

**ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«ОРЕНБУРГСКИЙ ГОСУДАРСТВЕННЫЙ АГРАРНЫЙ УНИВЕРСИТЕТ»**

**МЕТОДИЧЕСКИЕ УКАЗАНИЯ ДЛЯ ОБУЧАЮЩИХСЯ
ПО ОСВОЕНИЮ ДИСЦИПЛИНЫ**

Б1.В.06 Теория информации

Направление подготовки (специальность) 09.03.01 Информатика и вычислительная техника

Профиль образовательной программы Автоматизированные системы обработки информации и управления

Форма обучения заочная

СОДЕРЖАНИЕ

1. Конспект лекций.....	3
1.1 Лекция № 1 Понятие информации. Модели детерминированных и случайных сигналов. Преобразование непрерывных сигналов в дискретные.....	3
1.2 Лекция № 2 Оценка информационных характеристик источников сообщений.....	5
1.3 Лекция № 3 Эффективное кодирование. Введение в теорию помехоустойчивого кодирования.....	8
2. Методические материалы по проведению практических занятий.....	9
2.1 Практическое занятие № ПЗ-1 Понятие информации. Модели детерминированных и случайных сигналов. Преобразование непрерывных сигналов в дискретные.....	9
2.2 Практическое занятие № ПЗ-2 Меры неопределенности дискретных множеств и непрерывных случайных величин. Количество информации как мера снятой неопределенности.....	10
2.3 Практическое занятие № ПЗ-3 Оценка информационных характеристик источников сообщений.....	10
2.4 Практическое занятие № ПЗ-4 Информационные характеристики каналов связи.....	11
2.5 Практическое занятие № ПЗ-5 Эффективное кодирование. Введение в теорию помехоустойчивого кодирования.....	11
2.6 Практическое занятие № ПЗ-6 Построение групповых кодов. Циклические коды.....	11
2.7 Практическое занятие № ПЗ-7 Матричные представления в теории кодирования. Кодирование линейными последовательными машинами.....	12

1. КОНСПЕКТ ЛЕКЦИЙ

1.1 Лекция №1 (2 часа).

Тема: «Понятие информации. Модели детерминированных и случайных сигналов. Преобразование непрерывных сигналов в дискретные»

1.1.1 Вопросы лекции:

1. Понятие информации.
2. Модели детерминированных и случайных сигналов.
3. Преобразование непрерывных сигналов в дискретные.

1.1.2 Краткое содержание вопросов:

1. Понятие информации.

Информация — это осознанные сведения об окружающем мире, которые являются объектом хранения, преобразования, передачи и использования.

Сведения — это знания, выраженные в сигналах, сообщениях, известиях, уведомлениях и т. д. Каждого человека окружает информация различных видов.

Основные виды информации по ее форме представления, способам ее кодирования и хранения, что имеет наибольшее значение для информатики, это:

- графическая или изобразительная — первый вид, для которого был реализован способ хранения информации об окружающем мире в виде наскальных рисунков, а позднее в виде картин, фотографий, схем, чертежей на бумаге, холсте, мраморе и др. материалах, изображающих картины реального мира;
- звуковая — мир вокруг нас полон звуков и задача их хранения и тиражирования была решена с изобретением звукозаписывающих устройств в 1877 г.; ее разновидностью является музыкальная информация — для этого вида был изобретен способ кодирования с использованием специальных символов, что делает возможным хранение ее аналогично графической информации;
- текстовая — способ кодирования речи человека специальными символами — буквами, причем разные народы имеют разные языки и используют различные наборы букв для отображения речи; особенно большое значение этот способ приобрел после изобретения бумаги и книгопечатания;
- числовая — количественная мера объектов и их свойств в окружающем мире; особенно большое значение приобрела с развитием торговли, экономики и денежного обмена; аналогично текстовой информации для ее отображения используется метод кодирования специальными символами — цифрами, причем системы кодирования могут быть разными;

- видеoinформация— способ сохранения «живых» картин окружающего мира, появившийся с изобретением кино.

2. Модели детерминированных и случайных сигналов.

Детерминированным называется сигнал, значения которого в любые моменты времени являются известными величинами. В противном случае сигнал называют случайным или стохастическим (от греческого слова *stochastic* – догадка). Каждый конкретный вид случайного сигнала $X(t)$, представляющего собой функцию времени, называют реализацией. Каждую реализацию можно представить бесконечной совокупностью зависимых или независимых случайных величин.

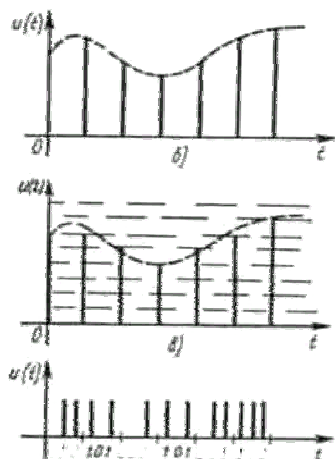
Случайный сигнал описывается статистически с помощью различных вероятностных характеристик.

3. Преобразование непрерывных сигналов в дискретные.

В любую систему информация поступает в виде сигналов. Различные параметры физических процессов с помощью датчиков обычно преобразуются в электрические сигналы. Как правило, ими являются непрерывно изменяющиеся ток или напряжение, но возможно поступление и импульсных сигналов, как, например, в радиолокации. Печатный текст отображается буквами, цифрами и другими знаками.

Хотя поступающую информацию можно хранить, передавать и обрабатывать как в виде непрерывных, так и в виде дискретных сигналов, на современном этапе развития информационной техники предпочтение отдается дискретным сигналам, поэтому сигналы, как правило, преобразуются в дискретные. С этой целью каждый непрерывный сигнал подвергается операциям квантования по времени (дискретизации) и по уровню.

Под *дискретизацией* подразумевают преобразование функции непрерывного времени в функцию дискретного времени, представляемую совокупностью величин, называемых координатами, по значениям которых исходная непрерывная функция может быть восстановлена с заданной точностью. Роль координат часто выполняют мгновенные значения функции, отсчитанные в определенные моменты времени.



Под *квантованием* подразумевают преобразование некоторой величины с непрерывной шкалой значений в величину, имеющую дискретную шкалу значений. Оно сводится к замене любого мгновенного значения одним из конечного множества разрешенных значений, называемых *уровнями квантования*.

Изменение вида сигнала $u(t)$ (рис.5.1,а) в результате проведения операции дискретизации показано на рис. 5.1,б, а в результате совместного проведения операций дискретизации и квантования — на рис. 5.1, в.

Число уровней квантования на рис. 5.1,в равно 8. Обычно их значительно больше. Передача такого множества различных по уровню импульсов даже на небольшие расстояния применяется крайне редко. Если провести нумерацию уровней, то их передача сведется к передаче чисел. Тогда, выразив эти числа в какой-либо системе счисления, можно обойтись меньшим множеством передаваемых сигналов. Как правило, дискретный сигнал преобразуется в последовательность чисел, выраженных в двоичном коде. Каждое дискретное значение сигнала представляется в этом случае последовательностью сигналов двух уровней. Наличие или отсутствие импульса на определенном месте интерпретируется единицей или нулем в соответствующем разряде двоичного числа.

Цифровая форма представления сигнала $u(t)$ (рис. 5.1,а) показана на рис. 5.1,г. Для восьми уровней достаточно трех двоичных разрядов. Импульсы старших разрядов расположены крайними справа.

1. 2 Лекция №2 (2 часа).

Тема: «Оценка информационных характеристик источников сообщений»

1.2.1 Вопросы лекции:

1. Энтропия источника
2. Избыточность источника
3. Производительность источника.

1.2.2 Краткое содержание вопросов:

1. Энтропия источника

Большинство реальных источников формирует сообщения с различным количеством информации, однако, при решении практических задач необходимо знать среднее количество информации приходящееся на одно сообщение. Среднее количество информации определяется как математическое ожидание количества информации в сообщении.

$$M[I(a_i)] = \sum_{i=1}^{M_a} P(a_i) I(a_i) = - \sum_{i=1}^{M_a} P(a_i) \log_2 P(a_i) = H(A); \text{ бит/сообщ.} \quad (9)$$

где M_a — количество возможных сообщений источника.

Величина $H(A)$ называется **энтропией источника** и характеризует среднее количество информации приходящейся на одно сообщение.

Приведенное выражение используется для определения энтропии источников дискретных сообщений. Для непрерывных сообщений $a(t)$ энтропия теоретически стремится к бесконечности, т. к. сообщение может принимать бесконечное число значений, следовательно $P(a_i) \rightarrow 0$, а $I(a_i) \rightarrow \infty$. Однако если сообщение подвергнуть дискретизации и представить его конечным числом квантованных значений по уровню L , то можно определить среднее количество информации в одном отсчете (*энтропию отсчета*):

$$H_{отсч}(A) = - \sum_{i=1}^{L_{кв}} p_i \log_2 p_i, \text{ бит/сообщ.} \quad (10)$$

где p_i — вероятность появления в квантованном сообщении i -го уровня.

$L_{кв}$ — количество уровней квантования.

Если осуществить предельный переход устремив L к бесконечности, то получится величина, называемая **дифференциальной энтропией**.

$$h(A) = - \int_{-\infty}^{\infty} p(a) \log_2 p(a) da, \text{ бит/сообщ.} \quad (11)$$

Энтропия является объективной информационной характеристикой источника сообщений. Она всегда положительна.

Свойства энтропии.

1. Энтропия равна нулю, если одно сообщение достоверно ($P(a_i)=1$), а другие не возможны.

2. Энтропия максимальна, когда все сообщения равновероятны, и растет с увеличением равновероятных сообщений.

3. Энтропия обладает свойством аддитивности, т. е. энтропии различных источников можно складывать.

2. Избыточность источника

Под избыточностью понимают наличие в сообщении «лишних» элементов, т. е. элементов не несущих смысловой нагрузки (например, союзы, предлоги). Данные элементы могут быть восстановлены за счет статистических взаимосвязей между другими элементами сообщения. Например, союзы, предлоги и знаки препинания можно восстановить, зная правила построения предложений. Таким образом:

избыточность — это мера сокращения сообщения без потери информации, за счет статистических взаимосвязей между элементами сообщения.

Количественной мерой информации является коэффициент избыточности:

$$c_{\text{и}} = [H_{\text{max}}(A) - H(A)] / H_{\text{max}}(A) \quad (12)$$

где $H(A)$ — энтропия, вычисленная на основе учета статистических характеристик сообщений;

$H_{\text{max}}(A)$ — максимальная энтропия источника, которая согласно второму свойству равна:

$$H_{\text{max}}(A) = \log_2 M_a; \text{ бит/сообщ.} \quad (13)$$

Наличие избыточности при передаче сообщений имеет свои положительные и отрицательные стороны. Сообщение, обладающее избыточности требует большего времени передачи и, соответственно большего времени занятия канала. Однако повышение избыточности приводит к увеличению помехоустойчивости сообщения. Она способствует обнаружению и исправлению ошибок в принятых сообщениях. Это связано с тем, что для формирования сообщения используются не все возможные комбинации символов, а лишь определенные (разрешенные), которые заносятся в специальные справочники (словари). При приеме сообщения с элементами (словами) которых нет в справочниках говорит о наличии ошибки и ошибочный элемент может быть заменен похожим по написанию или подходящим по смыслу. Все языки обладают избыточностью равной $c_{\text{и}} \gg 0,5$.

3. Производительность источника.

Производительность источника — это среднее количество информации создаваемое источником в единицу времени.

Производительность источника дискретных сообщений определяется как:

$$H'_{\text{д.с.}}(A) = H(A) / t_{\text{ср}}; \text{ бит/с} \quad (14)$$

где $t_{\text{ср}}$ — средняя длительность сообщения:

$$t_{\text{ср}} = t_{\text{н}} / n; \text{ с} \quad (15)$$

где $t_{\text{н}}$ — время, в течении которого было сформировано n сообщений.

Производительность источника непрерывных сообщений определяется как:

$$H'_{\text{н.с.}}(A) = f_{\text{д}} H_{\text{отсч}}(A); \text{ бит/с} \quad (16)$$

где $f_{\text{д}}$ — частота дискретизации ($2F_{\text{max}}$).

Если сообщение квантуется равновероятными уровнями, т. е. $p_i=1/L$, то производительность источника может быть определена как:

$$N'_{н.с.}(A)=f_d \log_2 L; \text{ бит/с} \quad (17)$$

1.3 Лекция № 3 (2 часа)

Тема: «Эффективное кодирование. Введение в теорию помехоустойчивого кодирования»

1.3.1 Вопросы лекции:

1. Эффективное кодирование
2. Основы теории помехоустойчивого кодирования

1.3.2 Краткое содержание вопросов:

1. Эффективное кодирование

Эффективное кодирование – это процедуры направленные на устранение избыточности. Основная задача эффективного кодирования – обеспечить, в среднем, минимальное число двоичных элементов на передачу сообщения источника. В этом случае, при заданной скорости модуляции обеспечивается передача максимального числа сообщений, а значит максимальная скорости передачи информации.

Пусть имеется источник дискретных сообщений, алфавит которого K . При кодировании сообщений данного источника двоичным, равномерным кодом, потребуется $L_{pk} = \log_2 K$ двоичных элементов на кодирование каждого сообщения.

Если вероятности $P(a_i)$ появления всех сообщений источника равны, то энтропия источника (или среднее количество информации в одном сообщении) максимальна и равна $H_{\max}(x) = \log_2 K$.

В данном случае каждое сообщение источника имеет информационную емкость $\log_2 K = L_{pk}$ бит, и очевидно, что для его кодирования (перевозки) требуется двоичная комбинация не менее L_{pk} элементов. Каждый двоичный элемент, в этом случае, будет переносить 1 бит информации.

Если при том же объеме алфавита сообщения не равновероятны, то, как известно, энтропия источника будет меньше

$$H_{\text{ред}}(x) = -\sum_{i=1}^K p(a_i) \log_2 p(a_i) < H_{\max}(x)$$

Если и в этом случае использовать для перевозки сообщения L_{pk} -разрядные кодовые комбинации, то на каждый двоичный элемент кодовой комбинации будет приходиться меньше чем 1 бит.

Появляется избыточность, которая может быть определена по следующей формуле

$$D = \frac{H_{\max}(x) - H_{\text{реал}}(x)}{H_{\max}(x)} = 1 - \frac{H_{\text{реал}}(x)}{H_{\max}(x)} \left[\frac{\text{бит}}{\text{на элемент}} \right].$$

Среднее количество информации, приходящееся на один двоичный элемент комбинации при кодировании равномерным кодом

$$\frac{H_{\text{реал}}(x)}{L_{\text{рк}}} = \frac{H_{\text{реал}}(x)}{H_{\max}(x)}$$

2. Основы теории помехоустойчивого кодирования

Для согласования источника дискретных сообщений с каналом связи используют корректирующее (помехоустойчивое) кодирование сообщений (кодирование с обнаружением и (или) исправлением ошибок). Кодирование дискретных сообщений является одним из основных путей осуществления уверенного приёма сигналов в тяжёлых условиях связи - высоком уровне помех, значительных искажениях сигнала из-за флуктуаций параметров канала связи и т.д. Поэтому знание принципов построения кодированных сигналов, методов их формирования на передающей и декодирования на приёмной сторонах системы связи является необходимым и обязательным для современного инженера-связиста. Теоретическую основу помехоустойчивого кодирования составляет теорема К. Шеннона для канала с шумами, в которой утверждается, что для указанного канала можно найти такую систему оптимального кодирования, при которой сообщения будут переданы со сколь угодно большой степенью верности, если только производительность источника не превышает пропускной способности канала связи. Другой важный результат теории оптимального кодирования состоит в том, что принципиально сколь угодно малая вероятность неправильного декодирования может быть достигнута при использовании кодов, имеющих весьма длинные кодовые комбинации (кодовые слова).

2. МЕТОДИЧЕСКИЕ МАТЕРИАЛЫ ПО ПРОВЕДЕНИЮ ПРАКТИЧЕСКИХ ЗАНЯТИЙ

2.1 Практическое занятие №1 (2 часа).

Тема: «Понятие информации. Модели детерминированных и случайных сигналов. Преобразование непрерывных сигналов в дискретные»

2.1.1. Краткое описание проводимого занятия:

Задания для проведения текущего контроля успеваемости

1. Информацию делят на два вида ...

- Логическую и непрерывную

- Непрерывную и дискретную
- Непрерывную и логическую
- Дискретную и логическую

2. Теория информации рассматривается как существенная часть ...

- Электротехники
- Математики
- Кибернетики
- Физики

3. Математическая теория, посвященная измерению информации и каналам связи, называется ...

- Теорией вероятностей
- Кибернетикой
- Информатикой
- Теорией информации

2.2 Практическое занятие №2 (2 часа).

Тема: «Меры неопределенности дискретных множеств и непрерывных случайных величин. Количество информации как мера снятой неопределенности»

2.2.1. Краткое описание проводимого занятия:

Задания для проведения текущего контроля успеваемости

1. Количество информации при выборе одного из сообщений "выпала решка", "выпал орел" равно ...

- 1 байт
- 1 бит
- 2 бита
- 2 байта

2. Формула Шеннона ...

- $I = -(p_1 \log_2 p_1 - p_2 \log_2 p_2 - \dots - p_N \log_2 p_N)$
- $I = (p_1 \log_2 p_1 + p_2 \log_2 p_2 + \dots + p_N \log_2 p_N)$
- $I = (p_1 \log_2 p_1 - p_2 \log_2 p_2 - \dots - p_N \log_2 p_N)$
- $I = -(p_1 \log_2 p_1 + p_2 \log_2 p_2 + \dots + p_N \log_2 p_N)$

3. Формула Шеннона превращается в формулу Хартли если ...

- вероятности сообщений равны

- вероятности сообщений неравны
- количество сообщений равно 2
- количество сообщений равно 10

2.3 Практическое занятие №3 (2 часа).

Тема: «Оценка информационных характеристик источников сообщений»

2.3.1. Краткое описание проводимого занятия:

Задания для проведения текущего контроля успеваемости

1. Зачем кодируют информацию при передаче по каналам связи
 - чтобы более точно передать информацию
 - чтобы экономно передать информацию
 - чтобы более точно и экономно передать информацию
 - чтобы засекретить информацию
2. Чего не хватает в последовательности (источник, канал, приемник)
 - кодировщик и декодировщик
 - интегратор и дифференциатор
 - модем источника и модем приемника
 - модулятор и демодулятор
3. Чего не хватает в последовательности (источник, кодировщик, декодировщик, приемник)
 - шум
 - помехи
 - провод
 - канал

2.4 Практическое занятие №4 (2 часа).

Тема: «Информационные характеристики каналов связи»

2.4.1. Краткое описание проводимого занятия:

Задания для проведения текущего контроля успеваемости

1. Совокупность устройств, объединенных линиями связи, предназначенных для передачи информации от источника информации до ее приемника называется ...
 - запоминающим устройством
 - семантической информацией

- информационным каналом
- логическим устройством

2. Эффективность информационного канала характеризуется ...

- видом сигнала
- составом физической среды канала
- задержкой сигнала во времени
- скоростью и достоверностью передачи информации

3. Технические характеристики канала определяются ...

- принципом действия входящих в него
- надежностью работы
- задержкой сигнала во времени
- скоростью передачи информации

2.5 Практическое занятие №5 (2 часа).

Тема: «Эффективное кодирование. Введение в теорию помехоустойчивого кодирования»

2.5.1. Краткое описание проводимого занятия:

Задания для проведения текущего контроля успеваемости

1. Запись пути по дереву Хаффмана 0 и 1 означает ...

- вверх, вниз
- влево, вправо
- назад, вперед
- нет буквы, есть буква

2. Почему пример арифметического кодирования может показывать эффективность сжатия меньше энтропии

- В примере не учитывался код-маркер конца сообщения
- Арифметическое кодирование самое эффективное
- Рассматривался частный случай с хорошей сжимаемостью
- Нельзя сравнивать эффективность сжатия с энтропией

3. По какому принципу работает помехозащитное кодирование

- Уменьшает вероятность возникновения ошибки
- Уменьшает вероятность возникновения помех
- Уменьшает вероятность необнаруженной ошибки
- Улучшает пропускную способность канала

2.6 Практическое занятие №6 (2 часа).

Тема: «Построение групповых кодов. Циклические коды»

2.6.1. Краткое описание проводимого занятия:

Задания для проведения текущего контроля успеваемости

1. Совершенным называется групповой (m,n) код, исправляющий все ошибки веса ...
 - не большего k
 - не меньшего k и никаких других
 - не большего k и никаких других
 - не меньшего k
2. Квазисовершенным называется групповой (m,n) код, исправляющий все ошибки веса ...
 - не меньшего k , и некоторые ошибки веса $k + 1$
 - не большего k , и все ошибки веса $k + 1$
 - не меньшего k , и все ошибки веса $k + 1$
 - не большего k , и некоторые ошибки веса $k + 1$
3. Для любого целого положительного числа r существует совершенный (m,n) - код исправляющий одну ошибку, называемый ...
 - кодом Хаффмана
 - кодом Хэмминга
 - кодом Шеннона
 - кодом Фэнно

2.7 Практическое занятие №7 (2 часа).

Тема: «Матричные представления в теории кодирования. Кодирование линейными последовательными машинами»

2.7.1. Краткое описание проводимого занятия:

Задания для проведения текущего контроля успеваемости

1. Чем выгодно матричное кодирование
 - Проходит с большей скоростью
 - Требуется меньше объема памяти
 - Обнаруживает больше ошибок
 - Исправляет больше ошибок
2. Что такое криптография
 - Наука об изменении файла с целью сделать его непонятным для

непосвященных лиц

- Наука об изменении шифра с целью сделать его непонятным для непосвященных лиц
- Наука об изменении письма с целью сделать его непонятным для непосвященных лиц
- Наука об изменении кода с целью сделать его непонятным для непосвященных лиц

3. В шифре простой замены ...

- каждый знак письма меняется на число по выбранному правилу
- каждый знак письма меняется на двоичное число по выбранному правилу
- каждый знак письма меняется на знак другого языка по выбранному правилу
- каждый знак письма меняется на другой знак по выбранному правилу