X) > 0, следовательно, $H(X) < \log_2 m$. Э меньше максимально возможной:

$$C_u = \Pi(X) < \frac{\log_2 m}{\tau_c} = C_0.$$

Максимальное значение информационной скорости меньше пропускной способности. Пропускная способность не реализуется, канал используется не полностью.

Для повышения информационной скорости передачи следует стремиться уменьшить избыточность ИС, что достигается путем преобразования сообщений источника в оптимальные (независимые, равновероятные).

К. Шеннон доказал несколько теорем, первая из которых применима к каналу без помех.

1 Теорема Шеннона. Если производительность ИС меньше пропускной способности канала, т. е.

$$\Pi(X) < \frac{\log_2 m}{\tau_c},$$

то сообщение от источника можно преобразовать так, чтобы передавать их сколь угодно точно по дискретному идеальному каналу без шумов со скоростью достаточно близкой к пропускной способности.

Комментарий. Шеннон дает идеологическую основу для преобразования ИДС с избыточностью в оптимальный источник. В теории информации для этого используется так называемое эф-

фективное (статистическое) кодирование, в результате применения которого $C_u = C_0$, избыточность равна нулю, производительность максимальна.

Технический комментарий. Согласно этой теореме, с точки зрения технической реализации существует способ кодирования и декодирования сообщений, при котором вероятность ошибочного декодирования сколь угодно мала. Кроме того, Шенноном доказано, что, если производительность источника превышает пропускную способность $\Pi(X) > C_0$, таких способов кодирования не существует.

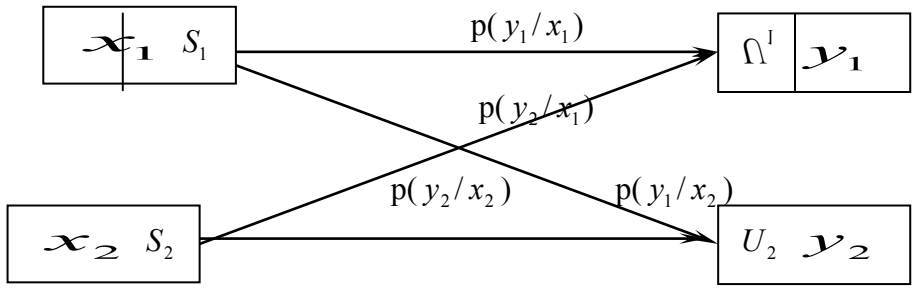
b) В реальных каналах с помехами КИ определяется понятием взаимной информации, поэтому пропускная способность при условии $C_m = \text{const}$:

$$C_0 = \frac{I(Y,X)}{T} = \frac{I(X,Y)}{T} = C_m * \max[H(X) - H(X/Y)] = C_m * \max[H(Y) - H(Y/X)].$$

Величины $H(X)$ и $H(X/Y)$, определяющие пропускную способность, зависят от многих факторов, в частности от статистической структуры ИС, от вида применяемой модуляции, от мощности помех, действующих в канале и т. д., т. е. зависит от группы факторов, влияющих на вероятность ошибки при приеме.

Введем выражение для пропускной способности бинарного симметричного канала информации без памяти (любая компьютерная сеть). Канал не обладает памятью, если нет зависимости от предыдущих символов.

Модель канала



где x - сообщение, S - сигнал, U- сигнал с шумом, y - принятое сообщение.

Вероятности правильного приема: $p(y_1/x_1)$ - вероятность принятия 0 при посылке 0, $p(y_2/x_2)$ - вероятность принятия 1 при посылке 1. Пусть $p_{ош} = p(y_2/x_1) = p(y_1/x_2) \equiv p$, $p_{прав} = p(y_1/x_1) = p(y_2/x_2) = 1 - p$. Такой канал называется симметричным.

Как связать вероятность ошибочного приема

$$C_0 = C_m * \max[H(Y) - H(Y/X)],$$

$$H(Y/X) = M\{H(Y/x_i)\} = \sum_{i=1}^{2} p(x_i) H(Y/x_i) = -\sum_{i=1}^{2} p(x_i) \sum_{j=1}^{2} p(y_j/x_i) \log_2 p(y_j/x_i) =$$

$$= -p(0)[(1-p)\log_2(1-p) + p\log_2 p] - p(1)[p\log_2 p + (1-p)\log_2(1-p)] =$$

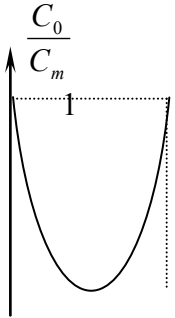
$$= -[p(0) + p(1)][p\log_2 p + (1-p)\log_2(1-p)] = -p\log_2 p - (1-p)\log_2(1-p).$$

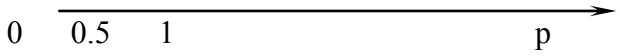
Считается, что p задано системой, следовательно, требуется максимизировать $H(Y)$. Если канал бинарный, сообщения равновероятны и независимы, то $\max H(Y) = 1$,

$$C_0 = C_m [p\log_2 p + (1-p)\log_2(1-p)].$$

Для источника с m состояниями

$$C_0 = C_m [\log m + p\log \frac{p}{m-1} + (1-p)\log(1-p)].$$





p очень мало, например, 0.01. С точки зрения технической реализации правая ветвь не используется. С увеличением p пропускная способность падает. Если вероятность ошибки совпадает с вероятностью правильного приема p = $p_{прав}$ = 0.5, пропускная способность C_0 = 0. Говорят, что в этом случае наступает обрыв канала, по каналу передавать информацию бессмысленно.

Вторая теорема Шеннона относится к каналу с помехами.

2 Теорема Шеннона. Если производительность ИС меньше пропускной способности канала $\Pi(X) < C_0$, то сообщение от этого источника можно преобразовать так, чтобы передавать их по каналу с помехами с любой степенью точности, т. е. за счет существования избыточности в сообщениях, вводимой специальным образом, можно уменьшить вероятность ошибки до сколь угодно малой величины.

С точки зрения технической реализации эта теорема означает, что существует способ кодирования и декодирования, при котором вероятность ошибочного декодирования может быть сколь угодно малой. Если $\Pi(X) > C_0$, то таких способов не существует.

Вторая теорема Шеннона является идеологической основой для существования помехоустойчивого (корректирующего) кодирования в каналах связи.

Модели дискретных каналов

Дискретным каналом называют совокупность средств, предназначенных для передачи дискретных сигналов. Такие каналы широко используются при передачи данных, в телеграфии, радиолокации.

Дискретные сообщения, состоящие из последовательности знаков алфавита источника сообщений первичного алфавита Z1 Z2 ... ZL преобразуются в кодирующем устройстве в последовательности символов.

Объём *m* алфавита символов (вторичным алфавитом) U1,U 2 ... Um, как правило меньше объема *L* алфавита знаков, но они могут и совпадать.

Материальным воплощением символа является элементарный сигнал, получаемый в процессе манипуляции - дискретного изменения определенного параметра переносчика информации. Элементарные сигналы формируются с учетом физических ограничений, накладываемых конкретной линией связи. В результате манипуляции последовательности символов ставится в соответствие сложный сигнал. Множество сложных сигналов конечно. Они различаются числом, составом и взаимным расположением элементарных сигналов.

Термины "элементарный сигнал" и "символ", так же как "сложный сигнал" и "последовательность символов", в дальнейшем будут использоваться как синонимы.

Информационная модель канала с помехами задается множеством символов на его входе и выходе и описанием вероятностных свойств передачи отдельных символов. В общем случае канал может иметь множество состояний и переходить из одного состояния в другое как с течением времени, так и в зависимости от последовательности передаваемых символов.

В состоянии канал характеризуется матрицей условных вероятностей $p(v_j / u_i)$ того, что переданный символ *Ui* будет воспринят на выходе как символ *Vj*. Значение вероятностей в реальных каналах зависит от многих различных факторов: свойств сигналов, являющихся физическими носителями символов (энергия, вид модуляции и т.д.), характера и интенсивности воздействующих на канал помех, способа определения сигнала на приемной стороне.

При наличии зависимости переходных вероятностей канала от времени, что характерно практически для всех реальных каналов, он называется нестационарным каналом связи. Если эта зависимость несущественна, используется модель в виде стационарного канала, переходные вероятности которого не зависят от времени. Нестационарный канал может быть представлен рядом стационарных каналов, соответствующих различным интервалам времени.

Канал называется с "памятью" если переходные вероятности в данном состоянии канала зависят от его предыдущих состояний. Если переходные вероятности постоянны, т.е. канал имеет только одно состояние, он называется стационарным каналом без памяти.

Стационарный дискретный двоичный канал без памяти однозначно определяется четырьмя условными вероятностями: p*(0/0) ,p(1/0),p(0/1), p(1/1).*

Если вероятности искажения можно принять равными, т.е. *p(0/1) ~ p(1/0)=q,* то такой канал называется двоичным симметричным каналом при *p(0/1) -p(1/0)* канал называется несимметричным). Символы на его выходе правильно принимают с вероятностью *1-p=q*. Математическая модель упрощается.

Именно этот канал исследовался наиболее интенсивно не столько в силу своей практической значимости (многие реальные каналы описываются им весьма приближенно), сколько в силу простоты математического описания.

Важнейшие результаты, полученные для двоичного симметрического канала, распространены на более широкие классы каналов.

Следует отметить еще одну модель канала, которая в последнее время приобретает все большее значение. Это дискретный канал со стиранием. Для него характерно, что алфавит выходных символов отличается от алфавита входных символов. На входе, как и ранее, символы 0 и 1, а на выходе канала фиксируются состояния, при которых сигнал с равным основанием может быть отнесен как к единице так и к нулю. На месте такого символа не ставится ни нуль, ни единица: состояние отмечается дополнительным символом стирания S. При декодировании значительно легче исправить такие символы, чем ошибочно определенные.

Информационные характеристики непрерывного источника сообщений и канала связи

К информационным характеристикам источников непрерывных сообщений (ИНС) относятся дифференциальная энтропия (ДЭ) и эпсилон-энтропия (ЭЭ).

Обобщим понятие энтропии и КИ на ансамбль непрерывных сообщений.

ИНС может вырабатывать любую реализацию сообщения x(t) из неограниченного множества. В этом случае на множестве реализаций задана плотность распределения вероятности или интегральный закон:

$x(t) \in X \to w(x)\ [F(x)]$.

Заменим непрерывное состояние ИС x(t) дискретными $x_1, x_2, \ldots, x_m$ отстоящими друг от друга на интервал Δx. В этом случае можно определить вероятность нахождения (появления) отсчета $x_1, x_2, \ldots, x_m$ как $p_i = w(x_i)\Delta x$.

Поэтому на основании определения среднего КИ можно определить эту величину в данном случае следующим образом:

$$M\{-\log p_i\} = H(X) = -\sum_{i=1}^{m} w(x_i)\Delta x \log(w(x_i)\Delta x).$$

Устремляя к нулю и учитывая условия нормировки для плотности распределения вероятности

$$\sum_{i=1}^{m} w(x_i)\Delta x = 1,$$

мы получим, что

$$H(X) = -\int_{-\infty}^{+\infty} w(x)\log w(x)dx - \lim_{\Delta x \to 0} \log \Delta x.$$

$-\lim_{\Delta x \to 0} \log \Delta x \to +\infty$ и не зависит от w(x), следовательно, среднее КИ и Э непрерывного сообщения являются бесконечно большими величинами. Но на практике интересуются величиной приращения Э. Поэтому величину $\log \Delta x$, которая зависит от Δx, не учитывают и вводят понятие ДЭ

$$h(x) \equiv -\int_{-\infty}^{+\infty} w(x)\log w(x)dx. \qquad (1)$$

Замечание. ДЭ в отличие от обычной Э нельзя рассматривать как меру собственной информации. Она не обладает многими свойствами обычной Э, в частности она может принимать и отрицательные значения. Информационный смысл имеет разность двух ДЭ, чем и объясняется название. Что касается свойства аддитивности, то оно сохраняется, т. е. ДЭ нескольких сечений случайного процесса равна сумме их ДЭ, вычисленной с учетом вероятностной зависимости между сечениями. Если решить вариационную задачу поиска распределения вероятности, доставляющую максимум функции ДЭ случайного процесса при ограничениях вида:

$$\int_{-\infty}^{+\infty} w(x)dx = 1,$$

$$\int_{-\infty}^{+\infty} x^2 w(x)dx = {\sigma_x}^2,$$

окажется, что решением является

$w(x) = \frac{1}{\sigma_x\sqrt{2\pi}}\exp(-\frac{x^2}{{\sigma_x}^2})$ - Нормальный закон или закон Гаусса.

Максимальное значение ДЭ: $\max k(x) = \log\sqrt{2\pi e\sigma_x^2}$.

Только дисперсия влияет на величину $\max k(x)$.

Понятие УЭ для ИНС для фиксированного состояния вводится по аналогии с соотношением (1).

$$h(X/y) = -\int_{-\infty}^{+\infty} w(x/y)\log w(x/y)dx, \qquad (2)$$

y - параметр.

КИ, содержащееся во всех возможных сообщениях источника для всех возможных состояний источника X имеет вид:

$$h(X/Y) = M\{h(X/y)\} = \int_{-\infty}^{+\infty} w(y)h(X/y)dy = -\int_{-\infty-\infty}^{+\infty+\infty} w(y)w(x/y)\log w(x/y)dxdy. \quad (3)$$

Энтропия объединения ИНС определяется согласно приведенным ранее формула следующим образом:

$h(X*Y) = h(X) + h(Y/X)$,

$h(X)$ - ДЭ, $h(Y/X)$ - УЭ, определенная по (3).

Взаимное КИ I(X,Y) между двумя ИНС X и Y вводят как разность безусловной и условной ДЭ:

$$I(X,Y) \equiv h(X) - h(X/Y). \qquad (4)$$

Две реализации сообщений: принятое *y(t)* и переданное *x(t)* называются эквивалентными, если различие между ними несущественно в смысле выбранного критерия (обычно это критерий СКО). Вводится понятие отклонения $\varepsilon(t) \equiv x(t) - y(t)$, задается ε_0^2. Если $\overline{\varepsilon^2}(t) < \varepsilon_0^2$, то реализации считаются эквивалентными.

ЭЭ $H_\varepsilon(x)$ называется минимальное среднее КИ, содержащееся в одном отсчете сообщения y(t) относительно сообщения x(t), при котором они еще эквивалентны. В соответствии с соотношением (4)

$H_\varepsilon(x) \equiv \min I(x,y)$,

где I(x,y) – взаимное КИ между реализациями x и y.

$H_\varepsilon(x) = h(x) - \max\limits_{\{w(x/y)\}} h(x/y)$.

ЭЭ определяет количество существенной информации, содержащейся в одном отсчете непрерывного сообщения.

Рассмотрим ИНС x(t), представляющий собой стационарный гауссовский процесс с ограниченной дисперсией σ_x^2. Процесс x(t) может быть представлен как

$x(t) = y(t) - \varepsilon(t)$,

где $\varepsilon(t)$ – шум воспроизведения. Поэтому УДЭ $h(x/y)$, присутствующая в определении ЭЭ, при заданном сообщении x(t) полностью определяется шумом воспроизведения $\varepsilon(t)$. Поэтому $\max h(x/y) = \max h(\varepsilon)$. Ранее было показано, что максимум $\max h(\varepsilon)$ достигается на гауссовском случайном процессе $\varepsilon(t)$:

$\max h(x/y) = \max h(\varepsilon) = \log_2 \sqrt{2\pi e \sigma_\varepsilon^2}$.

В этом случае выражение для ЭЭ достигает максимума, если x(t) и $\varepsilon(t)$ – гауссовские процессы. ЭЭ гауссовского источника с ограниченной дисперсией:

$$H_\varepsilon(x) = \log_2 \sqrt{2\pi e \sigma_x^2} - \log_2 \sqrt{2\pi e \sigma_\varepsilon^2} = 0.5 \log_2 \frac{\sigma_x^2}{\sigma_\varepsilon^2}. \qquad (5)$$

Соотношение характеризует КИ, приходящееся на один отсчет. Отношение двух дисперсий характеризует отношение сигнал/шум, при котором сообщения y(t) и x(t) можно считать эквивалентными. Это величина зависит от характера передаваемых сообщений, и при независимых отсчетах сообщений содержащаяся в них информация суммируется.

Производительность ИНС определяется как КИ, выдаваемое этим источником в единицу времени (в системе СИ 1 с) при заданном критерии эквивалентности. Это означает, что здесь мы будем оперировать понятием ЭЭ.

Если средняя скорость выдачи независимых отсчетов сообщений (техническая скорость) равняется C_m, то ε – производительность ИС для гауссовского процесса:

$\Pi_\varepsilon(x) = C_m H_\varepsilon(x) = C_m (h(x) - \log_2 \sqrt{2\pi e \sigma_\varepsilon^2})$.

Отсчеты берутся в соответствии с теоремой Котельникова:

При равном спектре сообщения в полосе $0 \ldots F_в$ отсчеты являются некоррелированными, а для гауссовского ИС – независимыми. Для гауссовского процесса:

$$\Pi_\varepsilon(x) = 2 F_в H_\varepsilon(x), \qquad \text{с учетом (5)}$$

$$\Pi_\varepsilon(x) = F_в \log\left(\frac{\sigma_x^2}{\sigma_\varepsilon^2}\right) = F_в \log \rho.$$

КИ, выдаваемое гауссовским ИС за время T, определяется:

$$I_0(x) = T^* \Pi_\varepsilon(x) = T^* F_в \log \rho. \qquad (6)$$

Величина $I_0(x)$ совпадает по определению с объемом сигнала, если считать, что динамический диапазон сигнала равен $\log \rho$. Выражение определяет максимальное КИ, выдаваемое ИС за время T, т. к. производительность гауссовского источника больше производительности любого другого источника с той же мощностью.

Пропускная способность непрерывных каналов передачи информации. Формула Шеннона

Пропускная способность (ПС) непрерывного канала – максимальное КИ, проходящей по каналу за единицу времени, если x(t) – переданное, y(t) – принятое сообщение определяются своими отсчетами, взятыми по теореме Котельникова через интервал $\Delta t = \frac{1}{2\Delta F_{экв}}$. Обычно $\Delta F_{экв} = \Delta F_в$.

Для канала передачи сообщений с КИ I(x,y) определяется величиной – взаимной информацией и представляет собой КИ, проходящей по каналу за время Т и определяется суммой КИ, переданной за каждый отсчет.

ПС канала, взятая за один отсчет, по определению равна максимуму взаимного КИ:

$$C_0^{(1)} = \max_{\{w(x/y)\}} I(x,y) = \max[h(x) - h(x/y)] = \max[h(y) - h(y/x)].$$

ПС всего канала в целом равна сумме $C_0^{(1)}$ всех отсчетов: $C_0 = \sum C_0^{(1)}$.

Вычислим ПС гауссовского канала без памяти, в котором действует аддитивная помеха типа белый гауссовский шум в полосе $\Delta F_{экв}$. Предположим, что сигнал, передаваемый по каналу, является гауссовским процессом с ограниченной мощностью σ_x^2. Мощность шума, действующего в канале $\sigma_ш^2 = P_ш$, где $P_ш$ - мощность шума на сопротивлении 1 Ом. Учитывая аддитивность действия шума и максимум ошибки воспроизведения для гауссовского процесса: $\max h(\varepsilon) = \log_2 \sqrt{2\pi e \sigma_\varepsilon^2}$,

можно вычислить ПС на отсчет

$$C_0^{(1)} = \max[h(y) - h(y/x)], \qquad (7)$$

y(t) = x(t)+n(t).

На гауссовском распределении $h(y/x) = \log_2 \sqrt{2\pi e \sigma_ш^2}$.

Поскольку x(t) и n(t) – гауссовские процессы, следовательно при независимом сигнале и шуме $h(y) = \log_2 \sqrt{2\pi e \sigma_y^2}$,

т. к. y(t) – гауссовский процесс, где $\sigma_y^2 = \sigma_x^2 + \sigma_ш^2 = P_c + P_ш$. Формулу (7) можно переписать в следующем виде:

$$C_0^{(1)} = \log_2 \sqrt{2\pi e(P_c + P_ш)} - \log_2 \sqrt{2\pi e P_ш} = \log_2 \sqrt{\frac{P_c + P_ш}{P_ш}} = 0.5 \log_2 (1 + \frac{P_c}{P_ш}). \qquad (8)$$

Информация, переданная за несколько отсчетов, максимальна, если отсчеты сигнала независимы. Это достигается при равномерном спектре сигнала x(t) в полосе $\Delta F_{экв}$. Поэтому, сложив величины, определяемые соотношением (8), для $2\Delta F_{экв}$ независимых отсчетов по Котельникову, получим ПС канала в целом:

$$C_0 = 2\Delta F_{экв}\ C_0^{(1)} = \Delta F_{экв}\ \log_2 (1 + \frac{P_c}{P_ш}). \qquad (9)$$

(9) – формула Шеннона.

Подытожим основные предположения, которые делались при выводе формулы Шеннона:

Сигнал x(t) имеет такой же характер, как и действующий в канале шум n(t), и описывается гауссовским законом распределения вероятности.

Спектр сигнала и спектр шума огранивчены сверху верхней частотой пропускания канала и действуют в полосе $F_{экв}$.

Сигнал x(t) и шум n(t) имеют равномерный энергетический спектр, т. е. оба являются шумоподобными.

Полоса пропускания канала согласована с шириной спектра сигнала.

Средняя мощность сигнала и средняя мощность шума ограничены величинами: $P_c = \sigma_x^2$, $P_ш = \sigma_ш^2$, и они взаимодействуют аддитивно.

Формула Шеннона определяет ПС канала для шумоподобных сигналов. Такие каналы называются идеальными гауссовскими каналами с ограниченной мощностью. Предельные возможности согласования ИНС с непрерывным каналом устанавливает третья теорема Шеннона.

Теорема Шеннона. Если при заданном критерии эквивалентности сообщений ε_0^2 $\varepsilon-$ производительность этого источника меньше ПС канала, т. е. $\Pi_\varepsilon(x) < C_0$, то существует такой способ кодирования и декодирования в обобщенном смысле (т. е. преобразование сообщения в сигнал и обратно), при котором неточность воспроизведения сообщения сколь угодно близка к ε_0^2. При $\Pi_\varepsilon(x) > C_0$ такого способа не существует.

Комментарии к формуле Шеннона:

Из формулы видно, что при неограниченной мощности сигнала ПС канала неограниченно возрастает, и ПС равна нулю, если отношение сигнал/шум равно нулю.

Согласно подходу Шеннона единственной причиной ошибок в канале являются шумы, действующие в канале, а сам канал считается неискажающим.

Формула Шеннона указывает на возможность обмена полосы пропускания канала на мощность сигнала, и наоборот.

Линейная зависимость C_0 от $\Delta F_{экв}$ (на основном участке) и логарифмическая зависимость C_0 от отношения сигнал/шум указывает на то, что эффективным является обмен мощности сигнала на полосу пропускания канала.

Замечание. Для реального гауссовского канала с ограниченной пиковой мощностью сигнала ПС оказывается несколько иной, чем по формуле Шеннона. В этом случае ПС канала может быть расчитана по формуле:

$$C_0 = \Delta F_{экв} \log_2(1 + \alpha_c \frac{P_c}{P_ш}).$$

α_c – коэффициент, учитывающий ухудшение информационных свойств применяемого класса сигналов по сравнению с гауссовским шумоподобным сигналом: $0 \le \alpha_c \le 1$. Как показывают расчеты, $\alpha_c \approx 0.3$ для гармонического сигнала . Для импульсных сигналов $\alpha_c \approx 0.03$. Для гауссовского шумоподобного сигнала $\alpha_c = 1$, и применяется классическая формула Шеннона.

Анализ формулы Шеннона:

$\Delta F_{экв} = \text{const}$. При $h^2 = \frac{P_c}{P_ш} << 1$ $\log_2(1+h^2) \approx h^2$. При $h^2 >> 1$ $\log_2(1+h^2) \approx \log_2 h^2$.

$\frac{C_0}{\Delta F_{экв}}$

$P_c = \text{const}$, $N_0 = \text{const}$. Тогда формула Шеннона может быть переписана в следующем виде:

$$C_0 = \Delta F_{экв} \log_2(1 + \frac{P_c}{N_0 \Delta F_{экв}}) = \log_2(1 + \frac{P_c}{N_0 \Delta F_{экв}})^{\Delta F_{экв}} = \log_2\left[(1 + \frac{P_c}{N_0 \Delta F_{экв}})^{\frac{N_0 \Delta F_{экв}}{P_c}}\right]^{\frac{P_c}{N_0}}.$$

Рассмотрим предел

$$\lim_{\Delta F_{экв} \to \infty} C_0 = \log_2 e^{\frac{P_c}{N_0}} = \frac{P_c}{N_0} \log_2 e \approx 1.443 \frac{P_c}{N_0}.$$

Из последнего соотношения следует, что для передачи одного бита в секунду необходимо обеспечить мощность сигнала $P_c \ge \frac{N_0}{\log_2 e} = \frac{N_0}{1.443} = 0.69 N_0$.

При малых значениях полосы пропускания $\Delta F_{экв}$ ПС C_0 пропорциональна $\Delta F_{экв}$. При дальнейшем увеличении рост C_0 замедляется.

Согласование статистических свойств источника сообщений и канала связи.

Согласование статистических свойств и отражающих их информационных характеристик источника сообщений и канала связи проводится с целью улучшения качества осуществляется по

трем основным показателям: достоверности, средней скорости передачи и сложности технической реализации системы, определяющей ее стоимость и надежность. Хотя с точки зрения практики сложность технической реализации может иметь решающее значение, при определении предельных возможностей системы целесообразно ограничиться только первыми двумя показателями.

Достоверность дискретного канала обычно оценивается значением вероятности ошибочного приема одного символа (элементарного сигнала). В случае передачи непрерывных сообщений о достоверности судят по значению среднеквадратической ошибки при воспроизведении сообщения

$$M[E]=M[(W(t)-Z(t))]$$

где W(t) - сообщение, поступающее с выхода канала;

Z(t) - сообщение на выходе канала;

Достоверность характеризует помехоустойчивость информационной системы.

Под скоростью передачи подразумевают среднее количество информации, передаваемое по каналу в единицу времени. Именно эта (а не техническая) скорость формирования символов подлежит согласованию с пропускной способностью канала.

Скорость передачи информации характеризует эффективность системы**.**

Если высоких требований в отношении скорости передачи и достоверности к системе передачи не предъявляется, то согласование статистических (информационных) характеристик источника сообщений и канала связи не является принципиально необходимым.

При преобразовании сообщений в сигналы в этом случае могут преследоваться две основные цели. Одна из них заключается в том, чтобы преобразовать сообщения в такую систему символов (код), чтобы она обеспечивала простоту и надежность аппаратурной реализации информационных устройств и приемлемую их эффективность:

простоту аппаратуры различения элементарных сигналов, соответствующих отдельным символам, приемлемое время при их передаче, простоту выполнения в этой системе арифметических и логических действий. Техническая реализация процесса кодирования в таком простейшем виде при непрерывном входном сигнале осуществляется аналого-цифровыми преобразователями.

Другой целью преобразования сообщения является защита их от санкционированного доступа. Такое преобразование называют шифрованием. Оно может проводится как на уровне знаков, так и на уровне символов.

В случае отсутствия необходимости в статистическом согласовании источника сообщений с каналом связи вопросы повышения качества функционирования системы решаются для дискретного канала о входа модулятора до выхода демодулятора.

Считается, что символы на вход модулятора поступают равновероятно и статистические связи между ними отсутствуют.

Из множества сигналов, удовлетворяющих заданным ограничениям по мощности и полосе частот, для отображения символов, отбираются такие, которые в предположении воздействия на них аддитивного гауссова шума обеспечивают наибольшую достоверность приема отдельного символа. Одновременно определяется и структура оптимального приемника. Наиболее полно эти вопросы рассмотрены для случая двоичного канала ($m=2$)

Увеличение эффективности и помехоустойчивости системы передачи информации, как показал Шеннон, возможно за счет введения в канал связи кодирующего, а следовательно и декодирующего устройств, цел которых в статистическом согласовании свойств источника сообщений и канала связи.

Доказанными им теоремами обоснованно существование оптимального способа кодирования, при котором достигается скорость передачи информации, сколь угодно близкая к пропускной способности данного канала связи. Под способом кодирования при этом подразумевается совокупность операций по преобразованию сообщений в сигналы и обратного преобразования смеси сигнала с помехами в сообщения, включая операции в части канала <модулятор - демодулятор>.

К сожалению, указанные теоремы не дают конструктивных рекомендаций относительно пу-

тей реализации оптимального способа кодирования. Определить соответствующую совокупность операций, а следовательно, и структуру оптимальной системы связи пока не удалось даже при ряде допущений, существенно упрощающих модели каналов. Для упрощения задачи переходят к оптимизации системы по частям путем нахождения наилучшего кода при условии оптимально спроектированной части канала <модулятор - демодулятор>.

Выяснить также целесообразность разделения процедур кодирования, обусловленных статистическими свойствами источника сообщений, и процедур кодирования, зависящих от статистических свойств канала связи. Такое разделение способствует лучшему пониманию существа процессов преобразования. С практической точки зрения оно ценно тем, что позволяет реализовать как кодирующие, так и декодирующие устройства из двух из двух фактически независимых блоков: кодера источника (КИ) и декодера источника (ДКИ), кодера канала (КК) и декодера канала (ДКК).

Рассмотрим теперь особенности статистического согласования различных источников сообщений и каналов связи.

Предположим, что дискретные сообщения, поступающие с источника, обладают избыточностью, а вредным действием помех в канале можно пренебречь, что будет близко к реальности при отношении сигнал/помеха, значительно превышающим единицу. В этом случае учитывать проблему обеспечения помехоустойчивости нет необходимости и остается задача повышения эффективности.

В основной теореме Шеннона о кодировании для дискретного канала без помех утверждается, что посредством преобразования сообщений в статистической независимые и равновероятные символы можно повысить скорость передачи вплоть до пропускной способности этого канала (подробно об этом будем рассматривать позже).

Техническая реализация указанной возможности осуществляется кодером источника, обеспечивающим такое кодирование, при котором за счет устранения избыточности снижается среднее число символов, требующихся для выражения знака сообщения. При отсутствии помех это непосредственно дает выигрыш во времени передачи (или в объеме запоминающего устройства), что повышает эффективность системы. Поэтому такое кодирование получило название эффективного или оптимального.

При наличии помех в канале оно позволяет преобразовать входную информацию в последовательность символов, наилучшим образом (в смысле максимального сжатия) подготовленную для дальнейших преобразований.

Оптимальное кодирование

Теория кодирования предполагает, что в канале нет помех. Идеологическую основу этим кодам дал Шеннон в 1 Теореме. Согласно 1 Теореме Шеннона, ИС можно преобразовать в источник с избыточностью, равной нулю.

Основные принципы эффективного кодирования:

Необходимо обеспечить минимальную среднюю длину кодового слова. Для этого избыточность ИС должна быть сведена к минимуму, теоретически к нулю. Поэтому эффективный код (ЭК) должен состоять из кодовых слов, в которых все символы равновероятны и независимы. Результатом является равенство: $C_u = C_0$.

Ни одна из кодовых комбинаций не должна получаться из другой, более короткой путем добавления новых символов. ЭК не требуют разделяющих сигналов (маркеров) между словами, и при этом должно выполняться их однозначное декодирование. Коды, удовлетворяющие этому свойству, называются префиксными, т. к. ни одно кодовое слово не является передней частью (префиксом) другого слова.

ЭК являются неравномерными, т. е. для передачи разных символов сообщений используются кодовые комбинации разной длины. При этом наиболее вероятные сообщения кодируются самыми короткими кодовыми словами, вследствие чего средняя длина кодового слова в сообщении уменьшается. Это позволяет решить задачу равенства информационной скорости передачи и ПС канала.

Средняя длина кодового слова определяется следующим выражением:

$$\bar{n} = \sum_{i=1}^{k} n_i p(x_i)$$

где n_i – длина кодового слова x_i, $\sum_{i=1}^{k} p(x_i) = 1$. При этом вводится понятие избыточности кода источника:

$$R_k \equiv \frac{\bar{n} - n_{\min}}{\bar{n}} = 1 - \frac{n_{\min}}{\bar{n}}$$.

R_k оценивается на выходе декодера.

При построении неравномерного кода более вероятные сообщения кодируются короткими блоками, менее вероятные – длинными, в результате чего средняя длина блока $\bar{n}$ уменьшается. Однако при выполнении процедуры кодирования необходимо обеспечить однозначность декодирования, т. е. соблюсти свойство префиксности кода.

Поясним эту проблему на примере. Пусть алфавит ИС содержит 6 букв (сообщений), передаваемых независимо друг от друга, $\sum_{i=1}^{6} p(x_i) = 1$.

А	p(А) = 0.4
Б	p(Б) = 0.3
В	p(В) = 0.1
Г	p(Г) = 0.08
Д	p(Д) = 0.07
Е	p(Е) = 0.05

$$H(X) = -\sum_{i=1}^{6} p(x_i) \log_2 p(x_i) = 2.16$$

Чтобы закодировать сообщения равномерным двоичным кодом необходимо затратить на каждое сообщение 3 символа. В соответствии с 1 Теоремой Шеннона эти сообщения можно закодировать двоичными символами так, чтобы в среднем на каждое сообщение затрачивалось 2.16 бит на символ. Попробуем сделать это, не задумываясь пока над однозначностью декодирования. Присвоим наиболее вероятным символам самые короткие блоки в соответствии с 3 принципом эффективного кодирования, делая код неравномерным, код разной длины.

	Код	
А	0	
Б	1	
В	00	(*)
Г	01	
Д	10	
Е	11	

Таким образом, для передачи сообщений А и Б, имеющих суммарную вероятность 0.7, используется один символ. Для передачи остальных сообщений, имеющих суммарную вероятность 0.3, используется два символа. В результате средняя длина кодового слова равна:

$\bar{n}$ = 0.7*1 + 0.3*2 = 1.3 < 2.16.

Получилось, что сообщение закодировано еще более экономично, чем позволяет теореме Шеннона. Парадокс объясняется тем, что выбранный код не пригоден для передачи сообщений, т. к. он не обеспечивает однозначного декодирования.

В принципе, используя код (*), можно обеспечить однозначность декодирования, если после каждого сообщения передавать некоторый разделительный символ (запятую, маркер), разделяющий эти сообщения. Тогда это будет не двоичный код, а троичный. Таким образом поступил,

например, Морзе, в коде которого кроме точки и тире используется пробел. Тем не менее, можно построить код и однозначно декодировать принимаемые сообщения без использования запятой. Для этого достаточно (хотя и не необходимо) строить код таким образом, чтобы он удовлетворял префиксному свойству (см. п. 2). Ни одно из используемых кодовых слов не должно являться началом другого кодового слова. Это свойство не выполняется у кода (*), т. к., например, слово, соответствующее сообщению А, является началом сообщения В. Существует несколько алгоритмов построения неравномерных кодов с префиксным свойством. Среди них оптимальным (позволяющим лучше всего приблизиться к границе, определяемой энтропией) является алгоритм Хаффмана. Более простым и несколько худшим является алгоритм Шеннона- Фано.

Алгоритм кодирования Шеннона - Фано

Сообщения алфавита источника записаны в порядке невозрастающих вероятностей и разделяются на две части так, чтобы суммарные вероятности сообщений в каждой из них были по возможности почти одинаковыми. Сообщениям первой части приписывается в качестве первого кодового слова 0, сообщениям второй части – 1.

Далее каждая из этих частей (если она содержит более одного сообщения) делится на две по возможности равновероятные части, и в качестве второго символа для первой из них берется 0, а для второй части – 1.

Этот процесс продолжается до тех пор, пока в каждой из полученных частей не останется по одному сообщению.

Для выше приведенного примера алгоритм заключается в следующем:

На первом этапе разделения на две части в первой части окажется одно сообщение А с вероятностью 0.4, а во второй части – все остальные сообщения с суммарной вероятностью 0.6. Если включить в первую часть два сообщения А и Б, то отклонения от равновероятности в группах будут еще больше. Припишем сообщению А 0 в качестве первого кодового символа, а всем остальным – 1.

На втором этапе разделим сообщения Б, В, Г, Д, Е на две равновероятные части, включив в первую часть сообщение Б, а во вторую часть – остальные. В этих группах суммарные вероятности для обеих частей одинаковы – по 0.3. Припишем сообщению Б в качестве второго кодового символа 0, а остальным – 1.

На третьем этапе сообщения В, Г образуют одну часть, сообщения Д, Е – вторую. В результате приходим к алгоритму

Сообщения	1 этап	2 этап	3 этап	4 этап
А	0			
Б	1	0		
В	1	1	0	0
Г	1	1	0	1
Д	1	1	1	0
Е	1	1	1	1

Этот код, как можно убедиться, по своему построению обладает префиксным свойством.

Среднее число символов, приходящееся на одно сообщение, с учетом их вероятностей равняется 2.2, т. е. превышает значение Э ИС меньше, чем на 2%. К значению 2.16 можно было подойти еще ближе, если бы при составлении кода сопоставлялись с кодовыми словами не одиночные сообщения, а блоки, состоящие из нескольких сообщений (кодированные блоки).

Замечание 1. Следует отметить, что ЭК, например, Шеннона- Фано, позволяют сократить только ту избыточность ИС, которая связана с неравными вероятностями появления сообщения. Избыточность другого происхождения, обусловленная памятью и статистической зависимостью между элементами сообщений, сокращается другими методами, в основном укрупнением блоков.

Замечание 2. Рассмотренная методика Шеннона- Фано не всегда приводит к однозначному построению кода, т. к. при разбиении на подгруппы можно сделать большей по вероятности как верхнюю, так и нижнюю подгруппу. От этого недостатка свободна методика кодирования по Хаффману.

Алгоритм кодирования Хаффмана

Эта методика гарантирует однозначное построение кода с наименьшим для данного распределения вероятностей средним числом символов, приходящихся на букву. Метод Хаффмана всегда приводит к получению оптимального множества кодовых слов в том смысле, что никакое другое множество не имеет меньшей средней длины слова на элементарное сообщение источника.

Для двоичного кода m = 2 методика сводится к следующему:

буквы алфавита ИС выписывают в основной левый столбец в порядке убывания вероятности.

две последние нижние буквы объединяют в одну вспомогательную букву, которой приписывают суммарную вероятность.

вероятности букв, не участвовавших в объединении, и полученная суммарная вероятность (вспомогательная буква, полученная выше) снова располагаются в порядке убывания вероятности во втором следующем столбце таблицы, а две последние буквы этого столбца объединяются в одну вспомогательную путем сложения вероятностей. Объем алфавита уменьшается на единицу.

производят новое укрупнение алфавита путем объединения двух новых нижних символов с наименьшей вероятностью. Получают новый алфавит меньшего объема.

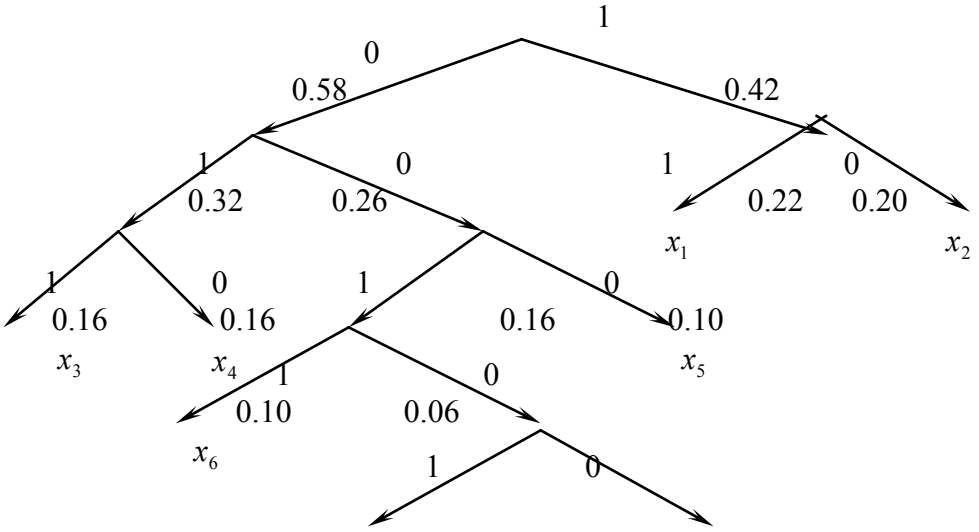
процесс продолжается до тех пор, пока не получается единственная вспомогательная буква с вероятностью 1.

Поясним эту методику на примере алфавита с восьмью состояниями.

x_i	p(x_i)	Вспомогательные столбцы							Кодовые слова
		1	2	3	4	5	6	7	
x_1	0.22	0.22	0.22	0.26	0.32	0.42	0.58	1	01
x_2	0.20	0.20	0.20	0.22	0.26	0.32	0.42		00
x_3	0.16	0.16	0.16	0.20	0.22	0.26			111
x_4	0.16	0.16	0.16	0.16	0.20				110
x_5	0.10	0.10	0.16	0.16					100
x_6	0.10	0.10	0.10						1011
x_7	0.04	0.06							10101
x_8	0.02								10100

Для заполнения последнего столбца таблицы удобно воспользоваться кодовым деревом. Для этого проводят линии, соединяющие символы при последовательном укрупнении алфавита. Концы ветвей кодового дерева являются символами исходного ИС алфавита $x_1,\ldots,x_8$. Приписывая далее ветвям дерева, исходящим из каждого промежуточного узла, различные символы алфавита кода (0 или 1), получают кодовые слова для каждого символа. Из точки, соответствующей вероятности 1 (вершина кодового дерева), вниз направляются две ветви, причем ветви с большей вероятностью присваивается кодовый символ 1, а с меньшей – 0. Такое последовательное ветвление продолжается до тех пор, пока мы не дойдем до вероятности каждой буквы.

Для рассмотренного выше примера кодовое дерево имеет следующий вид.



0.04 0.02

x_7 x_8

Двигаясь по кодовому дереву сверху вниз к каждой букве исходного алфавита, можно теперь записать для нее кодовую комбинацию по Хаффману. Нетрудно убедиться в том, что полученные кодовые комбинации по Хаффману неравномерные и префиксные.

Замечание. Другая формулировка Теоремы Шеннона, характеризующая предельные возможности ЭК.

Теорема Шеннона. Сообщения от источника с энтропией H(X) всегда можно закодировать последовательностями символов с объемом алфавита m (алфавита кодера) так, что среднее число символов, приходящееся на знак сообщения $\overline{n}$ будет сколь угодно близкой к величине $\frac{H(X)}{\log_2 m}$, т. е. не менее ее:

$$\overline{n}_{симв} \geq \frac{H(X)}{\log_2 m}.$$

Для кодера с двумя состояниями (m = 2) это утверждение сводится к тому, что $\overline{n}_{симв} \geq$ H(X).

Теорема не указывает конкретного способа кодирования, но из нее следует, что при выборе каждого символа кодовой комбинации необходимо стремиться, чтобы он нес максимальную информацию.

Эффективное кодирование блоками

Из теоремы Шеннона следует, что избыточность в последовательности кодовых символов можно устранить, если перейти к кодированию достаточно большими блоками.

Рассмотрим процедуру эффективного кодирования сообщений, образованными с помощью алфавита, состоящего всего лишь из двух букв с вероятностями появления каждой из них: $p(x_1) = 0.9$, $p(x_2) = 0.1$. Так как вероятности не равны, то последова-тельность из таких букв будет обладать избыточностью. Однако побуквенным кодированием мы никакого эффекта не получим, т. к. на передачу каждой буквы требуется как минимум один символ: 0 или 1. В то же время, Э такого ИС:

$$H(X) = \sum_{i=1}^{6} p(x_i) \log_2 \frac{1}{p(x_i)} = 0.9 \log_2 \frac{10}{9} + 0.1 \log_2 10 = 0.47$$

При кодировании блоками боки формируются следующим образом, процедура кодирования выглядит следующим образом (блок по 2):

Блок	Вероятность	Кодовая комбинация
$x_1\ x_1$	0.81	1
$x_1\ x_2$	0.09	01
$x_2\ x_1$	0.09	001
$x_2\ x_2$	0.01	000

Кодовая комбинация образована по алгоритму Фано-Шеннона.

Поскольку x_1 и x_2 независимы, то вероятности блоков равны произведению вероятностей знаков. В данном случае среднее число символов на блок и на букву:

$\overline{n}_{блок} = 0.81*1+0.09*2+0.09*3+0.01*3 = 1.29,$

$\overline{n}_{букв} = \frac{1.29}{2} = 0.645.$

Если произвести формирование блока по 3 буквы, ты мы достигнем еще большего эффекта в смысле приближения среднего числа символов, приходящихся на букву, к энтропии.

Блок	Вероятность	Кодовая комбинация
$x_1\ x_1\ x_1$	0.729	1
$x_2\ x_1\ x_1$	0.081	011
$x_1\ x_2\ x_1$	0.081	010
$x_1\ x_1\ x_2$	0.081	001
$x_2\ x_2\ x_1$	0.009	00011
$x_2\ x_1\ x_2$	0.009	00010
$x_1\ x_2\ x_2$	0.009	00001
$x_2\ x_2\ x_2$	0.001	00000

$\overline{n}_{блок} = 1.59,\ \overline{n}_{букв} = 0.53.$

Средняя длина, приходящаяся на букву, отличается от энтропии на 12%. Теоретический минимум 0.47 может быть достигнут при кодировании блоками, содержащими бесконечное число знаков.

Замечание. Увеличение эффективности кодирования при укрупнении блоков не связано с учетом все более далеких статистических связей (т. к. знаки алфавита не коррелированы), а определяется лишь тем, что набор вероятностей, получающихся при укрупнении блоков, можно делить на более близкие по суммарным вероятностям подгруппы.

Достоинства ЭК:

При эффективном кодировании, учитывающем вероятности появления букв алфавита ИС, удается построить коды с максимальной УЭ, приходящейся на символ.

Обеспечивается преобразование сообщения в сигнал с меньшей, чем у исходного сообщения, избыточностью (в пределе без избыточности).

На передачу сообщения затрачивается минимальное количество символов

Решается задача согласования ИС с каналом связи, в результате чего скорость передачи информации может приближаться к ПС канала.

Не требуется введение специальных разделительных символов (маркеров) для разделения одной кодовой комбинации от другой, т. к. ни одна комбинация ЭК не совпадает с началом другой, более длинной (свойство префиксности).

Недостатки ЭК:

ЭК являются неравномерными, т. е. кодовые комбинации имеют различное число символов. Если линия связи работает с постоянной скоростью передачи, то на выходе кодера необходимо иметь буферное запоминающее устройство (упругую задержку) для записи в него пульсирующих по длительности кодовых групп и для последующего считывания в канал символов с постоянной скоростью. Аналогичный буфер (упругая задержка) должен быть и на приемной стороне.

Наибольший эффект такие коды дают при кодировании исходного сообщения длинными блоками, т. к. при этом достигается равная вероятность и статистическая независимость блоков. Однако, блочное кодирование вызывает необходимость накапливать буквы алфавита источника прежде чем поставить им в соответствие определенную кодовую комбинацию. Это приводит к большим задержкам при передаче и приеме сообщений, что затрудняет (а иногда исключает) применение таких кодов в системах, работающих в реальном масштабе времени.

ЭК являются помехонезащищенными. Это означает, что любая одиночная ошибка при приеме переводит передаваемую комбинацию в другую, не равную ей по длительности, что влечет за собой неправильное декодирование целого ряда последующих кодовых слов. Такое специфическое влияние помех называется треком (пакетом) ошибок.

В чистом виде ЭК можно применять только в каналах без помех, и на практике такое кодирование является предварительной ступенью для последующего помехоустойчивого кодирования

Помехоустойчивое кодирование

В реальных условиях прием двоичных символов всегда происходит с ошибками из-за действия помех. В этом случае вместо символа 1 принимается 0 и наоборот. Ошибки могут возникать из-за помех, действующих в канале связи, из-за изменений характеристик канала связи (например, из-за замираний), из-за снижения уровня мощности ПРД, из-за нестабильности АЧХ и ФЧХ канала. В дискретных каналах общепринятым критерием оценки качества передачи является допустимая вероятность ошибки приема, нормированная на символ. Например, при телеграфной передаче эта величина составляет 10^{-3} на знак, при передаче данных – порядка 10^{-6}, при записи на CD-диск - 10^{-16}, в спутниковых каналах - 10^{-8}. Для обеспечения таких значений вероятностей используются различные меры, основной из которых служит повышение качества приема информации. Эти методы можно разбить на две группы:

1. Методы увеличения помехоустойчивости приема дискретной информации, связанные с выбором уровня сигнала, увеличения отношения $\frac{сигнал}{помеха + шум}$, увеличения ширины полосы пропускания канала и т. д. Все это – энергетические характеристики.
2. Методы обнаружения и исправления ошибок, основанные на искуственном введении избыточности в передаваемом сообщении. Избыточность передаваемого сообщения (сигнала) можно увеличить различными способами.

Последний метод базируется на том факте, что объем сигнала:

$V_c = P_c * \Delta F * T$.

Возможность увеличения мощности и полосы сигнала ограничена, поэтому идут на увеличение третьего сомножителя. Увеличение длительности сигнала осуществляется следующими способами:

a) многократными повторениями кодовых комбинаций;
b) одновременной передачей кодовых комбинации по нескольким каналам на разных частотах;
c) помехоустойчивым кодированием.

Здесь в (c) избыточность используется наиболее эффективно. Введение избыточности при использовании помехоустойчивых кодов обязательно связано с увеличением числа разрядов кодовой комбинации. Самыми распространенными помехоустойчивыми кодами являются так называемые разделимые коды, в которых кодовые комбинации состоят из двух частей: информационной и проверочной. Как правило, такие коды имеют кодовые комбинации, содержащие n символов, из которых первые k символов являются информационными, а за ними располагаются n - k проверочных символов. Эти коды обозначаются (n,k).

Основные характеристики корректирующих кодов (КК):

1. число разрешенных и запрещенных кодовых комбинаций;
2. избыточность;
3. минимальное кодовое расстояние (хэмминговое);
4. число обнаруживаемых и число исправляемых ошибок.

Число разрешенных и запрещенных кодовых комбинаций. Для блочных двоичных кодов, в которых информация передается словами (блоками) с числом символов равном n, общее число возможных кодовых комбинаций определяется:

$N_0 = 2^n$

Кодовые комбинации, используемые при кодировании, называются разрешенными, все другие комбинации называются запрещенными. Очевидно, что число разрешенных кодовых комбинаций при наличии k кодовых разрядов равно:

$N_k = 2^k$.

Поэтому число запрещенных кодовых комбинаций определяется:

$N_з = N_0 - N_k = 2^n - 2^k$.

Если кодовая комбинация на выходе канала связи оказывается запрещенной, то это указывает на помеху при передаче.

Избыточность корректирующего кода:

$$R_k \equiv \frac{r}{n} = \frac{n-k}{n} = 1 - \frac{k}{n}.$$

Минимальное кодовое расстояние d_0. Для того, чтобы можно было обнаруживать и исправлять ошибки, разрешенная кодовая комбинация должна как можно больше отличаться от запрещенной. Если ошибки в канале связи возникают независимо, то вероятность преобразования одной кодовой комбинации в другую будет тем меньше, чем большим числом символов они различаются.

Кодовым расстоянием (КР) называется количество единиц в сумме двух кодовых комбинаций по модулю 2.

Пример.

Сумма по модулю 2 кодовых пятиразрядных слов:

$01011 \oplus 10010 = 11001,$

Получено 3 единицы, следовательно, кодовое расстояние d = 3.

Минимальное КР d_0 - наименьшее из КР, найденных по всем разрешенным кодовым комбинациям при попарном их сравнении. В безызбыточном коде все комбинации являются разрешенными, поэтому $d_0 = 1$. Например, если n = 3, $N_0 = 2^3 = 8$, кодовые комбинации:

000, 001, 010, 011, 100, 101, 110, 111.

У такого кода любая разрешенная кодовая комбинация при искажении одного символа переходит в другую разрешенную комбинацию. Такой код называется примитивным (или первичным), он не обладает корректирующей способностью. Для того чтобы код обладал корректирующими свойствами, необходимо ввести избыточность, которая увеличила бы d от двух и выше. Таким образом, минимальное КР – важнейшая характеристика КК.

Число обнаруживаемых и число исправляемых ошибок. При применении двоичных кодов учитывают только дискретные искажения, при которых 1 переходит в 0 или наоборот. Такой переход в одном элементе кодовой комбинации называется единичной ошибкой. В общем число позиций (элементов) кодовой комбинации, на которых под действием помехи одни символы оказались замененными на другие, называют кратностью ошибки g, $0 \le g \le n$. Таким образом, фактически кратность ошибки -хэммингово расстояние между переданной и принятой кодовой комбинацией:

$$000 \xrightarrow[g=2]{} 110.$$

Если $d_0 = 2$, то ни одна из разрешенных кодовых комбинаций при одиночной ошибке (g = 1) не переходит в другую разрешенную кодовую комбинацию.

Например, для трехразрядного кода, приведенного выше, можно организовать подмножество разрешенных кодовых комбинаций по принципу четности в них единиц.

000, 011, 101, 110 - разрешенные комбинации,

001, 010, 100, 111 - запрещенные комбинации.

В этом случае $d_0 = 2$, такой код обнаруживает только одиночные ошибки и другие ошибки с нечетной кратностью.

В общем случае, при необходимости обнаруживать ошибки кратностью до g включительно хэммингово расстояние d_0 должно быть по крайней мере на единицу больше, чем g, т. е.

$$d_0 \ge g_0 + 1,$$

где g_0 - кратность обнаруживаемой ошибки.

Теорема 1 (характеризующая обнаружительную способность кода). Если код имеет хэммингово расстояние $d_0 > 1$, и используется декодирование по методу обнаружения ошибки, то все ошибки кратностью $g_0 < d_0$ обнаруживаются гарантированно, а что касается ошибок кратностью $g \ge d_0$, то одни из них обнаруживаются, другие – нет. Таким образом, условие обнаружения ошибок кратностью g_0 формулируется:

$g_0 \le d_0 - 1$.

Рассмотрим процедуру исправления ошибок на примере симметричного канала передачи информации без памяти (модель рассматривалась ранее в разделе ПС дискретного канала). В этом случае оптимальным является правило декодирования по наименьшему кодовому расстоянию. Это эквивалентно максимизации функции правдоподобия. Согласно этому критерию, запрещенную кодовую комбинацию декодируют как ту разрешенную, которая находится на наименьшем кодовом расстоянии от нее.

Декодер Виттерби. Исправляющая способность при этом правильном декодировании определяется соотношением:

$d_0 \ge 2 g_u + 1$,

где g_u - кратность исправляемых ошибок. Корректирующий код подбирается под канал, в котором он используется.

2 Теорема (о исправляющей способности кода). Если код имеет хэмминговое расстояние $d_0 > 2$, и используется декодирование с исправлением ошибок по наименьшему кодовому расстоянию (алгоритм максимального правдоподобия им. Виттерби), то все ошибки кратностью $g_u < \frac{d_0}{2}$ исправляются гарантированно. Что касается ошибок большей кратности, то после исправления кодовая комбинация может приобрести вид, который не соответствует исходной кодовой комбинации. Таким образом, сама процедура исправления выполнена. Окончательно:

$g_u \le \frac{d_0 - 1}{2}$.

Из 1 и 2 Теорем следует, что кратность обнаруживаемых ошибок $g_o < d_0$, кратность исправляемых ошибок $g_u < \frac{d_0}{2}$. Это означает, что, если код исправляет ошибки кратности g_u, то он может обнаружить количество ошибок в два раза больше, чем g_u, т. е. $g_o = 2 g_u$.

Корректирующие возможности кодов. Вопрос о максимально необходимой избыточности, при которой код обладает необходимыми корректирующими свойствами, является одним из важнейших в теории кодирования, и не получил до сих пор полного решения. Получен лишь ряд верхних и нижних границ (оценок), которые устанавливают связь между максимально возможным хэмминговым расстоянием и избыточностью кода.

Верхняя граница Плоткина

$d_0 \le \frac{n * 2^{k-1}}{2^k - 1}$

дает выражения для максимального значения d_0 .

Верхняя граница Хэмминга

$2^k \le \frac{2^n}{\sum_{i=0}^{(d_o-1)/2} C_k^i}$

устанавливает максимальное число разрешенных кодовых комбинаций при заданных n и d_0.

В частности, для кодов Хэмминга, обладающими кодовыми расстояниями 3 и 4, справедливы соотношения: $d_0 = 3 \rightarrow r \ge \log_2(n+1)$, $d_0 = 4 \rightarrow r \ge \log_2 2*n$,

r – число проверочных символов, n – число информационных символов.

Граница Варшамова-Гильберта определяет нижнюю границу для числа проверочных разрядов r = n – k в случае кодов большой разрядности, необходимого для обеспечения заданного кодового расстояния d: $r \ge \log_2 \sum_{k=0}^{d-2} C_{n-k}^i$.

Определены оценки, которые дают представление о верхней границе d_0 при фиксированных n и k и оценка снизу для числа проверочных символов r при заданных k и d.

Корректирующие коды Хэмминга

Построение этих кодов базируется на принципе проверки на четность числа единиц в информационной группе кодового блока.

Поясним эту идею на примере простейшего корректирующего кода, называемого кодом с проверкой на четность (иначе с контролем по паритету). В таком коде к кодовым комбинациям безызбыточного первичного k-разрядного кода добавляется один дополнительный разряд (проверочный символ для проверки на четность). Если число единиц исходного кодового сообщения четное, то в дополнительном разряде формируют проверочный символ 0, а если нечетное, то в дополнительный разряд пишут 1. В результате общее количество символов 1 в любой передаваемой кодовой комбинации всегда будет четным. Таким образом, правило формирования проверочного символа строится следующим образом: $r_1 = i_1 \oplus i_2 \oplus ... \oplus i_k$,

где i – информационный символ, k – число информационных символов.

Добавление дополнительного разряда увеличивает общее число возможных комбинаций вдвое по сравнению с числом комбинаций исходного первичного кода, а условие четности разделяет все комбинации на разрешенные и запрещенные. Следовательно, код с проверкой на четность позволяет обнаруживать одиночную ошибку при приеме, т. к. она нарушает условие четности, переводя разрешенную комбинацию в запрещенную. Критерием правильности принятой кодовой комбинации является равенство нулю результата суммирования S по модулю 2 всех n символов кода, включая проверочный символ r_1. При наличии одиночной ошибки величина S принимает следующее значение:

S = $r_1 \oplus i_1 \oplus i_2 \oplus ... \oplus i_k$,

S – синдром , локатор ошибки. При S = 0 ошибки нет.

Этот код описывается согласно нашей терминологии: (n,k) = (n,n - 1). Так как d_0 = 2, то может быть обнаружена однократная ошибка, исправление производиться не может. Код с проверкой на четность может быть использован только для обнаружения однократных ошибок. Увеличивая число дополнительных проверочных символов и формируя их по определенным правилам, равными 0 или 1, можно усилить корректирующие свойства кода и позволить ему исправлять ошибки. На этом основано построение кодов Хэмминга.

Рассмотрим коды Хэмминга для случая исправления одиночной ошибки. В этом случае для каждого целого существует n = $2^r - 1$, k = $2^r - 1 - r$ разрядов.

Например, при r = 3 код Хэмминга имеет вид: (7,4). В этом случае четыре информационных символа дополняются тремя проверочными:

$(i_1, i_2, i_3, i_4) - (r_1, r_2, r_3)$.

Алгоритм кодирования Хэмминга, т. е. правило получения проверочных символов из информационных, имеет следующий вид:

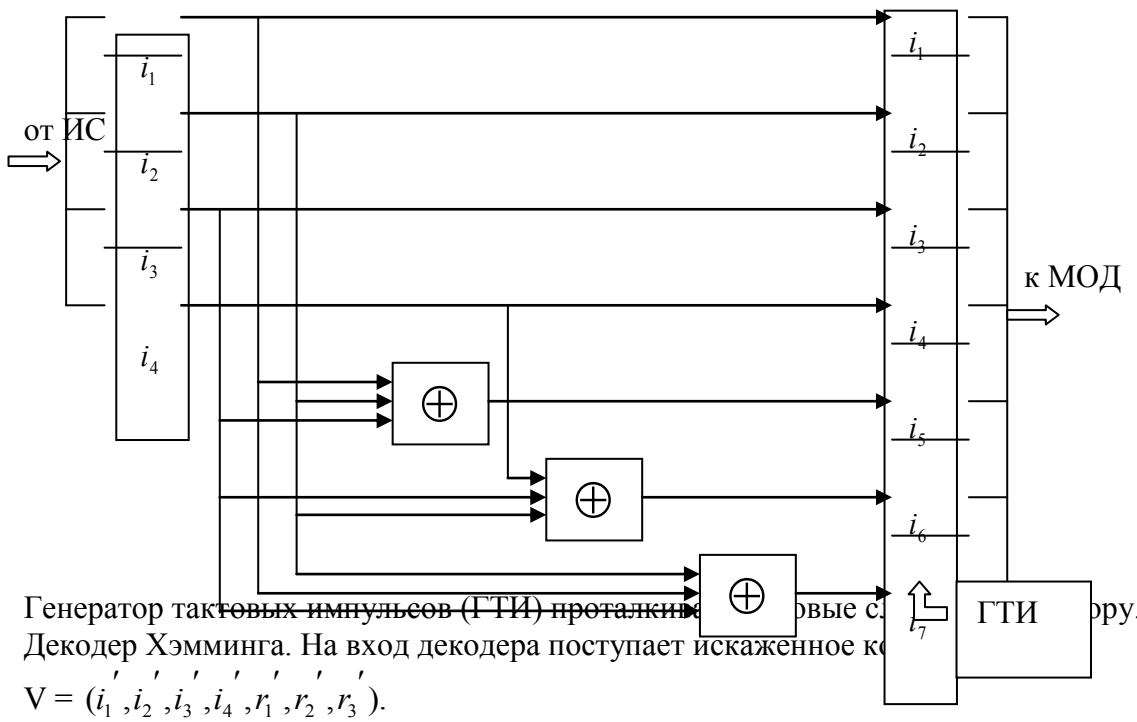
$$\begin{cases} r_1 = i_1 \oplus i_2 \oplus i_3, \\ r_2 = i_2 \oplus i_3 \oplus i_4, \\ r_3 = i_1 \oplus i_2 \oplus i_4. \end{cases}$$

Таким образом, существует $2^4 = 16$ всевозможных кодовых слов в коде x(7,4).

Номер перестановки	i_1	i_2	i_3	i_4	r_1	r_2	r_3
1	0	0	0	0	0	0	0
2	0	0	0	1	0	1	1
3	0	0	1	0	1	1	0
4	0	0	1	1	1	0	1
5	0	1	0	0	1	1	1
6	0	1	0	1	1	0	0

7	0	1	1	0	0	0	1
8	0	1	1	1	0	1	0
9	1	0	0	0	1	0	1
10	1	0	0	1	1	1	0
11	1	0	1	0	0	1	1
12	1	0	1	1	0	0	0
13	1	1	0	0	0	1	0
14	1	1	0	1	0	0	1
15	1	1	1	0	1	0	0
16	1	1	1	1	1	1	1

Принцип построения этой таблицы может быть технически реализован с помощью нижеприведенной схемы кодера Хэмминга (7,4).



Генератор тактовых импульсов (ГТИ) проталкивает кодовые с[illegible]ору.

Декодер Хэмминга. На вход декодера поступает искаженное ко[illegible]

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

Далее в декодере в режиме исправления ошибок строится трехсимвольная последовательность $\vec{S} = (S_1, S_2, S_3)$ - синдром ошибок:

$$\begin{cases} S_1 = r_1' \oplus i_1' \oplus i_2' \oplus i_3', \\ S_2 = r_2' \oplus i_2' \oplus i_3' \oplus i_4', \\ S_3 = r_3' \oplus i_1' \oplus i_2' \oplus i_4'. \end{cases}$$

Этот синдром представляет собой сочетание результатов проверки на четность соответствующих символов кодовой группы и характеризует определенную конфигурацию ошибок. При r = 3 имеется $2^3 = 8$ всевозможных синдромов, при этом $\vec{S} = (0,0,0)$ указывает на отсутсвие ошибок при приеме.

Всякому ненулевому синдрому $\vec{S}$ соответствует определенная конфигурация ошибок, которая и исправляется далее в декодере. Например, для (7,4) ненулевые синдромы и соответствующие конфигурации ошибок имеют следующий вид:

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

символе							

Каждая из ошибок имеет свой единственный синдром. При технической реализации декодера возможно создание такого цифрового корректора ошибок (дешифратора синдрома), который по соответствующему синдрому исправляет соответствующий символ в принятой кодовой комбинации. После внесения исправлений проверочные символы можно не выводить на выход декодера.

Циклические коды

Циклический код – такой групповой код, все базовые комбинации которого могут быть получены из одной путем циклического сдвига ее элементов.

Циклический сдвиг кодовой комбинации – перемещение ее элементов справа налево без нарушения порядка их следования, так что крайний левый элемент занимает место крайнего правого.

В теории циклических кодов принято записывать кодовые комбинации в виде полинома некоторой фиктивной переменной x:

$$C_1 a_1 \oplus C_2 a_2 \oplus C_3 a_3 \oplus \ldots \oplus C_i a_i \oplus \ldots \oplus C_q a_q \neq 0$$

где ai – значение символа кодовой комбинации на позиции i при нумерации справа налево;

xi – 1 – фиктивная переменная в степени номера позиции i без единицы.

Пример

Представить в виде полинома кодовую комбинацию 1011101.

$$a(x) = a^7 x^{7-1} + a^6 x^{6-1} + a^5 x^{5-1} + a^4 x^{4-1} + a^3 x^{3-1} + a^2 x^{2-1} + a^1 x^{1-1} =$$

$$= a^7 x^6 + a^6 x^5 + a^5 x^4 + a^4 x^3 + a^3 x^2 + a^2 x + a^1 =$$

$$= 1 \cdot x^6 + 0 \cdot x^5 + 1 \cdot x^4 + 1 \cdot x^3 + 1 \cdot x^2 + 0 \cdot x + 1 =$$

$$= x^6 + x^4 + x^3 + x^2 + 1.$$

Максимальная степень x в полиноме на единицу меньше числа элементов в кодовой комбинации.

Запись комбинации в виде полинома понадобилась для того, чтобы отобразить формализованным способом операцию циклического сдвига исходной кодовой комбинации. Допустим, задана исходная кодовая комбинация и соответствующий ей полином

$$a(x) = a^n x^{n-1} + a^{n-1} x^{n-2} + \ldots + a^i x^{i-1} + \ldots + a^2 x + a^1.$$

Умножим a(x) на x:

$$a(x) \cdot x = a^n x^n + a^{n-1} x^{n-1} + \ldots + a^i x^i + \ldots + a^2 x^2 + a^1 x$$

Так как максимальная степень x в кодовой комбинации длиной n не превышает (n – 1), то из правой части полученного выражения для получения исходного полинома необходимо вычесть an(xn – 1). Вычитание an(xn – 1) называется взятием остатка по модулю (xn – 1). Сдвиг исходной комбинации на i тактов можно представить следующим образом: a(x)xi – an(xn – 1), то есть умножением a(x) на xi и взятием остатка по модулю (xn – 1). Взятие остатка необходимо при получении многочлена степени, большей или равной n.

Возвращаясь к определению циклического кода и учитывая запись операций циклического сдвига кодовых комбинаций, можно записать порождающую матрицу циклического кода в следующем виде:

$$V = \begin{bmatrix} p(x) \\ p(x)\cdot x - C_2\left(x^n - 1\right) \\ p(x)\cdot x^2 - C_3\left(x^n - 1\right) \\ \dots \\ p(x)\cdot x^i - C_{i+1}\left(x^n - 1\right) \\ \dots \\ p(x)\cdot x^{m-1} - C_m\left(x^n - 1\right) \end{bmatrix},$$

где p(x)– исходная кодовая комбинация, на базе которой получены все остальные (m – 1) базовые комбинации;

Ci = 0 или Ci = 1 (“0”, если результирующая степень полинома p(x)xi не превосходит (n – 1), “1”, если превосходит).

Комбинация *p(x)* называется порождающей (образующей, генераторной) комбинацией. Для построения циклического кода достаточно верно выбрать *p(x)*. Затем все остальные кодовые комбинации получаются такими же, как и в групповом коде.

Порождающий полином должен удовлетворять следующим требованиям:

- p(x) должен быть ненулевым;
- вес p(x) не должен быть меньше минимального кодового расстояния: $\nu\left[p(x)\right] \geq d_{\min}$;
- p(x) должен иметь максимальную степень k (k – число избыточных элементов в коде);
- p(x) должен быть делителем полинома (xn – 1).

Выполнение условия 4 приводит к тому, что все рабочие кодовые комбинации циклического кода приобретают свойство делимости на *p(x)* без остатка. Учитывая это, можно дать другое определение циклического кода.

Циклический код – код, все рабочие комбинации которого делятся на порождающую без остатка.

Свойство делимости всех комбинаций на *p(x)* позволяет просто записать порождающую матрицу кода в каноническом виде:

$$V_{KAH} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 & c_1^1 & c_1^2 & \dots & c_1^k \\ 0 & 1 & 0 & \dots & 0 & c_2^1 & c_2^2 & \dots & c_2^k \\ 0 & 0 & 1 & \dots & 0 & c_3^1 & c_3^2 & \dots & c_3^k \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 1 & \dots & 1 & c_m^1 & c_m^2 & \dots & c_m^k \end{bmatrix} = \left[E_m \quad C_{m\times k}\right].$$

Для определения строк контрольной подматрицы $C_{m\times k}$ поступают следующим образом:

- любая строка единичной подматрицы записывается в виде полинома: $E_i(x) = x_m\text{-}i$;
- справа к ней приписывается *k* нулей, что равносильно умножению на x_k: $E_i(x)x_k$;
- результат делится на порождающий полином *p(x)*: $E_i(x)\cdot x^k = \xi(x)\cdot p(x) + r_i(x)$; при этом остаток $r_i(x)$ имеет степень не выше (k – 1) и содержит *k* элементов;
- рабочая комбинация циклического кода состоит из *m* элементов единичной подматрицы и из *k* элементов остатка $r_i(x)$ (он приписывается в “чистом” виде): $E_i(x)\cdot x^k \oplus r_i(x) = \xi(x)\cdot p(x)$

В настоящее время для облегчения процедуры построения циклических кодов их авторами найдены различные порождающие полиномы *p(x)* в зависимости от требований к коду. В частности, существуют таблицы с полиномами *p(x)* для циклических кодов с $d_{min} = 3$ (коды с $d_{min} = 3$

наиболее часто применяются на практике). Их можно найти в литературе по теории информации. Для построения кодов с $d_{min} = 4$ достаточно умножить выбранный полином *p(x),* найденный в таблице порождающих полиномов кодов с d*min* = 3, на (x + 1), что равносильно проверке на общую четность.

Суммируя изложенное, приведем процедуру выбора *p(x):*

определить число информационных элементов *m* (M = 2m) и число избыточных элементов *k* (если $d_{min} = 3$, 2k = m + k + 1);

найти *p(x)* степени *k* по таблице (если полиномов этой степени несколько, то выбирается любой).

Разработаны циклические коды, обеспечивающие произвольное минимальное кодовое расстояние $d_{min} = 5$. Они получили название БЧХ (Боуза-Чоудхури-Хоквингема, по имени разработчиков). Порождающие полиномы для таких кодов в зависимости от предъявляемых к ним требований, можно найти в таблице.

k	n	m	s	d_{min}	Образующий полином	
					Символическая запись	Запись в виде полинома
3	7	4	1	3	13	1011
4	15	11^{1}	1	3	23	10011
8	15	7	1	3	721	111010001
10^{1}	15	5	3	7	2467	10100110111
5	31	26^{2}	1	3	45	100101
10^{1}	31	21^{2}	2	5	3551	11101101001
15^{1}	31	16^{1}	3	7	107657	1000111110101111
25^{2}	31	11^{1}	5	11	5423325	101100010011011010101

МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ ПО ПРОВЕДЕНИЮ ЛАБОРАТОРНЫХ РАБОТ

Лабораторные работы имеют различный уровень сложности и на их выполнение требуется различное количество часов. Каждая предполагает самостоятельную работу студентов по освоению лекций и теоретического материала, вынесенного на самостоятельное изучение. Текущий контроль знаний осуществляется путем опроса студентов перед началом лабораторного занятия по вопросам, перечень которых приведен в каждой лабораторной работе.

Лабораторная работа 1. Оценка информационных характеристик систем

Определим пропускную способность канала как максимальное количество информации, которое можно передавать по нему в единицу времени:

$$C = \max\{I_{xy}\} / t_x \quad (\text{бит/с})$$

Для канала без помех справедливо условие $I_{xy} = H_x$, а потому его пропускная способность:

$C_{бп} = \max\{H_x\} / t_x = \log_2 m / t_x$

В частном случае передачи двоичных разрядов (m = 2) справедливо

$C_{бп} = 1/t_x$

Для нас важно, как соотносится величина $C_{бп}$ с потоком информации источника H`z, который определяется по формуле

$H`z = Hz/t_z$ (бит/с)

Пропускная способность канала используется полностью, когда H`z = C. Между тем, уменьшение энтропии H_z может привести к сокращению информационного потока. Чтобы его увеличить, требуется сократить время t_z. Если учесть, что $t_z = t_x * l_{ср}$, где $l_{ср}$ - средняя длина кода символа, то становится ясно: чтобы полнее использовать пропускную способность канала для любого источника, нужно рационально кодировать сообщения, по возможности сокращая величину $l_{ср}$.

Если записать условие полного использования пропускной способности канала H`z = C в развернутом виде, то для канала без помех оно будет иметь вид:

$H_z/t_z = \log_2 m/t_x$

а с учетом $t_z = t_x * l_c p$ и $\log_2 m = 1$ (при *m=2*) мы получим условие:

$l_{ср} = H_z$

Рассмотрим теперь вариант, когда помехи в канале вызывают появление ошибок с вероятностью p_0. В этом случае:

$C = \max \{Hx - Hx/y\}/ tx = (\log_2 m - Hx/y) / t_x$

Рассмотрим наиболее распространенный случай так называемого двоичного симметричного канала. При этом *m = 2* ($\log_2 m = 1$), а вероятности ошибки “переход "1" в "0" ” “переход "0" в "1" ” одинаковы.

Если теперь рассмотреть в качестве случайного события передачу разряда кода с ошибкой (вероятность p_0), то для определения энтропии, получим:

$Hx/y = Hy/x = -p_0 \log_2 p_0 - (1 - p_0) \log_2(1 - p_0)$

С учетом этого можно записать:

$C = [1 - p_0 \log_2 p_0 - (1 - p_0)\log_2(1 - p_0)]/tx$

Таким образом, пропускная способность симметричного двоичного канала с помехами определяется только скоростью передачи разрядов кода ($V_x = 1/t_x$) и вероятностью ошибок.

Клод Шеннон показал, что за счет кодирования пропускную способность канала с помехами также можно использовать максимально полно (напомним, что сама она будет ниже, чем у канала без помех).

Способ кодирования, который позволяет этого добиться, основан на использовании избыточных кодов, когда каждый информационный блок защищается контрольными разрядами и чем больше длина блока, тем меньше удельный вес этих избыточных разрядов, позволяющих обнаружить и исправить ошибки.

Источник И передает в канал непрерывное сообщение Z(t).

Формирователь сигналов Фс преобразует его в сигнал X (t), приспособленный для передачи по аналоговому каналу.

В линии связи ЛС на сигнал воздействуют случайные аддитивные помехи e(t) (для помех такого типа справедливо соотношение Y(t) = X(t) + e(t)).

Устройство распознавания сигнала восстанавливает сообщение Z(t) по полученному Y(t).

В этой схеме стадия кодирования вообще не рассматривается. Однако подход (кстати, предложенный опять-таки Клодом Шенноном) основан на тех же принципах, что и для дискретного канала, потому нам целесообразно рассмотреть этот вопрос именно здесь.

Вернемся к определению пропускной способности канала связи:

$$Сбп = \max\{I_{xy}\} / t_x = \max\{H_x\} / t_x$$

Величина tx в нашем случае соответствует шагу дискретизации сигнала d_t. Согласно теореме Котельникова, непрерывный сигнал можно полностью восстановить по его дискретным отсчетам, если шаг дискретизации d_t вдвое меньше периода самой высокочастотной составляющей f_m сигнала ($d_t = 1/2f_m$). Учитывая, что любой физический канал связи всегда имеет ограниченную полосу частот, которые он в состоянии пропустить, величину f_m (а следовательно и d_t) можно определить исходя из характеристик канала.

Если значение d_x конечно, то непрерывный канал можно рассматривать как дискретный с объемом алфавита $m = x_m/d_x + 1$. Если к тому же в канале отсутствуют помехи (Hx/y = 0), то можно записать:

$$C = \max \{Hx\} / d_t = 2f_m * \log_2 m = 2f_m * \log_2 (x_m/ d_x + 1)$$

Отсюда видно, что пропускная способность непрерывного канала без помех (d_x -> 0) стремится к бесконечности. Однако, в реальном канале помехи присутствуют всегда, при этом сколько бит информации удается "нагрузить" на один дискретный отсчет, зависит от соотношения мощности полезного сигнала на входе приемника и помехи $P_c/P_п$.

Клод Шеннон показал, что в случае наиболее "неприятной" помехи типа "белый шум", чья мощность равномерно распределена во всей полосе частот канала, справедливо соотношение:

$C_n = f_m \log_2(P_c/P_п + 1)$

Доказательство этой теоремы Шеннона о пропускной способности непрерывного канала весьма громоздко и мы не станем его рассматривать. Остановимся на анализе самой формулы. Итак пропускная способность непрерывного канала с помехами:

- пропорциональна ширине полосы частот канала f_m;
- возрастает с увеличением отношения полезный сигнал/помеха (в этом случае будет уверенно распознаваться на фоне помех);
- не равна нулю даже при $P_c << P_п$ (передачу информации принципиально можно вести сигналами более слабыми, чем помехи).

ЗАДАНИЕ

1. Ознакомиться с теоретической частью, используя дополнительную литературу.
2. Исходя из полученных у преподавателя исходных данных (количества передаваемых сообщений N), рассчитать пропускную способность дискретного канала связи с шумами и непрерывного канала связи без шумов.
3. Провести программный контроль выполнения пункта 2 на примере исходных данных полученных у преподавателя.
4. Отчет.

Контрольные вопросы.

1. Сформулируйте теорему Шеннона для канала без помех.
2. Как отличается трактовка величины Hz для случаев "посимвольного" и "цепочечного" эффективного кодирования?
3. Почему при вероятности ошибки p_0 =1 пропускная способность канала имеет ту же величину, что и при p_0 = 0? Как практически можно использовать такой канал?
4. В чем суть теоремы Шеннона для канала с помехами?
5. Как практически можно избежать потери информации в канале с помехами?

Лабораторная работа 2. Построение оптимального кода.

Оптимальным кодированием называется процедура преобразования символов первичного алфавита m_1 в кодовые слова во вторичном алфавите m_2, при которой средняя длина сообщений во вторичном алфавите имеет минимально возможную для данного m_2 длину.

Оптимальными именуются коды, представляющие кодируемые понятия кодовыми словами минимальной средней длины. Оптимальные коды относятся к классу *префиксных кодов*, т.е. каждая кодовая комбинация имеет свою длину и ни одна не является началом другой, более длиной.

В сообщениях, составленных из кодовых слов оптимального кода, статистическая избыточность сведена к минимуму, в идеальном случае – к нулю.

Основная теорема кодирования для каналов связи без шумов доказывает принципиальную возможность построения оптимальных кодов. Из нее однозначно вытекают методика построения и свойства оптимальных кодов.

Одно из основных положений этой теории заключается в том, что при кодировании сообщения, разбитого на N - буквенные блоки, можно, выбрав N достаточно большим, добиться, чтобы среднее число двоичных элементарных сигналов, приходящихся на одну букву исходного сооб-

щения, было сколь угодно близким к H/log m. Разность L - $\frac{H}{\log m}$ будет тем меньше, чем больше H, а H достигает максимума при равновероятных и взаимонезависимых символах. Отсюда вытекают основные свойства оптимальных кодов:

минимальная средняя длина кодового слова оптимального кода обеспечивается в случае, когда избыточность каждого кодового слова сведена к минимуму (в идеальном случае - к нулю);

кодовые слова оптимапльного кода должны строиться из равновероятных и взаимонезависимых символов.

Из свойств оптимальных кодов вытекают принципы их построения. 1 - выбор каждого кодового слова необходимо производить так, чтобы содержащееся в нем количество информации было максимальным; 2 - буквам первичного алфавита, имеющим большую вероятность, присваивается более короткие кодовые слова во вторичном алфавите.

Принципы оптимального кодирования определяют методику построения оптимальных кодов. *Построение оптимального кода по методу Шеннона-Фано* для ансамбля из M сообщений сводится к следующей процедуре:

) множество из M сообщений располагают в порядке убывания вероятностей;

2) первоначальный ансамбль кодируемых сигналов разбивают на две группы таким образом, чтобы суммарные вероятности сообщений обеих групп были по возможности равны;

3) первой группе присваивают символ 0, второй - символ 1;

4) каждую из групп делят на две подгруппы так, чтобы их суммарные вероятности были по возможности равны;

5) первым подгруппам каждой из групп вновь присваивают 0, а вторым – 1, в результате получают вторые цифры кода. Затем каждую из четырех подгрупп вновь делят на равные (с точки зрения суммарной вероятности) части и т.д. - до тех пор, пока в каждой из них останется одна буква.

Пример. Построим оптимальный код для передачи сообщений, в которых вероятности появления букв первичного алфавиты равны: A_1=1/4, A_2=1/4, A_3=1/8, A_4=1/8, A_5=1/16, A_6=1/16, A_7=1/16, A_8=1/16.

Решение. Построение ведем по общей методике. Оптимальный код для данных условий представлен в табл. 1.

буква	Вероятность появления буквы	Кодовое слово после разбиения				Число знаков в кодовом слове	L(i) pi
		1-го	2-го	3-го	4-го		
A1	1/4	0	0			2	0,5
A2	1/4	0	1			2	0,5
A3	1/8	1	0	0		3	0,375
A4	1/8	1	0	1		3	0,375
A5	1/16	1	1	0	0	4	0,25
A6	1/16	1	1	0	1	4	0,25
A7	1/16	1	1	1	0	4	0,25
A8	1/16	1	1	1	1	4	0,25

Проверка оптимальности кода осуществляется путем сравнения энтропии кодируемого (первичного) алфавита со средней длиной кодового слова во вторичном алфавите.

Для рассматриваемого примера энтропия источника сообщений

H=- $\sum_{i=1}^{N} p_i \log p_i$ =2,75 бит/символ.

Среднее число двоичных знаков на букву кода

$\sum_{i}^{N} l(i) * p_i =$

L= 2*0.5+2*0.375+4*0.25=2,75 бит/символ,

где l(i) – длина i-й кодовой комбинации;

pi - вероятность появления i-го символа комбинации длиной в l(i).

Таким образом, H=L, т.е. код, оптимален для данного ансамбля сообщений.

Коды, представляющие первичные алфавиты с неравномерным распределением символов, имеющие минимальную среднюю длину кодового слова во вторичном алфавите, называются оптимальными неравномерными кодами (ОНК).

Максимально эффективными будут те ОНК, у которых

$\log_2 m \sum_{i=1}^{N} l(i) p_i = l_{ср}$=H,

где m и N – символы соответственно вторичного и первичного алфавитов.

Эффективность ОНК оценивают при помощи *коэффициента статистического сжатия*

$$K_{c.c} = \frac{H_{\max}}{l_{cp}} = \frac{\log_2 N}{\log_2 m \sum_{i=1}^{N} l(i) p_i},$$

характеризующего уменьшение количества двоичных знаков на символ сообщения при применении ОНК по сравнению с применением методов нестатистического кодирования, и *коэффициента относительной эффективности*

$$K_{о.э} = \frac{H}{l_{cp}} = \frac{-\sum_{i=1}^{N} p_i \log_2 p_i}{\log_2 m \sum_{i=1}^{N} l(i) p_i},$$

показывающего, насколько используется статистическая избыточность передаваемого сообщения.

Метод Шеннона-Фано не единственный способ построения оптимальных кодов. Хорошо известна и широко применяется методика построения ОНК при помощи кодовых деревьев. Впервые она была описана Хаффменом.

Хаффмен показал, что для получения минимально возможной длины кода основания m с числом взаимонезависимых букв первичного алфавита N

$$l_{cp} = \sum_{i=1}^{N} l(i) p_i$$

необходимо и достаточно выполнение следующих условий:

1) если выписать символы в порядке убывания вероятностей $p_i > p_j$, то при i < j, l(i) < l(j);

2) два последних, но не больше чем m кодовых слова равны по длительности и различаются лишь значениями последнего символа, при этом

$$\frac{N - n_0}{m - 1} = a$$

где m – число качественных признаков вторичного алфавита, а n_0 – число наименее вероятных сообщений, объединяемых на первом этапе построения кодового дерева; кроме того a – целое положительное число;

3) любая возможная последовательность N-1 кодовых слов должна сама быть кодовой комбинацией.

Исходя из данных условий, Хаффмен предложил следующий метод построения ОНК. Символы первичного алфавита выписываются в порядке убывания вероятностей. Последние n0 символов, где $2 \le n_0 \le m$ и $N - n_0 / m - 1$ - целое число, объединяют в некоторый новый символ с вероятностью, равной сумме вероятностей объединяемых символов. Последние символы с учетом

образованного символа вновь объединяют и получают новый, вспомогательный символ. Опять выписывают символы в порядке убывания вероятностей с учетом вспомогательного - и так до тех пор, пока вероятности m оставшихся символов после $N-n_0/m-1$-го выписывания не дадут в сумме 1.

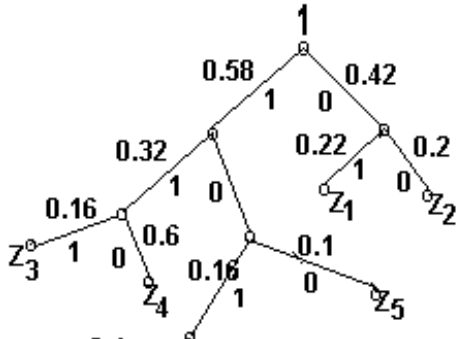
На практике обычно не производят многократного выписывания вероятностей символов, а обходятся геометрическими построениями, суть которых для кодов с числом качественных признаков m=2 сводится к тому, что символы кодируемого алфавита попарно объединяются в новые, начиная с символов, имеющих наименьшую вероятность, а затем, с учетом вероятностей вновь образованных символов, опять производят попарное объединение символов с наименьшими вероятностями и таким образом строят двоичное кодовое дерево, в вершине которого стоит символ с вероятностью 1.

Пример. Используя методику Хаффмена, осуществить эффективное кодирование ансамбля знаков с вероятностями соответственно: z1=0,22; z2=0,20; z3=0,16; z4=0,16; z5=0,10; z6=0,10; z7=0,04; z8=0,02.

Решение.

Знаки	вероятности	Вспомогательные столбцы						
		I	II	III	IV	V	VI	VII
Z_1	0.22	0 .22	0 .22	**0 .26**	**0 .32**	**0 .42**	**0 .58**	**1**
Z_2	0.20	0 .20	0 .20	0 .22	0 .26	0 .32	0 .42	
Z_3	0.16	0 .16	0 .16	0 .20	0 .22	0 .26		
Z_4	0.16	0 .16	0 .16	0 .16	0 .20			
Z_5	0.10	0 .10	**0 .16**	0 .16				
Z_6	0.10	0 .10	0 .10					
Z_7	0.04	**0 .06**						
Z_8	0.02							

Для наглядности построим кодовое дерево. Из точки соответствующей вероятности 1 направляем две ветви, причем ветви с большей вероятностью присваиваем символ 1, а с меньшей 0. Такое последовательное ветвление продолжаем до тех пор, пока не дойдем до вероятности каждой буквы.



Теперь, двигаясь по кодовому дереву сверху вниз, можно записать для каждой буквы соответствующую ей кодовую комбинацию:

Z_1	Z_2	Z_3	Z_4	Z_5	Z_6	Z_7	Z_8
01	00	111	110	100	1011	10101	10100

При построении ОНК для вторичных алфавитов с m=2 методы Шеннона-Фано и Хаффмена дают в большинстве случаев одинаковые результаты.

ЗАДАНИЕ

1. Получить у преподавателя задание и ознакомиться с ним.
2. Построить код по методу Шеннона-Фано и проверить его оптимальность.
3. Построить код по методу Хаффмена и кодовое дерево.
4. Провести программный контроль выполнения 2, 3 пунктов на примере случайных сообщений.
5. Подготовить отчет и сдать работу.

Контрольные вопросы.

1. Какой код называется оптимальным?
2. Способы построения оптимальных кодов.
3. В чем заключается сущность оптимального кодирования и практический результат его применения?
4. Как оценивается эффективность ОНК?
5. Какие коды называются оптимальными неравномерными кодами?

Лабораторная работа 3. Построение двоичного группового кода (кода Хемминга)

Код Хемминга - один из наиболее распространенных систематических кодов, имеющих простой и удобный для технической реализации алгоритм обнаружения и исправления одиночной ошибки.

Код Хемминга строится так, чтобы полученный при проверках результат (r1,r2,...rn-k) прямо указал номер искаженного разряда и тем самым упростил декодирование.

Уравнения кодирования для определения проверочных разрядов находят приравниванием проверочных уравнений нулю при отсутствии ошибок. Проверочные разряды различаются внутри кодовой комбинации на местах, соответствующих номеру.

Для вычисления основных параметров кода задается количество либо информационных символов, либо информационных комбинаций – $N = 2^{n_u}$. При помощи следующих формул вычисляются n и nk:

$$2^{n_k} \geq n+1 \qquad 2^n = 2^{n_k} * 2^{n_u}$$

Соотношение между n, nk и nи для кода Хэмминга представлены в табл.1.

Таблица 1

N	$n_и$	n_k	n	$n_и$	n_k
1	0	1	9	5	4
2	0	2	10	6	4
3	1	2	11	7	4
4	1	3	12	8	4
5	2	3	13	9	4

6	3	3	14	10	4
7	4	3	15	11	4
8	4	4	16	11	5

Зная основные параметры корректирующего кода, определяют, какие позиции сигналов будут рабочими, а какие – контрольными. Практика показала, что номера контрольных символов удобно выбирать по закону 2_i, где $i=0,1, 2, 3\ldots$ - натуральный ряд чисел. Номера контрольных символов в этом случае равны 1, 2, 4, 8, 16, 32… Затем определяют значения контрольных коэффициентов (0 или 1), руководствуясь следующим правилом: сумма единиц на проверочных позициях должна быть четной. Если эта сумма четна – значение контрольного коэффициента 0, в противном случае – 1.

Проверочные позиции выбирают следующим образом. Составляют табличку для ряда натуральных чисел в двоичном коде. Число ее строк равно $n=n_u+n_k$. Первой строке соответствует проверочный коэффициент a_1, второй a_2 и т.д.

0001 a_1	0101 a_5	1001 a_9
0010 a_2	0110 a_6	1010 a_{10}
0011 a_3	0111 a_7	1011 a_{11}
0100 a_4	1000 a_8	

Затем выявляют проверочные позиции, выписывая коэффициенты по следующему принципу: в первую проверку входят коэффициенты, которые содержат 1 в младшем разряде, т.е. a_1, a_3, a_5, a_7, a_9, a_{11} и т.д.; во вторую – содержащие 1 во втором разряде, т.е. a_2, a_3, a_6, a_7, a_{10} и т.д.; в третью – содержащие 1 в третьем разряде, и т.д. Номера проверочных коэффициентов соответствуют номерам проверочных позиций, что позволяет составить общую таблицу проверок (табл. 2).

Таблица 2

номер проверки	Проверочные позиции (П)	номер контрол. символа
1	1, 3, 5, 7, 9,11,. . .	1
2	2, 3, 6, 7, 10, 11, 14, 15, 18, 19, 22, 24,…	2
3	4, 5, 6, 7, 12, 13, 14, 15, 20, 21, 22,23,…	4
4	8, 9, 10, 11, 12, 13, 14, 15, 24, 25, 26, 27, 28, 29, 30, 31, 40, 41, 42,…	8

Пример. Требуется исправить любую одиночную ошибку при передаче комбинации 0101, т.е. $n_и=4$.

Решение. Согласно табл.1 минимальное число контрольных символов $n_k=3$, при этом n=7. Контрольные коэффициенты будут расположены на позициях 1, 2, 4. Составляем макет корректирующего кода и записываем его во вторую колонку в табл. 3. Пользуясь табл.2, определим значения коэффициентов К1, К2, К3.

Первая проверка: сумма П1+П3+П5+П7 должна быть четной, а сумма К1+0+1+1 будет четной при К1=0.

Вторая проверка: сумма П2+П3+П6+П7 должна быть четной, а сумма К2+0+0+1 будет четной при К2=1.

Третья проверка: сумма П4+П5+П6+П7 должна быть четной, а сумма К3+1+0+1 будет четной при К3=0.

Окончательное значение искомой комбинации корректирующего кода записываем в третью колонку в таблице 3.

Таблица 3

Позиция символов корректирующего кода	Кодовое слово

	без значений контрольных коэффициентов	со значениями контрольных коэффициентов
1	К1	0
2	К2	1
3	0	0
4	К3	0
5	1	1
6	0	0
7	1	1

Предположим, что в канале связи под действием помех произошло искажение и вместо 0100101 было принято 0100111. Для обнаружения ошибки производят проверки на четность.

Первая проверка: сумма П1+П3+П5+П7=0+0+1+1 четна. В младший разряд номера ошибочной позиции записываем 0.

Вторая проверка: сумма П2+П3+П6+П7=1+0+1+1 нечетна. Во второй разряд номера ошибочной позиции записываем 1.

Третья проверка: сумма П4+П5+П6+П7=0+1+1+1 нечетна. В третий разряд номера ошибочной позиции записываем 1.

Номер ошибочной позиции 101=6. Следовательно, символ шестой позиции следует изменить на обратный, и мы получим правильную кодовую комбинацию.

ЗАДАНИЕ

1. Ознакомиться с теоретической частью, используя дополнительную литературу.

2. Построить код Хемминга по заданным исходным данным (число информационных разрядов k).

3. Составить систему уравнений кодирования для определения проверочных разрядов для кода Хемминга по пункту 2.

4. Провести программный контроль выполнения 2 и 3 пунктов на примере случайных кодовых комбинаций.

5. Отчет.

Контрольные вопросы.

1. На каких позициях проверочные символы в коде Хэмминга?
2. Что такое информационные и проверочные символы?
3. Какими графическими и геометрическими способами можно представить коды? Приведите пример.
4. Что такое кодовое расстояние, как оно определяется между двумя комбинациями двоичного кода?
5. Каким соотношением связаны информационные, проверочные символы и минимальное кодовое дерево?

Лабораторная работа 4. Построить групповой код и показать процесс исправления ошибки в произвольном разряде корректирующего кода

Систематический код - групповой *n*-значный код, в котором из n символов, образующих кодовую комбинацию, n_u символов информационные, а $n_k = n - n_u$ - избыточные, предназначенные для проверки.

Систематические коды удобно задавать при помощи производящей матрицы. Число строк матрицы равно $n_и$, число столбцов равно n.

$$C=\begin{Vmatrix} a_{11}\ a_{12}\ \dots\ a_{1n_u}\ P_{11}\ P_{12}\ \dots\ P_{1n_k} \\ a_{21}\ a_{22}\ \dots\ a_{2n_u}\ P_{21}\ P_{22}\ \dots\ P_{2n_k} \\ \dots\ \dots\ \dots\ \dots\ \dots\ \dots\ \dots\ \dots \\ a_{n_u1}\ a_{n_u2}\ \dots\ a_{n_un_u}\ P_{n_u1}\ P_{n_u2}\ \dots\ P_{n_un_k} \end{Vmatrix}$$

Производящая матрица С может быть представлена при помощи двух матриц И и П (информационной и проверочной). Число столбцов матрицы П равно n_k, число столбцов матрицы И равно $n_и$.

$$C=\begin{Vmatrix} a_{11} & a_{12} & \dots & a_{1n_u} \\ a_{21} & a_{22} & \dots & a_{2n_u} \\ \dots & \dots & \dots & \dots \\ a_{n_u 1} & a_{n_u 2} & \dots & a_{n_u n_u} \end{Vmatrix} \begin{Vmatrix} P_{11} & P_{12} & \dots & P_{1n_k} \\ P_{21} & P_{22} & \dots & P_{2n_k} \\ \dots & \dots & \dots & \dots \\ P_{n_u 1} & P_{n_u 2} & \dots & P_{n_u n_k} \end{Vmatrix} = \| \text{И} \| \, \| \text{П} \|$$

Теорией и практикой установлено, что в качестве матрицы И удобно брать единичную матрицу

$$\begin{Vmatrix} 10\dots & 000 \\ 01\dots & 000 \\ \dots & \dots \\ 00\dots & 001 \end{Vmatrix}$$

При выборе матрицы П исходят из следующих рассуждений: чем больше единиц в разрядах проверочной матрицы П, тем ближе соответствующий порождаемый код к оптимальному.

Критерием оптимальности таких кодов является соблюдение условия

$$2^{n-n_u}-1 \geq \sum_{i=1}^{r} C_n^i$$

r – число ошибок.

С другой стороны, число единиц в матрице П определяет число сумматоров по модулю 2 в шифраторе и дешифраторе, т.е. чем больше единиц в матрице П, тем сложнее аппаратура. Но даже если основным требованием к аппаратуре будет ее простота, вес каждой строки матрицы П должен быть не менее $W_П \geq d_0 - W_И$, где $W_И$ – вес соответствующей строки матрицы И. Если матрица И – единичная, то $W_И=1$ (при $W_И>1$ усложнилось бы как построение кодов, так и их техническая реализация).

Производящая матрица позволяет получить все возможные комбинации кода суммированием по модулю 2 всех возможных сочетаний строк матрицы.

Пример. Построить матрицу для группового кода, способного исправлять одиночную ошибку при передаче 16 символов первичного алфавита.

Кодовое расстояние $d_0=3$. Так как число информационных разрядов кода nи=4 ($16=24=2^{n_u}$), то число строк производящей матрицы С должно быть равно 4. Число столбцов матрицы С равно *n*; *n* – длина кода, в свою очередь, равна $n_и+n_k$; n_k – число корректирующих разрядов, равное

$n_k=\log 2\{5+[\log 25]\}=\log 28=3$.

Следовательно, число столбцов, содержащих контрольные разряды, должно быть равно 3, а общее число столбцов матрицы С равно $n_и+n_k=4+3=7$.

Так как вес каждой строки проверочной матрицы П должен быть $W_П \geq d_0 - W_И$, то в качестве ее строк могут быть выбраны трехзначные двоичные комбинации с числом единиц ≥2: 111; 110; 101; 011.

Как видно из примера, основным требованиям могут удовлетворять несколько матриц. Выбор той или иной из матриц, возможных для данного $n_и$, n_k, и d_0, определяется по дополнительным требованиям: минимум корректирующих разрядов или максимальная простота аппаратуры.

В процессе декодирования систематического кода осуществляются проверки, идея которых в общем виде может быть представлена следующим образом:

$$P_j \oplus \sum_{i=1}^{n_u} P_{ij} a_i = S_j, \qquad j=1, 2,\dots nk . \qquad (2)$$

Для каждой конкретной матрицы существует своя, одна - единственная система проверок. Проверки производятся по следующему правилу: в первую вместе с проверочным разрядом p_1 входят информационные разряды, соответствующие единицам первого столбца проверочной матрицы П, во вторую - второй проверочный разряд p_2 и информационные разряды, соответствующие единицам второго столбца проверочной матрицы и т.д. Число проверок равно числу проверочных разрядов корректирующего кода n_k. В результате проверок образуется проверочный вектор S1, S2,…S^{n_k}, который называют синдромом. Если вес синдрома равен нулю, то принятая комбинация считается безошибочной. Но если хотя бы один разряд проверочного вектора содержит единицу, принятая комбинация содержит ошибку. Исправление ошибки производится по виду синдрома, так как каждому ошибочному разряду соответствует один - единственный проверочный вектор.

Вид синдрома для каждой конкретной матрицы может быть определен при помощи контрольной матрицы Н, представляющей собой транспонированную матрицу П, дополненную единичной матрицей I, число столбцов которой равно числу проверочных разрядов кода

$$H = \left\| \Pi^T I_{n_k} \right\|$$

Столбцы такой матрицы - значение синдрома для разряда, соответствующего номеру столбца матрицы Н.

Пример. Групповой код построен по матрице

$$C = \begin{Vmatrix} 1\,0\,0\,0\,0\,1\,1 \\ 0\,1\,0\,0\,1\,0\,1 \\ 0\,0\,1\,0\,1\,1\,0 \\ 0\,0\,0\,1\,1\,1\,1 \end{Vmatrix}$$

Показать процесс исправления ошибки в произвольном разряде корректирующего кода, информационная часть которого - четырехразрядные комбинации натурального двоичного кода.

Решение. Кодовое расстояние $d_0 = 3$. Число проверочных символов $n_к=3$; $n_и=4$.

Производящая матрица С в виде информационной матрицы И и проверочной матрицы П может быть представлена следующим образом

$$C = \begin{Vmatrix} 1\,0\,0\,0\,0\,1\,1 \\ 0\,1\,0\,0\,1\,0\,1 \\ 0\,0\,1\,0\,1\,1\,0 \\ 0\,0\,0\,1\,1\,1\,1 \end{Vmatrix} = \begin{Vmatrix} 1\,0\,0\,0 \\ 0\,1\,0\,0 \\ 0\,0\,1\,0 \\ 0\,0\,0\,1 \end{Vmatrix} = \begin{Vmatrix} 0\,1\,1 \\ 1\,0\,1 \\ 1\,1\,0 \\ 1\,1\,1 \end{Vmatrix}$$

И П

Согласно принципу построения системы проверки (2) система проверок для кодов, построенных по матрице С, будет иметь вид

$$P_1 \oplus a_2 \oplus a_3 \oplus a_4 = S_1$$
$$P_2 \oplus a_1 \oplus a_3 \oplus a_4 = S_2$$
$$P_3 \oplus a_1 \oplus a_2 \oplus a_4 = S_3.$$

Чтобы знать, какая комбинация значений разрядов синдрома S_1, S_2, S_3 будет соответствовать ошибке в определенном разряде принятой комбинации, строим контрольную матрицу Н, ее строками являются столбцы матрицы П, дополненные единичной транспонированной матрицей I, размерность которой определяется числом избыточных разрядов кода, т.е. в нашем случае равная 3. Таким образом,

a1	a2	a3	a4	P1	P2	P3
0	1	1	1	1	0	0

$$H=\begin{Vmatrix} 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{Vmatrix}$$

Если разряды синдрома соответствуют первому столбцу матрицы H, т.е. $S_1=0$, $S_2=1$, $S_3=1$, то ошибка в первом разряде принятой комбинации; если синдром имеет вид 101, что соответствует второму столбцу матрицы H, то ошибка во втором разряде; синдром 001 соответствует ошибке в третьем проверочном разряде кода.

В качестве примера проверки корректирующих свойств кода используем комбинации кода вида

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

Пусть сбои произошли в первом разряде первой комбинации и в четвертом разряде второй, т.е.:

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

Находим проверочные векторы согласно системе проверок.

Для первой комбинации: P1=1, P2=1, P3=0.

$$P_1 \oplus a_2 \oplus a_3 \oplus a_4 = 1 \oplus 1 \oplus 0 \oplus 0 = 0$$

$$P_2 \oplus a_1 \oplus a_3 \oplus a_4 = 1 \oplus 0 \oplus 0 \oplus 0 = 1$$

$$P_3 \oplus a_1 \oplus a_2 \oplus a_4 = 0 \oplus 0 \oplus 1 \oplus 0 = 1$$

Синдром – 0 1 1 показывает, что в первом разряде символ следует заменить на обратный.

Для второй комбинации:

$1 \oplus 0 \oplus 1 \oplus 1 = 1$

$1 \oplus 0 \oplus 1 \oplus 1 = 1$

$0 \oplus 0 \oplus 0 \oplus 1 = 1$

Синдром – 1 1 1, ошибка в четвертом разряде.

ЗАДАНИЕ

1. Ознакомиться с теоретической частью, используя дополнительную литературу.

2. Исходя из полученных у преподавателя исходных данных (количества передаваемых сообщений N), рассчитать необходимое число информационных и контрольных разрядов для систематического кода, обнаруживающего и исправляющего одиночные ошибки.

3. Составить порождающую и проверочную матрицы, а также уравнения проверки по пункту 2, исходя из количества информационных разрядов.

4. Провести программный контроль выполнения 2 и 3 пунктов на примере некоторых случайных кодовых комбинаций рассчитанной ранее разрядности.

5. Отчет.

Контрольные вопросы.

1. Приведите классификацию корректирующих кодов по способу введения и использования избыточности, по структуре кода.

2. Какие среди систематических кодов имеют наибольшую практическую значимость и почему?

3. Что такое синдром ошибки?

4. Как получают проверочную матрицу при формировании систематических кодов и чем объясняется такое требование ее построения?

5. Какими выражениями удобно пользоваться для практических расчетов при определении числа контрольных разрядов с d=3? d=4?

Лабораторная работа 5. Построить циклический код по заданным характеристикам и проверить его свойства по обнаружению и исправлению ошибок

Циклические коды широко применяются при передаче данных в сетях и системах телеобработки данных. По способу и системе коррекции ошибок они относятся к блочным неразделимым кодам.

Основной принцип использования основывается на формировании комбинации кода путем циклического сдвига разрядов влево образующего многочлена. Эта операция аналогична процедуре умножения на X:

$$(X4 + X3 + X2 + 1) * X = X5 + X4 + X3 + X$$
$$0011101 \qquad\qquad 0111010$$

Таким образом, при соответствующем выборе образующего многочлена любая разрешенная комбинация может быть получена в результате умножения образующего многочлена на некоторый другой многочлен.

Основная идея обнаружения и исправления ошибок заключается в делении комбинации кода на образующий многочлен, т.е.:

$$\frac{G(X)*X}{K(X)} = Q(X) + \frac{R(X)}{K(X)}$$

где G(X) - комбинация кода;
K(X) - образующий многочлен;
Q(X) - результат деления;
R(X) - остаток.

Если остаток равен нулю, то исследуемая комбинация разрешенная и код не содержит ошибки. В противном случае имеется ошибка.

Пример. Найти образующий многочлен для следующих параметров кода: d_0=3, n=7.

Решение. Вычислим число проверочных *m* и информационных *k* символов.

m = log (n + 1) = 3
k = n - m = 7 - 3 = 4.

По таблице неприводимых многочленов найдем для m = 3 и d = 3 образующий многочлен вида 1101 или K(X) = X3 + X2 + 1.

Вычислим проверочные разряды и получим образующую матрицу путем умножения всех комбинаций кода на образующий многочлен.

0000		0000000
0001		0001101
0010		0011010
0011		0010111
0100		0110100
0101		0111001
0110		0101110
0111		0100011
1000	* (1101) =	1101000
1001		1100101
1010		1110010
1011		1111111
1100		1001110
1101		1010001
1110		1000110
1111		1001011

Проверим возможность кода на обнаружение и исправление ошибок. Возьмем комбинацию 0111001. Разделим ее на образующий многочлен

```
0111001 ¦1101
  1101   +------
 ---------  101
0001101
```

```
  1101
--------
 0000
```

Остаток равен 0 - следовательно, это разрешенная комбинация.

Исказим третий разряд.

```
 0111101 ¦ 1101
  1101    +------
 -----      101
 0001001
    1101
   ----
    0100
```

Остаток свидетельствует об обнаружении ошибки.

Правила построения циклических кодов исправляющих одну ошибку

1. Расчет соотношения между разрядами:

$n = m + k$,

где m - число проверочных разрядов;

k - число информационных разрядов;

$m = [\log (n + 1)]$

или

$m = [\log \{(k + 1) + [\log (k + 1)]\}]$.

2. Выбор образующего многочлена производится по таблицам неприводимых двоичных многочленов, где m - степень многочлена, d - число единиц в комбинации.

3. Выбор параметров единичной матрицы производится исходя из условия, что число столбцов матрицы определяется числом информационных разрядов.

4. Определение элементов дополнительной матрицы производится по остаткам от деления последней строки транспонированной матрицы на образующий многочлен (это еще один способ формирования образующей матрицы).

5. Образующая матрица составляется путем дописывания элементов дополнительной матрицы справа от единичной матрицы или путем умножения элементов единичной матрицы на образующий многочлен.

6. Комбинациями исходного кода являются строки образующей матрицы и всевозможные суммы по модулю 2 различных сочетаний строк образующей матрицы.

7. Обнаружение и исправление ошибок происходит по остаткам от деления принятой комбинации G(X) на образующий многочлен K(X). Если деление без остатка, то ошибки нет. Для исправления ошибки:

а) принятая комбинация делится на образующий многочлен;

б) подсчитывается вес остатка. Если $W \in S$, где S – допустимое число исправляемых ошибок, то принятая комбинация складывается по модулю 2 с полученным остатком. Сумма даст исправленную комбинацию;

в) если $W > S$, делим полученную в результате циклического сдвига комбинацию на образующий многочлен. Если в остатке $W \in S$, то складываем делимое с остатком. Затем производим циклический сдвиг вправо комбинации, полученной в результате суммирования последнего делимого с остатком. Если после первого циклического сдвига и последующего деления остаток получается таким, что его вес $W > S$, то процедура повторяется до тех пор, пока $W \in S$. Затем производится циклический сдвиг вправо на столько разрядов, на сколько была сдвинута принятая комбинация. В результате получаем исправленную комбинацию.

ЗАДАНИЕ

1. Получить у преподавателя задание и ознакомиться с ним.

2. Вычислить параметры кода *d, m, k, p, l, S*. Найти образующий многочлен, воспользовавшись таблицей неприводимых многочленов.

3. Проверить, имеются ли ошибки в исследуемой комбинации, при наличии ошибок - исправить их.

4. Провести программный контроль выполнения 4, 5 пунктов на примере случайных кодовых комбинаций.

5. Подготовить отчет и сдать работу.

Контрольные вопросы.

1. В чем заключаются основные идеи обнаружения и исправления ошибок циклическим кодом?

2. Что такое кодовое расстояние?

3. Чем отличается представление циклическим кодом для d = 3 и d = 5? где d - кодовое расстояние?

4. Какие существуют способы формирования комбинаций циклического кода?

5. В чем достоинство циклических кодов?

6. Что такое транспонированная матрица для циклического кода и ее размерность?

МЕТОДИЧЕСКИЕ УКАЗАНИЯ К ПРАКТИЧЕСКИМ ЗАНЯТИЯМ

Тема 1. Оценка информационных характеристик каналов

Основные понятия. Канал связи, информационная скорость, техническая скорость, помехи, пропускная способность, канальная матрица, матрица объединения, информационная полнота, условные вероятности, количество информации.

Вопросы для самоконтроля

1. Решение некоторых логических задач с помощью подсчёта информации.
2. Энтропия и информация о письменной русской речи.
3. Пропускная способность линий связи.
4. Передача информации при наличии помех.
5. Информационная скорость и техническая скорость.

Практические задания

Задача 1. Чему равна условная энтропия и энтропия объединения, если канальная матрица имеет вид

$$p(b/a) = \begin{|1 & 0 & 0 & 0|\\ |0 & 1 & 0 & 0|\\ |0 & 0 & 1 & 0|\\ |0 & 0 & 0 & 1|\end{}$$

Задача 2. Канал связи, в котором передаются сигналы А1, А2, А3, А4, описан следующей канальной матрицей:

$$p(a,в) = \begin{|0.01 & 0.1 & 0.11 & 0.02|\\ |0.02 & 0.02 & 0.05 & 0.07|\\ |0.2 & 0.08 & 0.07 & 0.03|\\ |0.02 & 0.03 & 0.06 & 0.01|\end{}$$

Найти долю информационных потерь, которые припадают на сигнал А2 при передаче сигналов А1 - А4 по данному каналу связи.

Задача 3. Определить пропускную способность дискретного канала связи, описанному матрицей

$$p(A,B) = \begin{|0.1 & 0 & 0|\\ |0.1 & 0.3 & 0|\\ |0 & 0.1 & 0.4|\end{}$$

Задача 4. Определить все возможные информационные характеристики канала связи, в котором взаимосвязь источника с приемником может быть описана матрицей вида:

$$p(A,B) = \begin{|0.2 & 0.1 & 0|\\ |0.2 & 0 & 0.2|\\ |0.1 & 0.1 & 0.1|\end{}$$

Задача 5. Вероятность появления сигналов на входе приемника сообщений равна соответственно рв1=0.2, рв2=0.3, рв3=0.5. Канал связи описан следующей канальной матрицей:

$$p(A/B)=\begin{|0.97 & 0 & 0.01|\\ |0.02 & 0.98 & 0.01|\\ |0.01 & 0.02 & 0.98|\end{}$$

Определить информационную скорость.

Задача 6. В результате статистических испытаний канала связи получены следующие условные вероятности перехода одного сигнала в другой: $p(b_1/a_1)=0.85$, $p(b_2/a_1)=0.1$, $p(b_3/a_1)=0.05$, $p(b_1/a_2)=0.09$, $p(b_2/a_2)=0.91$, $p(b_3/a_2)=0$, $p(b_1/a_3)=0.08$, $p(b_3/a_3)=0.92$. Построить канальную матрицу и определить общую условную и взаимную энтропию сообщений, передаваемых по данному каналу связи, если на выходе источника сигналы появились с равной вероятностью.

Задача 7. Обладают ли информационной полнотой следующие информационные характеристики:

$$p(a1)=0.3,\ p(a2)=0.2,\ p(a3)=0.5 \quad p(a/b)=\begin{|1 & 0 & 0|\\ |0 & 1 & 0|\\ |0 & 0 & 1|\end{}.$$

Тема. Оценка количества информации в дискретном канале связи

Энтропия, условная энтропия, взаимная энтропия, информационные потери, канальная матрица, матрица объединения, информационная полнота, условные вероятности, количество информации, канал связи, дискретный.

Вопросы для самоконтроля

1. Энтропия как мера неопределённости.
2. Энтропия сложных событий. Условная энтропия.
3. Понятие об информации I(a,b).
4. Свойства энтропии: H(a,b), $H_b(a)$.
5. Количественная оценка информации.

Задача 1. Влияние помех в канале связи описывается следующим распределением условных вероятностей:

$$p(b/a)=\begin{|0.98 & 0.01 & 0.01|\\ |0.15 & 0.75 & 0.1|\\ |0.3 & 0.2 & 0.5|\end{}$$

Вычислить количество информации, которое переносится одним символом сообщения при равновероятном появлении символов в сообщении. Вычислить количество информации в сообщении, составленном из 400 букв первичного алфавита.

Задача 2. Влияние помех в канале связи описывается следующим распределением условных вероятностей:

$$p(b/a)=\begin{|0.98 & 0.01 & 0.01|\\ |0.15 & 0.75 & 0.1|\\ |0.3 & 0.2 & 0.5|\end{}$$

Вычислить количество информации, которое переносится одним символом сообщения при вероятностях $p(a1)=0.7$; $p(a2)=0.2$; $p(a3)=0.1$. Вычислить количество информации в сообщении, составленном из 400 букв первичного алфавита.

Задача 3. Определить количество информации при передаче k сообщений по каналу связи, описанному следующей канальной матрицей:

$$p(b/a)=\begin{|0.25 & 0.25 & 0.25 & 0.25|\\ |0.25 & 0.25 & 0.25 & 0.25|\\ |0.25 & 0.25 & 0.25 & 0.25|\\ |0.25 & 0.25 & 0.25 & 0.25|\end{}$$

если на выходе источника сообщений символы встречаются с вероятностями $p(A1)=0.8$, $p(A2)=0.1$, $p(A3)=p(A4)=0.5$.

Задача 4. Взаимодействие двух систем A и B описывается следующей матрицей:

$$\begin{|0.4 & 0.1 & 0|\end{}$$

$$p(A,B)=\begin{vmatrix}0 & 0.2 & 0.1\\ 0 & 0 & 0.2\end{vmatrix}$$

Определить безусловную энтропию системы А и системы В, и H(a /b).

Задача 5. В сообщении, составленном из пяти качественных признаков, последние используются с разной частотой, т.е. вероятности их различны и равны соответственно p1=0.8, p2=0.15, p3=0.03, p4=0.015 и p5=0.005. Всего в сообщении принято 20 знаков. Определить количества информации на букву сообщения и во всем сообщении. Каково было бы количество информации в данном сообщении, если бы все признаки имели равную вероятность?

Задача 6. Определить частную условную энтропию относительно каждого символа источника сообщений при передаче по каналу связи, описанному следующей канальной матрицей:

$$p(A,B)=\begin{vmatrix}0.2 & 0 & 0\\ 0.1 & 0.2 & 0\\ 0 & 0.1 & 0.4\end{vmatrix}$$

Тема. Методы оптимального кодирования

Код, метод Шеннона-Фано, метод Хаффмена, избыточность, Энтропия, условная энтропия, взаимная энтропия, информационные потери, канальная матрица, матрица объединения, информационная полнота, условные вероятности, количество информации.

Вопросы для самоконтроля

1. За счет чего при оптимальном кодировании уменьшается средняя длина кодовой комбинации?
2. До какого предела может быть уменьшена средняя длина кодовой комбинации при оптимальном кодировании?
3. В чем преимущество методики построения оптимального кода, предложенной Хаффменом , по сравнению с методикой Шеннона-Фано?
4. Какому основному условию должны удовлетворить оптимальные коды?
5. Перечислите сложности, возникающие при использовании оптимальных кодов.

Задача 1. Построить код для передачи сообщений, составленных из алфавита с распределением вероятностей: A_1=0,18; A_2=0,18; A_3=0,18; A_4=0,18; A_5=0,1; A_6=0,09; A_7=0,09. Построение провести по методу Шеннона-Фано.

Задача 2. Построить ОНК для передачи сообщений, составленных из некоторого условного словаря со следующими вероятностями появления слов в тексте: запятая – 0,37; товарищ – 0,13; свою – 0,125; в – 0,11; и – 0,08; труд – 0,06; бой - 0,05; хранить – 0,023; беззаветно – 0,002.

Задача 3. Построить код для передачи сообщений, составленных из алфавита с распределением вероятностей: A_1=0,18; A_2=0,18; A_3=0,18; A_4=0,18; A_5=0,1; A_6=0,09; A_7=0,09. Построение провести по методу Хаффмена с числом качественных признаков равное 3.

Задача 4. Построить ОНК по методу Хаффмена для вторичного алфавита с числом качественных признаков равное 3, если символы кодируемого алфавита имеют следующее распределение вероятностей: A_1=0,37; A_2=0,25; A_3=0,18; A_4=0,1; A_5=0,06; A_6=0,02; A_7=0,02.

Задача 5. Построить ОНК по методам Шеннона-Фано и Хаффмена, если символы источника сообщений появляются с вероятностями: $A_1=A_2=A_3=A_4$=0,19; $A_5=A_6=A_7$=0,08. Сравнить, насколько полученные коды близки к оптимальному.

Тема. Методы обнаружения и исправления одиночных ошибок

Проверочная матрица, система проверок, четность, образующая матрица, информационная матрица, кодовое расстояние, синдром.

Вопросы для самоконтроля

1. Какой код называют систематическим?
2. Как построить проверочную матрицу кода?
3. Дайте определение образующей матрицы.
4. Что такое синдром?
5. Что определяет кодовое расстояние?

Задача 1. Пусть требуется передать 16 сообщений. Построить систематический код, исправляющий одну ошибку.

Задача 2. Требуется исправить любую одиночную ошибку при передаче комбинации 0101, т.е. $n_u=4$.

Задача 3. Построить матрицу для группового кода, способного исправлять одиночную ошибку при передаче 32 символов первичного алфавита. И показать процесс исправления ошибки.

Задача 4. Определить кодовое расстояние между двоичными векторами: 110011011; 100110010.

Задача 5. Какой вид имеет производящая матрица группового кода, оптимального с точки зрения минимума корректирующих разрядов при максимуме информационных разрядов, для использования его в системе телемеханики, проектируемой для передачи не менее 2000 различных сообщений.

Тема. Методы образования циклического кода

Циклические коды, образующий многочлен, опознаватель ошибок, неприводимые многочлены, производящая матрица, суммирование по модулю два.

Вопросы для самоконтроля

1. Как построить образующую матрицу циклического кода?
2. Каким требованиям должен удовлетворять образующий многочлен циклического кода?
3. Как находят опознаватели ошибок в случае циклического кода?
4. Поясните процесс декодирования циклического кода.
5. Как найти все комбинации кода?

Задача 1. Используя метод образующих матриц, построить циклический код, исправляющий одинарные ошибки при передаче комбинаций четырехзначного двоичного кода на все сочетания.

Задача 2. Методом умножения образующего многочлена на многочлены четырехзначного двоичного кода на все сочетания построить циклический код.

Задача 3. Методом циклического сдвига строки образующей матрицы и зеркальной ее комбинации построить циклический код с $d_0=3$ и $n_u=5$.

Задача 4. Показать процесс исправления одиночной ошибки в принятой кодовой комбинации 1100101.

Задача 5. При помощи образующей матрицы, полученной в результате умножения единичной матрицы на образующий многочлен $x^3 + x + 1$, построить циклический код, исправляющий одиночную ошибку в любом из четырех информационных разрядов.

МЕТОДИЧЕСКИЕ РЕКОМЕНДАЦИИ К САМОСТОЯТЕЛЬНОЙ РАБОТЕ СТУДЕНТОВ

Самостоятельная работа студентов по дисциплине «Теория информации» проводится с целью:

систематизации и закрепления полученных теоретических знаний и практических умений по дисциплине;

углубления и расширения теоретических знаний;

формирования умений использовать полученные знания в новых условиях;

развития познавательных и творческих способностей;

формирования самостоятельности мышления, способности к саморазвитию, самореализации.

В учебном процессе выделяют два вида самостоятельной работы – **аудиторную,** которая выполняется под руководством преподавателя, и **внеаудиторную**, которая выполняется по заданию преподавателя, но без его непосредственного участия в определенные сроки и с последующей проверкой результатов на занятиях.

Перед выполнением обучающимися внеаудиторной самостоятельной работы преподаватель проводит инструктаж по выполнению заданий, которые включают цель задания, его содержание, сроки выполнения, объем работы, основные требования к результатам работы, критерии оценки результатов внеаудиторной самостоятельной работы.

В качестве форм контроля внеаудиторной самостоятельной работы обучающихся используется тестирование, самоотчеты, контрольные работы.

Основные формы самостоятельной учебной работы:

работа над конспектом лекции: лекции – основной источник информации по многим предме-

там, позволяющий не только изучить материал, но и получить представление о наличии других источников, сопоставить разные взгляды на основные проблемы данного курса. Лекции предоставляют возможность «интерактивного» обучения, когда есть возможность задавать преподавателю вопросы и получать на них ответы. Поэтому имеет смысл находить время для хотя бы беглого просмотра информации по материалу лекций (учебники, справочники и пр.) и непонятные, а также дискуссионные моменты обсуждать с преподавателем, другими студентами; □

подготовка к практическому занятию: производится, как правило, с использованием методических пособий, состоит в теоретической подготовке (особенно для семинаров) и выполнении практических заданий (решение задач, ответы на вопросы и т.д.).

доработка конспекта лекции с применением учебника, методической литературы, дополнительной литературы: этот вид самостоятельной работы студентов особенно важен в том случае, когда изучаемый предмет содержит много неоднозначно трактуемых вопросов, проблем. Тогда преподаватель заведомо не может успеть изложить различные точки зрения, и студент должен ознакомиться с ними по имеющейся литературе. Кроме того, рабочая программа дисциплины предполагает рассмотрение некоторых относительно несложных тем только во время самостоятельных занятий, без чтения лектором; □

подбор, изучение, анализ и конспектирование рекомендованной литературы;

самостоятельное изучение отдельных тем, параграфов; □

консультации по сложным, непонятным вопросам лекций, семинаров, зачетов; □

подготовка к зачету: данная форма СРС может быть весьма разнообразной по своей сути, так как сам зачет бывает различным. Он проводится обычно по итогам семестра перед сессией в письменной или устной форме, причем преподаватель может включать в него вопросы как практических занятий, так и лекционных (что особенно уместно, когда по данному предмету не сдается экзамен). Главное отличие зачета от экзамена – почти всегда не пяти-, а двухбалльная система оценки (сдал – не сдал), что делает его получение несколько более простым делом. С другой стороны, порой процедура его сдачи достаточно сложна, а иногда применяется и пятибалльная оценка (так называемый зачет с оценкой). Таким образом, для сдачи зачета необходимо, прежде всего, выполнить все требования преподавателя, что предполагает знание этих требований. Нужно как можно раньше выяснить, какие вопросы предстоит готовить и каковы правила самой процедуры (учитывается ли посещаемость, надо ли пропущенные занятия отрабатывать, а если надо, то каким образом и т.д.). Практика показывает, что хорошее посещение занятий является почти полной гарантией получения зачета, так как тогда можно быть в курсе всех требований преподавателя. И, напротив, большое количество пропусков может осложнить жизнь даже сильному студенту. Кроме того, необходимо учитывать, что проблемы могут появиться при распространенном подходе студента к практическим занятиям, когда многие работают первые месяцы вполсилы, накапливая задолженности по выполнению рефератов, практических заданий, конспектов и пр., а перед сессией пытаются все это сделать за одну неделю. Старайтесь распределять силы равномерно по всей дистанции семестра, и тогда зачетная неделя перед сессией будет не самой напряженной, а самой разгрузочной;

подготовка к экзамену: один из самых ответственных видов самостоятельной работы, и в то же время возможность сэкономить большое количество времени в период сессии, если эту подготовку начинать заблаговременно. Одно из главных правил – представлять себе общую логику предмета, что достигается проработкой планов лекций, составлением опорных конспектов, схем, таблиц. Фактически основной вид полготовки к экзамену – «свертывание» большого объема информации в компактный вид, а также тренировка в ее «развертывании» (примеры к теории, выведение одних закономерностей из других и т.д.). Владение этими технологиями обеспечивает, пожалуй, более половины успеха. Тем более что преподаватель обычно замечает в течение семестра целенаправленную подготовку такого студента и может поощрить его тем или иным способом. Надо также правильно распределить силы, не только готовясь к самому экзамену, но и позаботившись о допуске к нему (часто это хорошее посещение занятий, выполнение в назначенный срок практических заданий, активность на семинарах). Наконец, необходимо выяснить условия проведении, самого экзаменационного испытания, использовав для этой цели прежде всего консульта-

цию (хотя преподаватель обычно касается этой темы заранее): количество и характер вопросов, форма проведения (устно или письменно), возможность использовать при подготовке различные материалы и пособия (таблицы, схемы, тетради для практических занятий и т.д.).

ЛИТЕРАТУРА

1. Гуменюк А.С. Прикладная теория информации [Электронный ресурс]: учебное пособие/ А.С. Гуменюк, Н.Н. Поздниченко— Электрон. текстовые данные.— Омск: Омский государственный технический университет, 2015.— 189 c.— Режим доступа: http://www.iprbookshop.ru/58097.html.— ЭБС «IPRbooks»

2. Санников В.Г. Теория информации и кодирования [Электронный ресурс]: учебное пособие/ В.Г. Санников— Электрон. текстовые данные.— М.: Московский технический университет связи и информатики, 2015.— 95 c.— Режим доступа: http://www.iprbookshop.ru/61558.html.— ЭБС «IPRbooks»

3. Акулиничев Ю.П. Теория и техника передачи информации [Электронный ресурс]: учебное пособие/ Ю.П. Акулиничев, А.С. Бернагрдт— Электрон. текстовые данные.— Томск: Томский государственный университет систем управления и радиоэлектроники, Эль Контент, 2012.— 210 c.— Режим доступа: http://www.iprbookshop.ru/13984.html.— ЭБС «IPRbooks»

4. Чернышев А.Б. Теория информационных процессов и систем [Электронный ресурс]: учебное пособие/ А.Б. Чернышев, В.Ф. Антонов, Г.Б. Суюнова— Электрон. текстовые данные.— Ставрополь: Северо-Кавказский федеральный университет, 2015.— 169 c.— Режим доступа: http://www.iprbookshop.ru/63140.html.— ЭБС «IPRbooks»

5. Математические методы теории передачи сигналов. Часть 3 [Электронный ресурс]: учебное пособие/ — Электрон. текстовые данные.— М.: Московский технический университет связи и информатики, 2012.— 48 c.— Режим доступа: http://www.iprbookshop.ru/61494.html.— ЭБС «IPRbooks»

6. Балюкевич Э.Л. Теория информации [Электронный ресурс]: учебное пособие/ Э.Л. Балюкевич— Электрон. текстовые данные.— М.: Евразийский открытый институт, 2009.— 215 c.— Режим доступа: http://www.iprbookshop.ru/10863.html.— ЭБС «IPRbooks»

7. Гультяева Т.А. Основы теории информации и криптографии [Электронный ресурс]: конспект лекций/ Т.А. Гультяева— Электрон. текстовые данные.— Новосибирск: Новосибирский государственный технический университет, 2010.— 88 c.— Режим доступа: http://www.iprbookshop.ru/44987.html.— ЭБС «IPRbooks»

8. Теория информации [Электронный ресурс] : метод. указания к практ. занятиям / АмГУ, ФМиИ ; сост. С. Г. Самохвалова. - Благовещенск : Изд-во Амур. гос. ун-та, 2017. - 45 с.

Режим доступа: http://irbis.amursu.ru/DigitalLibrary/AmurSU_Edition/9589.pdf

СОДЕРЖАНИЕ