Федеральное агентство по рыболовству
БГАРФ ФГБОУ ВО «КГТУ»
Калининградский морской рыбопромышленный колледж

Утверждаю
Заместитель начальника колледжа
по учебно-методической работе
А.И.Колесниченко

Подписано цифровой подписью: Колесниченко Андрей Игоревич
Причина: Я утвердил этот документ
Дата: 18 Март 2025 г. 20:02:48
DN: E=uc_fk@roskazna.ru, S=77 Москва, ИНН ЮЛ=7710568760, ОГРН=1047797019830, STREET="Большой Златоустинский переулок, д. 6, строение 1", L=г. Москва, C=RU, O=Казначейство России, CN=Федеральное казначейство

ОПд.14 ОСНОВЫ ТЕОРИИ ИНФОРМАЦИИ

Методические указания для выполнения практических занятий
по специальности
09.02.06 Сетевое и системное администрирование

МО-09 02 06-ОП.14.ПЗ

РАЗРАБОТЧИКИ	Богатырева Т.Н.
ЗАВЕДУЮЩИЙ ОТДЕЛЕНИЕМ	Кругленя В.Ю.
ГОД РАЗРАБОТКИ	2024
ГОД ОБНОВЛЕНИЯ	2025

Содержание

Практическая работа №1. Способы хранения обработки и передачи информации.

Исходные данные: раздаточный материал

Содержание и порядок выполнения задания:

Задачи к теме "Единицы измерения информации: биты, байты, килобайты."

Основные равенства, которые необходимо знать и помнить при решении задач на исчисление вероятности:

N - исходная совокупность равновероятных событий;

P - вероятность одного равновероятного события из исходной совокупности **N**;

i - степень, в которую нужно возвести константу 2, чтобы получить **N**, **i** – показатель количества информации в битах

$$N=1/P$$

$$2^i=N$$

$$2^i=1/P$$

$$P=1/N$$

Задача 1. Сколько битов информации содержится в сообщении размером 8 байтов?

Задача 2. Сообщение, записанное буквами из 64-символьного алфавита, содержит 20 символов. Какой объём информации оно несёт?

Задача 3. Сколько символов содержит сообщение, записанное с помощью 16-символьного алфавита, если его объём составил 1/16 часть мегабайта?

Задача 4. Сколько байтов информации содержится в сообщении размером четверть мегабайта?

Задача 5. Объём сообщения, содержащего 2048 символов, составил 1/512 часть мегабайта. Какова мощность алфавита, с помощью которого записано сообщение?

Задача 6. Текст занимает 1/4 килобайта памяти компьютера в кодировке КОИ-8 (однобайтной). Сколько символов содержит этот текст?

Задача 7. Для хранения текста требуется 84000 бит. Сколько страниц займёт этот текст, если на странице размещается 30 строк по 70 символов в строке?

Задача 8. В корзине лежат шары. Все разного цвета. Сообщение о том, что достали синий шар, несёт 5 бит информации. Сколько всего шаров было в корзине?

Задача 9. Алфавит племени Мульти состоит из 8 букв. Какой объём информации несёт любая буква этого алфавита?

Практическая работа №2. Измерение количества информации.

Исходные данные: раздаточный материал

Содержание и порядок выполнения задания:

Задача 10. В корзине лежат 16 шаров. Все шары разного цвета. Сколько информации несёт сообщение о том, что достали красный шар?

Задача 11. У племени Мульти 32-символьный алфавит, племя Пульти пользуется 64-символьным алфавитом. Вожди племён обменялись письмами. Письмо племени Мульти содержало 80 символов, а письмо племени Пульти - 70 символов.

Сравнить объём информации, содержащийся в письмах.

Задача 12 В корзине лежат шары (белые и чёрные). Среди них - 4 белых. Сообщение о том, что достали белый шар, несёт 3 бита информации. Сколько всего шаров было в корзине?

Дано: $i_{бел}$=3 бита; $k_{бел}$=4 шара; $N_{чёрных+белых}$=?

Задача 13. В ящике лежат перчатки (белые и чёрные). Среди них – $k_{чёрн.}$=2 пары чёрных. Сообщение о том, что из ящика достали одну пару чёрных перчаток, несёт $i_{чёрн.}$=4 бита информации.

Сколько всего было пар перчаток (чёрных и белых) в ящике?

Задача 14. В ящике лежат 8 чёрных шаров и 24 белых. Сколько информации несёт сообщение о том, что достали чёрный шар?

Дано: $k_{чёрн}$=8; $k_{бел}$=24. Найти $i_{чёрн}$

Задача 15. В мешке лежат 64 монеты. Сообщение о том, что достали золотую монету, несёт 4 бита информации. Сколько золотых монет было в мешке?

Дано: N=64; $i_{зол}$=4. Найти: $k_{зол}$.

Задача 16. На остановке останавливаются автобусы с разными номерами. Сообщение о том, что к остановке подошёл автобус маршрута №1, несёт 4 бита информации. Вероятность появления на остановке автобуса маршрута №2 $P_{№2}$ в два раза меньше, чем вероятность появления автобуса маршрута №1 $P_{№1}$. Сколько бит информации несёт сообщение о появлении автобуса маршрута №2 на остановке?

Дано: $i_{№1}$=4 бита; $P_{№2}=P_{№1}/2$.

Задача 17

На остановке останавливаются автобусы с разными номерами. Сообщение о том, что к остановке подошёл автобус маршрута № 1, несёт 4 бита информации. Вероятность появления на остановке автобуса маршрута №2 в два раза больше, чем вероятность появления автобуса маршрута №1. Сколько бит информации несёт сообщение о появлении автобуса маршрута №2 на остановке?

Задача 18. Известно, что в ящике лежит 64 шара. Из них чёрных 16, белых 16, жёлтых 2, красных 4. Какое количество информации несут сообщение о том, что из ящика случайным образом достали чёрный шар, белый шар, жёлтый шар, красный шар?

Практическая работа №3. Определение пропускной способности канала.

Исходные данные: раздаточный материал

Содержание и порядок выполнения задания:

Тема: «Передача информации. Скорость передачи информации»

1. Повторение ранее изученного материала

1. Понятие информации

Информация – в общем случае, совокупность сведений о каких-либо событиях, явлениях, предметах, получаемых в результате взаимодействия с внешней средой. Формой представления информации является сообщение.

2. Виды и свойства информации

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

1. графическая;
2. звуковая;
3. текстовая;
4. числовая;

3. Единицы измерения количества информации

- 1 байт = 8 бит,
- 1 килобайт = 1024 байт,
- 1 мегабайт = 1024 Кбайт,
- 1 гигабайт = 1024 Мбайт,
- 1 терабайт = 1024 Гбайт,
- 1 петабайт = 1024 Тбайт.

2. Введение нового материала

Все виды информации кодируются в последовательности электрических импульсов: есть импульс (1), нет импульса (0), то есть в последовательности нулей и единиц. Такое кодирование информации в компьютере называется двоичным кодированием. Соответственно раз если можно данные импульсы сохранять и обрабатывать при помощи компьютерных устройств, значит их можно и передавать.

Для передачи информации необходимы:

Источник информации – система из которой информация передаётся.

Канал передачи информации – способ при помощи которого осуществляется передача информации.

Приемник информации – система, которая осуществляет получение необходимой информации.

Преобразование информации в сигналы, удобные для прохождения по линии связи, осуществляется передатчиком.

В процессе преобразования информации в сигнал происходит её кодирование. В широком смысле кодированием называется преобразование информации в сигнал. В узком смысле кодирование – это преобразование информации в сочетание определенных символов. В нашем случае в последовательность 1 и 0.

На приемной стороне осуществляется обратная операция декодирования, т.е. восстановление по принятому сигналу переданной информации.

Декодирующее устройство, (декодер) преобразует принятый сигнал к виду удобному для восприятия получателем.

Одними из самых важных свойств передачи информации являются скорость передачи информации и пропускная способность канала.

Скорость передачи данных - скорость, с которой передается или принимается информация в двоичной форме. Обычно скорость передачи данных измеряется количеством бит, переданных в одну секунду.

Минимальная единица измерения скорости передачи информации – 1 бит в секунду (1 бит/сек)

Пропускная способность канала связи - максимальная скорость передачи данных от источника к получателю.

Обе величины измеряются в бит/сек, что часто путают с Байт/сек и обращаются к поставщикам (провайдерам) услуг связи в связи с ухудшением скорости или несоответствии скорости передачи информации.

3. Решение задач

Решение задач на скорость передачи информации практически полностью совпадает с решением задач на скорость, время и расстояние.

S – размер передаваемой информации

V – скорость передачи информации

T – время передачи информации

Поэтому формулы: $S = V * t; V = \frac{S}{t}; t = \frac{S}{V}$ справедливы при решении задач на скорость передачи информации. Однако следует помнить, что все величины измерения должны совпадать. (если скорость в Кбайт/сек, то время в секундах, а размер в Килобайтах)

Рассмотрим пример задачи:

Сколько секунд потребуется модему, передающему сообщение со скоростью 28800 бит/сек, чтобы передать цветное изображение размером 640*480 пикселей, при условии, что цвет каждого пикселя кодируется 3 байтами.

Решение:

1. Определим количество пикселей в изображении:

640*480= 307200 пикселей

2. Т.к. каждый пиксель кодируется 3 байтами, определим информационный объем изображения:

307200 * 3 = 921600 байт

3. Заметим, что скорость передачи информации измеряется в бит/сек, а информационный вес изображения в байтах. Переведем скорость в байт/сек, для удобства подсчета:

28800: 8 = 3600 байт/сек

4. Определяем время передачи сообщения, если скорость равна 3600 байт/сек:

921600: 3600 = 256 сек

Ответ: 256 секунд потребуется

Задачи:

1. Известно, что длительность непрерывного подключения к сети Интернет с помощью модема для некоторых АТС не превышает 10 мин. Определите максимальный размер файла (Кбайт), который может быть передан за время такого подключения, если модем передает информацию в среднем со скоростью 32 Кбит/сек.

2. Скорость передачи данных через ADSL-соединение равна 64000 бит/сек. Через данное соединение передают файл размером 375 Кбайт. Определите время передачи файла в секундах.

3. Сколько секунд потребуется модему, передающему сообщение со скоростью 28800 бит/сек, чтобы передать 100 страниц текста в 30 сток по 60 символов каждая, при условии, что каждый символ кодируется одним байтом.

4. Скорость передачи данных через модемное соединение равна 56 Кбит/сек. Передача текстового файла через это соединение заняла 12 сек. Определите, сколько символов содержал переданный текст, если известно, что он был представлен в кодировке UNICODE.

5. Модем передает данные со скоростью 56 Кбит/сек. Передача текстового файла заняла 4,5 минуты. Определите, сколько страниц содержал переданный текст, если известно, что он был представлен в кодировке Unicode, а на одной странице – 3072 символа.

6. Средняя скорость передачи данных с помощью модема равна 36 Кбит/сек. Сколько секунд потребуется модему, чтобы передать 4 страницы текста в кодировке КОИ8, если считать, что на каждой странице в среднем 2 304 символа?

7. Разведчик Белов должен передать сообщение: «Место встречи изменить нельзя. Юстас.» пеленгатор определяет место передачи, если она длиться не менее 2 минут. С какой скоростью (бит/сек) должен передавать радиограмму разведчик?

Практическая работа №4. Интерполяционная формула Уиттекера-Шеннона, частота Найквиста

Исходные данные: раздаточный материал

Содержание и порядок выполнения задания:

Расчет вероятностей. Составление закона распределения вероятностей

Теоретические основы

Дискретной называют случайную величину, возможные значения которой есть отдельные изолированные числа, которые эта величина принимает с определенными ненулевыми вероятностями. Число возможных значений может быть конечным или бесконечным (счетным).

Законом распределения дискретной случайной величины называют перечень её возможных значений и соответствующих им вероятностей. Закон распределения может быть задан одним из следующих способов.

1. Таблицей

где

x	x_1	x_2	…	x_n
p	p_1	p_2	…	p_n

$$\sum_{i=1}^{n} p_i = 1.$$

2. Аналитически $P(X = x_i) = \varphi(x_i)$. Например:

а) *биномиальное распределение* (Формула Бернулли)

$P(X = k) = C_n^k p^k q^{n-k}$, 0<p<1, k=0, 1, 2, …, n;

б) распределение *Пуассона*

$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!},\ \lambda>0,\ k=0, 1, 2, \ldots .$$

3. С помощью *функции распределения F(x)*, определяющей для каждого значения *x* вероятность того, что случайная величина X примет значение, меньшее *x*, т. е. $F(x) = P(X < x)$.

Свойства F(x):

1) $0 \le F(x) \le 1$;

2) $F(x_2) \ge F(x_1)$, если $x_2 > x_1$;

3) $\lim\limits_{x \to -\infty} F(x) = 0,\ \lim\limits_{x \to +\infty} F(x) = 1.$

4. Закон распределения может быть задан графически - *многоугольником распределения (см. пример 1).*

Числовые характеристики дискретных случайных величин

Математическое ожидание $M(X)=\sum_i x_i p_i$;

Дисперсия $D(X)=M\left[X-M(X)\right]^2$ или $D(X)=M(X^2)-\left[M(X)\right]^2$;

Среднее квадратическое отклонение $\sigma(X)=\sqrt{D(X)}$.

Для биномиального распределения *M(X)=np*, *D(X)=npq*. Для распределения Пуассона *M(X)=λ, D(X)= λ.*

Пример 1.

Устройство состоит из трех независимо работающих элементов. Вероятность отказа каждого элемента в одном опыте равна 0,1. Составить закон распределения числа отказавших элементов в одном опыте, построить многоугольник распределения. Найти функцию распределения F(x) и построить её график. Найти M(X), D(X), σ(X).

Решение: Дискретная случайная величина X (число отказавших элементов в одном опыте) имеет следующие возможные значения: $x_1=0$ (ни один из элементов устройства не отказал), $x_2=1$ (отказал один элемент), $x_3=2$ (отказало два элемента) и $x_4=3$ (отказали три элемента).

Отказы элементов независимы один от другого, вероятности отказа каждого элемента равны между собой, поэтому применима формула Бернулли. Учитывая, что, по условию, n=3, p=0,1 (следовательно, q=1–0,1=0,9), получим: $P_3(0)=q^3=0{,}9^3=0{,}729$;

$P_3(1)=C_3^1 pq^2=3\cdot 0{,}1\cdot 0{,}9^2=0{,}243$;

$P_3(2)=C_3^2 p^2 q=3\cdot 0{,}1^2\cdot 0{,}9=0{,}027$; $P_3(3)=p^3=0{,}1^3=0{,}001$.

Контроль: $\sum_{i=1}^{4} p_i=1$; 0,729+0,243+0,027+0,001=1.

Искомый биномиальный закон распределения *X*:

X	0	1	2	3
p	0,729	0,243	0,027	0,001

Для построения многоугольника распределения строим прямоугольную систему координат. По оси абсцисс откладываем возможные значения x_i, а по оси ординат – соответствующие им вероятности p_i. Построим точки $M_1(0;0,729)$, $M_2(1;0,243)$, $M_3(2;0,027)$, $M_4(3;0,001)$. Соединив эти точки отрезками прямых, получаем искомый многоугольник распределения (Рис.1).

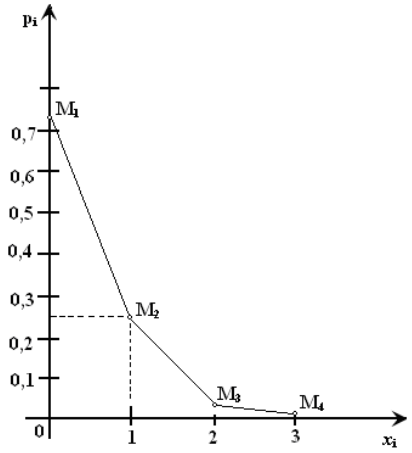


Рис.1

Найдем функцию распределения $F(x)=P(X<x)$.

Для $x \le 0$ имеем $F(x)=P(X<0)=0$;

для $0 < x \le 1$ имеем $F(x)=P(X<1)=P(X=0)=0{,}729$;

для $1 < x \le 2$ $F(x)=P(X<2)=P(X=0)+P(X=1)=0{,}729+0{,}243=0{,}972$;

для $2 < x \le 3$ $F(x)=P(X<3)=P(X=0)+P(X=1)+P(X=2)=0{,}972+0{,}027=0{,}999$;

для $x>3$ будет $F(x)=1$, т. к. событие достоверно.

$$F(x)=\begin{cases} 0 & при \ \ x \le 0, \\ 0{,}729 & при \ 0 < x \le 1, \\ 0{,}972 & при \ 1 < x \le 2, . \\ 0.999 & при \ 2 < x \le 3, \\ 1 & при \ x > 3. \end{cases}$$

График этой функции приведен на Рис. 2.

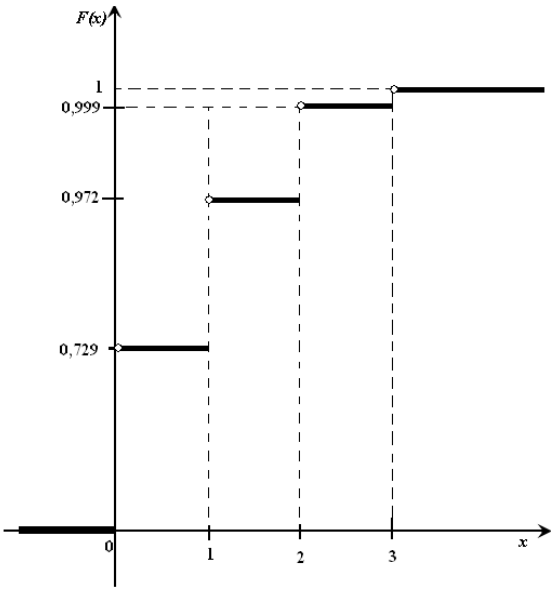


Рис. 2

Для биномиального распределения $M(X)=np=3\cdot0,1=0,3$; $D(X)=npq=3\cdot0,1\cdot0,9=0,27$; $\sigma(X)=\sqrt{D(X)}=\sqrt{0,27}\approx0,52$.

Пример 2.

В партии из 10 деталей имеется 8 стандартных. Наудачу отобраны две детали. Составить закон распределения случайной величины X – числа стандартных деталей среди отобранных. Найти $M(X)$, $D(X)$.

Решение: Случайная величина X – число стандартных деталей среди отобранных деталей – имеет следующие возможные значения: $x_1=0$; $x_2=1$; $x_3=2$. Найдем вероятности возможных значений X по формуле (пример 2) $P(X=k)=\frac{C_n^k\cdot C_{N-n}^{m-k}}{C_N^m}$ (N – число деталей в партии, n – число стандартных деталей в партии, m – число отобранных деталей, k – число стандартных деталей среди отобранных), находим: $P(X=0)=\frac{C_8^0\cdot C_2^2}{C_{10}^2}=\frac{1}{10\cdot9/(1\cdot2)}=\frac{1}{45}$;

$$P(X=1)=\frac{C_8^1\cdot C_2^1}{C_{10}^2}=\frac{8\cdot2}{45}=\frac{16}{45};$$

$$P(X=2)=\frac{C_8^2\cdot C_2^0}{C_{10}^2}=\frac{8\cdot7/(1\cdot2)}{45}=\frac{28}{45}.$$

Составим искомый закон распределения:

X	0	1	2
p	$\frac{1}{45}$	$\frac{16}{45}$	$\frac{28}{45}$

Контроль: $\frac{1}{45}+\frac{16}{45}+\frac{28}{45}=1$.

$$M(X)=\sum_{i=1}^{3}x_i p_i=0\cdot\frac{1}{45}+1\cdot\frac{16}{45}+2\cdot\frac{28}{45}=\frac{72}{45}=\frac{8}{5}.$$

$$D(X)=M(X^2)-[M(X)]^2;\ M(X^2)=\sum_{i=1}^{3}x_i^{\,2}p_i=0^2\cdot\frac{1}{45}+1\ \cdot\frac{16}{45}+2^2\cdot\frac{28}{45}=\frac{128}{45};$$

$$D(X)=\frac{128}{45}-\left(\frac{8}{5}\right)^2=\frac{128}{45}-\frac{64}{25}=\frac{64}{225}.$$

Практическая работа №5. Применение теоремы отчетов

Цель занятия: знакомство с основными понятия и принципами создания библиографического списка.

Исходные данные: Раздаточный материал

Содержание и порядок выполнения задания:

Теоретические основы

Теорема Котельникова

В 1933 году В.А. Котельниковым доказана теорема отсчетов [6, 32], имеющая важное значение в теории связи: непрерывный сигнал $s(t)$ с ограниченным спектром можно точно восстановить (интерполировать) по его отсчетам $s(k\Delta t)$, взятым через интервалы $\Delta t=\frac{1}{(2F)}$, где F – верхняя частота спектра сигнала.

В соответствии с этой теоремой сигнал $s(t)$ можно представить рядом Котельникова

$$s(t)=\sum_{k=-\infty}^{\infty}s\left(\frac{k}{2F}\right)\frac{\sin 2\pi F\left[t-\frac{k}{2F}\right]}{2\pi F\left[t-\frac{k}{2F}\right]}. \qquad (1.21)$$

Таким образом, сигнал $s(t)$, можно абсолютно точно представить с помощью последовательности отсчетов $s\left(\frac{k}{2F}\right)$, заданных в дискретных точках $\frac{k}{2F}$ (рис.1.16).

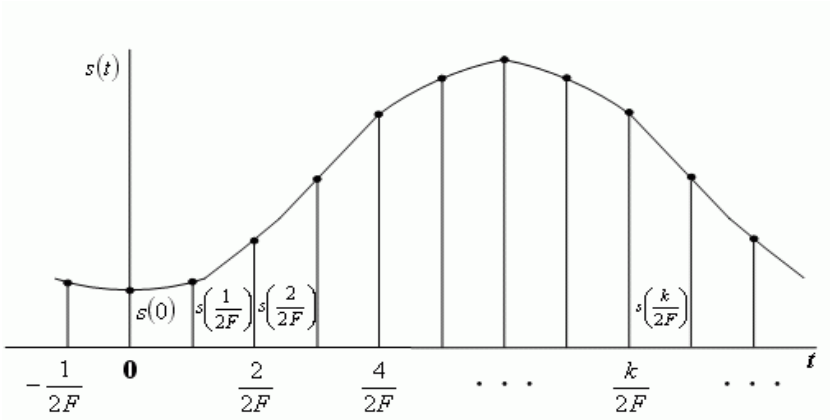


Рис. 1.16. Сигнал и его отсчеты

Функции

$$\psi(t)=\frac{\sin 2\pi F\left[t-\frac{k}{2F}\right]}{2\pi F\left[t-\frac{k}{2F}\right]} \qquad (1.22)$$

образуют ортогональный базис в пространстве сигналов, характеризующихся ограниченным спектром:

$$\Phi(f)=0 \text{ при } |f|>F. \qquad (1.23)$$

Обычно для реальных сигналов можно указать диапазон частот, в пределах которого сосредоточена основная часть его энергии и которым определяется ширина спектра сигнала. В ряде случаев спектр сознательно сокращают. Это обусловлено тем, что аппаратура и линия связи должны иметь минимальную полосу частот. Сокращение спектра выполняют, исходя из допустимых искажений сигнала. Например, при телефонной связи хорошая разборчивость речи и узнаваемость абонента обеспечиваются при передаче сигналов в полосе частот $\Delta F = 0{,}3\ldots3{,}4\,[\text{кГц}]$. Увеличение ΔF приводит к неоправданному усложнению аппаратуры и повышению затрат. Для передачи телевизионного изображения при стандарте в 625 строк полоса частот, занимаемая сигналом, составляет около 6 МГц.

Из вышесказанного следует, что процессы с ограниченными спектрами могут служить адекватными математическими моделями многих реальных сигналов.

Функция вида $\frac{\sin 2\pi F\left[t-\frac{k}{2F}\right]}{2\pi F\left[t-\frac{k}{2F}\right]}$ называется функцией отсчетов (рис.1.17).

Рис. 1.17. Функция отсчётов

Она характеризуется следующими свойствами. Если $k=0$, функция отсчетов имеет максимальное значение при $t=0$, а в моменты времени $t=\frac{i}{2F}$ ($i=1,2,\dots$) она обращается в нуль; ширина главного лепестка функции отсчетов на нулевом уровне равна $\frac{1}{F}$, поэтому минимальная длительность импульса, который может существовать на выходе линейной системы с полосой пропускания F, равна $\frac{1}{F}$; функции отсчетов ортогональны на бесконечном интервале времени.

На основании теоремы Котельникова может быть предложен следующий способ дискретной передачи непрерывных сигналов:

Для передачи непрерывного сигнала $s(t)$ по каналу связи с полосой пропускания F определим мгновенные значения сигнала $s(t)$ в дискретные моменты времени $t_k=\frac{1}{2F}$, ($k=0,1,2,\dots$). После этого передадим эти значения по каналу связи каким - либо из возможных способов и восстановим на приемной стороне переданные отсчеты. Для преобразования потока импульсных отсчетов в непрерывную функцию пропустим их через идеальный ФНЧ с граничной частотой F.

Можно показать, что энергия сигнала находится по формуле [6, 32]:

$$E=\int_{-\infty}^{\infty}s^2(t)dt=\frac{1}{2F}\sum_{k=-\infty}^{\infty}s^2\left(\frac{k}{2F}\right). \qquad (1.24)$$

Для сигнала, ограниченного во времени, выражение (1.24) преобразуется к виду:

$$E=\int_{1}^{2FT}s^2(t)dt=\frac{1}{2F}\sum_{k=1}^{2FT}s^2\left(\frac{k}{2F}\right). \qquad (1.25)$$

Выражение (1.25) широко применяется в теории помехоустойчивого приема

сигналов, но является приближенным, т.к. сигналы не могут быть одновременно ограничены по частоте и времени.

Практическое задание

1. Изобразить сигналы, синтезируемые в лабораторной работе:

а) синусоидальный сигнал частотой 5кГц;

б) видеоимпульсы прямоугольной формы длительностью 0,25; 0,5; 1,0 мс;

в) видеоимпульсы пилообразной формы длительностью 0,5 мс; 1,0 мс.

2. Рассчитать и построить идеальные выборочные сигналы для сигналов, указанных в п. 1а, 1б, 1в, при $f_{выб}$=5, 10, 20, 40 кГц.

Отчет

Отчет должен быть оформлен в текстовом редакторе и содержать:

— наименование работы;

— цель работы;

— задание;

— последовательность выполнения работы;

— ответы на контрольные вопросы;

— вывод о проделанной работе.

Контрольные вопросы

1. Сформулируйте теорему Котельникова для сигналов с ограниченным спектром.

2. Объясните погрешности синтезирования реальных сигналов по дискретным отсчетам.

Практическая работа №6. Поиск энтропии случайных величин

Исходные данные: Раздаточный материал

Содержание и порядок выполнения задания:

Вероятностный подход к измерению дискретной и непрерывной информации

Количество информации по Хартли и Шеннону

Понятие количество информации отождествляется с понятием информация. Эти два понятия являются синонимами. Мера информации должна монотонно возрастать с увеличением длительности сообщения (сигнала), которую естественно измерять числом символов в дискретном сообщении и временем передачи в непрерывном случае. Кроме того, на содержание количества информации должны влиять и статистические характеристики, так как сигнал должен рассматриваться как случайный процесс.

При этом наложено ряд ограничений:

1. Рассматриваются только дискретные сообщения.

2. Множество различных сообщений конечно.

3. Символы, составляющие сообщения равновероятны и независимы.

Хартли впервые предложил в качестве меры количества информации принять логарифм числа возможных последовательностей символов.

$I=\log m^k=\log N$ (1)

К.Шеннон попытался снять те ограничения, которые наложил Хартли. На самом деле в рассмотренном выше случае равной вероятности и независимости символов при любом k все возможные сообщения оказываются также равновероятными, вероятность каждого из таких сообщений равна P=1/N. Тогда количество информации можно выразить через вероятности появления сообщений $I=-\log P$.

В силу статистической независимости символов, вероятность сообщения длиной в k символов равна

$$P=\prod_{i=1}^{k} p_i$$

Если i–й символ повторяется в данном сообщении k_i раз, то $P=\prod_{i=1}^{m} p_i^{k_i}$

так как при повторении i символа k_i раз k уменьшается до m. Из теории вероятностей известно, что, при достаточно длинных сообщениях (большое число

символов k) $k_i \approx k \cdot p_i$ и тогда вероятность сообщений будет равняться $P = \prod_{i=1}^{m} p_i^{kp_i}$

Тогда окончательно получим

$$I = -\log P = -k\sum_{i=1}^{m} p_i \log p_i \qquad (2)$$

Данное выражение называется формулой Шеннона для определения количества информации.

Формула Шеннона для количества информации на отдельный символ сообщения совпадает с энтропией. Тогда количество информации сообщения, состоящего из k символов будет равняться $I = k \cdot H$

Количество информации, как мера снятой неопределенности

При передаче сообщений, о какой-либо системе происходит уменьшение неопределенности. Если о системе все известно, то нет смысла посылать сообщение. Количество информации измеряют уменьшением энтропии.

Количество информации, приобретаемое при полном выяснении состояния некоторой физической системы, равно энтропии этой системы:

$$I = -\sum_{i=1}^{n} p_i \log p_i$$

Количество информации I – есть осредненное значение логарифма вероятности состояния. Тогда каждое отдельное слагаемое $-\log p_i$ необходимо рассматривать как частную информацию, получаемую от отдельного сообщения, то есть

$$I_i = -\log p_i$$

Избыточность информации

Если бы сообщения передавались с помощью равновероятных букв алфавита и между собой статистически независимых, то энтропия таких сообщений была бы максимальной. На самом деле реальные сообщения строятся из не равновероятных букв алфавита с наличием статистических связей между буквами. Поэтому энтропия реальных сообщений $-H_p$, оказывается много меньше оптимальных сообщений – H_o.

Допустим, нужно передать сообщение, содержащее количество информации, равное I. Источнику, обладающему энтропией на букву, равной H_p, придется затратить некоторое число n_p, то есть

$$I = n_p H_p$$

Если энтропия источника была бы H_0, то пришлось бы затратить меньше букв на передачу этого же количества информации

$$n_o = \frac{I}{H_o} < n_p$$

$I = n_0 H_0$

Таким образом, часть букв $n_p - n_o$ являются как бы лишними, избыточными. Мера удлинения реальных сообщений по сравнению с оптимально закодированными и представляет собой избыточность D.

$$D = 1 - \frac{H_p}{H_o} = 1 - \frac{n_o}{n_p} = \frac{n_p - n_o}{n_p} \qquad (3)$$

Но наличие избыточности нельзя рассматривать как признак несовершенства источника сообщений. Наличие избыточности способствует повышению помехоустойчивости сообщений. Высокая избыточность естественных языков обеспечивает надежное общение между людьми.

Частотные характеристики текстовых сообщений

Важными характеристиками текста являются повторяемость букв, пар букв (биграмм) и вообще m-ок (m-грамм), сочетаемость букв друг с другом, чередование гласных и согласных, и некоторые другие. Замечательно, что эти характеристики являются достаточно устойчивыми.

Идея состоит в подсчете чисел вхождений каждой n^m возможных m-грамм в достаточно длинных открытых текстах $T=t_1t_2\ldots t_l$, составленных из букв алфавита $\{a_1, a_2, \ldots, a_n\}$. При этом просматриваются подряд идущие m-граммы текста

$t_1t_2\ldots t_m$, $t_2t_3\ldots t_{m+1}$, ..., $t_{i-m+1}t_{l-m+2}\ldots t_l$.

Если $\vartheta(a_{i1}a_{i2}\ldots a_{im})$ – число появлений m-граммы $a_{i1}a_{i2}\ldots a_{im}$ в тексте T, а L общее число подсчитанных m-грамм, то опыт показывает, что при достаточно больших L частоты

$\frac{\vartheta(a_{i1}a_{i2}\ldots a_{im})}{L}$ для данной m-граммы мало отличаются друг от друга.

В силу этого, относительную частоту считают приближением вероятности P ($a_{i1}a_{i2}...a_{im}$) появления данной m-граммы в случайно выбранном месте текста (такой подход принят при статистическом определении вероятности).

Для русского языка частоты (в порядке убывания) знаков алфавита, в котором отождествлены Е с Ё, Ь с Ъ, а также имеется знак пробела (-) между словами, приведены в таблице 1.

Таблица 1

- 0.175	О 0.090	Е, Ё 0.072	А 0.062
И 0.062	Т 0.053	Н 0.053	С 0.045
Р 0.040	В 0.038	Л 0.035	К 0.028
М 0.026	Д 0.025	П 0.023	У 0.021
Я 0.018	Ы 0.016	З 0.016	Ь, Ъ 0.014
Б 0.014	Г 0.013	Ч 0.012	Й 0.010
Х 0.009	Ж 0.007	Ю 0.006	Ш 0.006
Ц 0.004	Щ 0.003	Э 0.003	Ф 0.002

Некоторая разница значений частот в приводимых в различных источниках таблицах объясняется тем, что частоты существенно зависят не только от длины текста, но и от его характера.

Устойчивыми являются также частотные характеристики биграмм, триграмм и четырехграмм осмысленных текстов.

Порядок выполнения работы

Определить количество информации (по Хартли), содержащееся в заданном сообщении, при условии, что значениями являются буквы кириллицы.

«Фамилия Имя Отчество» завершил ежегодный съезд эрудированных школьников, мечтающих глубоко проникнуть в тайны физических явлений и химических реакций

Построить таблицу распределения частот символов, характерные для заданного сообщения. Производится так называемая частотная селекция, текст сообщения анализируется как поток символов и высчитывается частота встречаемости каждого символа. Сравнить с имеющимися данными в табл. 1.

На основании полученных данных определить среднее и полное количество информации, содержащееся в заданном сообщении. Оценить избыточность сообщения.

1. Построить таблицу распределения частот символов, характерных для заданного сообщения путём деления количества определённого символа в данном сообщении на общее число символов

По формуле

$$\sum_{i=1}^{m} p_i \log p_i$$

H= вычислил энтропию сообщения

2. Далее по формуле Шеннона для определения кол-ва информации

$$I = -\log P = -k\sum_{i=1}^{m} p_i \log p_i$$ вычислил кол-во информации в передаваемом сообщении

3. Вычислил избыточность D по формуле

$$D = 1 - \frac{H_p}{H_o} = 1 - \frac{n_o}{n_p} = \frac{n_p - n_o}{n_p}$$

Отчет

Отчет должен быть оформлен в текстовом редакторе и содержать:

— наименование работы;

— цель работы;

— задание;

— последовательность выполнения работы;

— ответы на контрольные вопросы;

— вывод о проделанной работе.

Контрольные вопросы

1. Дать определение понятие энтропия.
2. Что означает вероятностный способ измерения информации?
3. Что означает статическое определение вероятности?
4. Запишите уравнение Хартли.
5. Какие основные разработки внес в основу теории информации Шеннон?

Практическая работа №7. Энтропийное кодирование. Дифференциальная энтропия.

Исходные данные: Раздаточный материал

Содержание и порядок выполнения задания:

Пример 3.

В устройстве независимо друг от друга выходят из строя три элемента. Вероятность выхода из строя первого элемента – 0,3, второго – 0,2, третьего – 0,4. Составить закон распределения случайной величины X – числа вышедших из строя элементов.

Решение: случайная величина X имеет следующие возможные значения: $x_1=0$, $x_2=1$, $x_3=2$, $x_4=3$. $p_1=0,3$, $q_1=1-p_1=0,7$, $p_2=0,2$, $q_2=1-p_2=0,8$, $p_3=0,4$, $q_3=1-p_3=0,6$.

P(X=k) вычисляем по следующим формулам (см. пример 4)

$$P(X=0)=q_1\cdot q_2\cdot q_3=0,7\cdot 0,8\cdot 0,9=0,336;$$

$$P(X=1)=p_1\cdot q_2\cdot q_3+q_1\cdot p_2\cdot q_3+q_1\cdot q_2\cdot p_3=0,3\cdot 0,8\cdot 0,6+0,7\cdot 0,2\cdot 0,6+$$
$$+0,7\cdot 0,8\cdot 0,4=0,144+0,084+0,224=0,452;$$

$$P(X=2)=p_1\cdot p_2\cdot q_3+p_1\cdot q_2\cdot p_3+q_1\cdot p_2\cdot p_3=0,3\cdot 0,2\cdot 0,6+0,3\cdot 0,8\cdot 0,4+$$
$$+0,7\cdot 0,2\cdot 0,4=0,118;$$

$$P(X=0)=p_1\cdot p_2\cdot p_3=0,3\cdot 0,2\cdot 0,4=0,024.$$

Контроль: 0,336+0,452+0,118+0,024=1.

Искомый закон распределения:

X	0	1	2	3
p	0,336	0,452	0,118	0,024

Пример 4.

Среднее число заказов такси, поступающих на диспетчерский пункт в одну минуту, равно двум. Составить закон распределения случайной величины X – числа заказов, поступающих за 4 минуты. Найти M(X), D(X).

Решение: Поток заказов на такси можно считать *простейшим,* т. е. обладающим стационарностью, «отсутствием последствия» и ординарностью. *Интенсивность потока* (среднее число событий появляющихся в единицу времени) $\lambda=2$. Вероятность появления k событий простейшего потока за время t=4 определяется формулой Пуассона $P_t(k)=\frac{(\lambda t)^k\ e^{-\lambda t}}{k!}$, для данной задачи $P_4(k)=\frac{8^k\ e^{-8t}}{k!}$. Совокупность возможных значений X есть счетное множество, т.е. $x_1=0,\ x_2=1,\ \ldots,\ x_k=k+1,\ \ldots;$ тогда закон распределения случайной величины X – числа заказов, поступающих за 4 минуты принимает вид:

X	0	1	2	…	k	…
p	$e^{-\lambda t}$	$\frac{(\lambda t)\ e^{-\lambda t}}{1!}$	$\frac{(\lambda t)^2\ e^{-\lambda t}}{2!}$	…	$\frac{(\lambda t)^k\ e^{-\lambda t}}{k!}$	…

или

X	0	1	2	…	k	…
p	e^{-8}	$\frac{8\ e^{8}}{1!}$	$\frac{8^2\ e^{-8}}{2!}$	…	$\frac{8^k\ e^{-8}}{k!}$	…

Воспользовавшись таблицей 3 приложения, окончательно получим:

X	0	1	2	…	k	…
p	0,00035	0,002684	0,010735	…	$\frac{8^k\ e^{-8}}{k!}$	…

Наивероятнейшее число заказов такси за 4 минуты можно определить по получившемуся закону распределения (значения *x*, при которых *p* максимально): $k'_0=7$, $k''_0=8$. Для простейшего потока событий: математическое ожидание $M(X)=\lambda t=8$, дисперсия $D(X)=\lambda t=8$.

Пример 5.

Даны законы распределения независимых случайных величин *X* и *Y*. Составить закон распределения случайной величины *Z=X+2Y*. Найти *M(Z), D(Z).*

X	-3	0	1
p	0,1	0,03	0,06

Y	1	3	6
p	0,2	0,5	0,3

Решение: Закон распределения *V=2Y* получается из распределения *Y* путем умножения всех значений y_i на 2. Получаем:

V	2	6	12
p	0,2	0,5	0,3

.

Для составления закона распределения случайной величины *Z* вычислим все ее возможные значения по формуле $z_k = x_i + v_j$, $k = 1,2,...9$, $i, j = 1,2,3$.

Соответствующие данным значениям z_k вероятности p_k можно вычислить по формуле умножения вероятностей $P_k = P(Z = z_k) = P(X = x_i) \cdot P(V = v_j)$, т. к. события $X = x_i$ и $V = v_j$ - независимы (исходим из независимости случайных величин X и Y) и наступают совместно (событие $\{Z = z_k\}$ ={совместное наступление событий $X = x_i$ и $V = v_j$}). Тогда распределение Z принимает вид

Z	-1	3	9	2	6	12	3	7	13
p	0,02	0,05	0,03	0,06	0,15	0.09	0,12	0,3	0,18

Рассмотрим значения $z_2 = z_7 = 3$. События $Z = z_2$ и $Z = z_7$ несовместны, поэтому вероятность наступления хотя бы одного из этих событий вычисляется по правилу сложения вероятностей

$$P(Z = z_2 \cup Z = z_7) = P(Z = z_2) + P(Z = z_7) = 0,05 + 0,12 = 0,17$$

Искомый закон распределения случайной величины Z получается после размещения z_k по возрастанию.

Z	-1	2	3	6	7	9	12	13
p	0,02	0,06	0,17	0,15	0,3	0,03	0,09	0,18

Математическое ожидание *M(Z)* и дисперсию *D(Z)* можно найти по формулам:

$M(Z)=\sum_{k=1}^{8} z_k p_k$; $D(Z)=M(Z^2)-[M(Z)]^2$, где $M(Z^2)=\sum_{k=1}^{8} z_k{}^2 p_k$.

Рассмотрим другой способ.

M(Z) и *D(Z)* можно найти через *M(X), M(Y), D(X), D(Y).*

$$M(X)=\sum_{i=1}^{3} x_i p_i = -3\cdot 0{,}1+0\cdot 0{,}3+1\cdot 0{,}6=0{,}3$$

$$M(Y)=\sum_{j=1}^{3} y_j p_j = 0\cdot 0{,}2+3\cdot 0{,}5+6\cdot 0{,}3=3{,}3$$

$$D(X)=\sum_{i=1}^{3} x_i{}^2 p_i-(M(X))^2=(-3)^2\cdot 0{,}1+0^2\cdot 0{,}3+1^2\cdot 0{,}6-(0{,}3)^2=1{,}41$$

$$D(Y)=\sum_{j=1}^{3} y_j{}^2 p_j-(M(Y))^2=0^2\cdot 0{,}2+3^2\cdot 0{,}5+6^2\cdot 0{,}3-(3{,}3)^2=4{,}41$$

$$M(Z)=M(X+2Y)=M(X)+M(2Y)=M(X)+2M(Y)=0{,}3+2\cdot 3{,}3=6{,}9,$$

т. к. математическое ожидание суммы равно сумме математических ожиданий слагаемых; постоянный множитель можно вынести за знак математического ожидания.

$$D(Z)=D(X+2Y)=D(X)+D(2Y)=D(X)+4D(Y)=1{,}41+4\cdot 4{,}41=19{,}05,$$

т. к. дисперсия суммы независимых случайных величин равна сумме дисперсий слагаемых; постоянный множитель можно вынести за знак дисперсии, возведя его в квадрат.

Пример 6.

Стрелок ведет стрельбу с вероятностью попадания в цель 0,8 при каждом выстреле. Стрельба ведется до первого попадания, но делается не более 3 выстрелов. Составить закон распределения случайной величины X, если: а) X – число промахов; б) X – число попаданий; в) X – число произведенных выстрелов.

Решение: Вероятность попадания p=0,8; вероятность промаха q=1-p=0,2.

а) Случайная величина X – число промахов при трех выстрелах – имеет следующие возможные значения: $x_1=0$; $x_2=1$; $x_3=2$; $x_4=3$.

Событие X=0 равносильно попаданию с первой попытки, следовательно, P(X=0)=p=0,8.

Событие X=1 равносильно попаданию со второй попытки, т. е. совместному наступлению двух событий: промаха и попадания; следовательно, P(X=1)=q·p=0,2·0,8=0,16.

Событие X=2 равносильно попаданию с третьей попытки, т. е. P(X=2)=q·q·p=0,2·0,2·0,8=0,032.

Событие X=3 означает отсутствие попаданий, P(X=3)=q·q·q=$0,2^3$=0,008.

Искомый закон распределения X:

X	0	1	2	3
p	0,8	0,16	0,032	0,008

б) Случайная величина X – число попаданий – имеет следующие возможные значения: $x_1 = 0$ (допущено три промаха); $x_2 = 1$ (произошло попадание с первой, второй или третьей попытки).

Тогда P(X=0)= q^3 =$0,2^3$=0,008;

P(X=1)= p+q·p+q·q·p=0,8+0,16+0,032=0,992

или P(X=1)=1-P(X=0)=1-0,008=0,992.

Искомый закон распределения X:

X	0	1
P	0,008	0,992

в) Случайная величина X – число произведенных выстрелов – имеет следующие возможные значения: $x_1 = 1; x_2 = 2; x_3 = 3$.

Событие X=1 равносильно попаданию с первой попытки, т. е. P(X=1)=p=0,8.

Событие X=2 равносильно попаданию со второй попытки, т. е. P(X=2)=q·p=0,16.

Событие X=3 означает, что либо произошло попадание с третьей попытки, либо было три промаха. Тогда P(X=3)=q·q·p+q·q·q=0,032+0,008=0,04.

Искомый закон распределения X:

X	1	2	3
P	0,8	0,16	0,04

Практическая работа №8. Расчет вероятностей. Составление закона распределения вероятностей

Исходные данные: Раздаточный материал

Содержание и порядок выполнения задания:

Задачи

Вариант 1. Производятся последовательные независимые испытания приборов на надёжность. Каждый следующий прибор испытывается лишь в том случае, если предыдущий оказался надежным. Построить закон распределения случайного числа испытанных приборов, если вероятность выдержать испытание для каждого из них равна 0,9. Найти математическое ожидание числа испытанных приборов. Найти функцию распределения F(x) и построить ее график; найти M(X), σ(X); построить многоугольник распределения.

Вариант 2. Известно, что в партии из 20 телефонных аппаратов 5 недействующих. Случайным образом из этой партии взято 4 аппарата. Построить закон распределения случайной величины X – числа недействующих аппаратов из отобранных. Найти дисперсию этой случайной величины. В каких единицах она измеряется? Построить график функции распределения F(x) случайной величины X, многоугольник распределения.

Вариант 3. Сырье на завод привозят от трех независимо работающих поставщиков. Вероятность своевременного прибытия сырья от первого поставщика равна 0,4, от второго – 0,7, от третьего – 0,6. Найти математическое ожидание M(X), дисперсию D(X) числа своевременных поставок сырья. Найти функцию распределения и построить ее график.

Вариант 4. Завод получает сырье на автомашинах от трех независимо работающих поставщиков. Вероятность прибытия автомашины от первого поставщика равна 0,2, от второго – 0,3 и от третьего – 0,1. Составить распределение числа прибывших автомашин. Найти математическое ожидание и дисперсию полученной величины. Построить график функции распределения F(x).

Вариант 5. Вероятность изготовления бракованной детали p=0,1. Изготовлено 4 детали. X – случайное число бракованных деталей. Построить закон распределения случайной величины X, найти ее математическое ожидание и дисперсию. Построить график функции распределения, многоугольник распределения.

Вариант 6. Среднее число заявок, поступающих на предприятие бытового обслуживания за 1 час, ровно 2. Составить закон распределения случайной величины X – числа заявок, поступивших за 3 часа. Найти M(X), D(X) и наивероятнейшее число заявок за 3 часа.

Вариант 7. В среднем в магазин заходит 3 человека в минуту. Составить закон распределения случайной величины X – числа зашедших в магазин человек за 2 минуты. Построить многоугольник распределения. Найти M(X), D(X).

Вариант 8. Даны законы распределения независимых случайных величин

X	-3	0	1
P	0,1	0,3	0,6

Y	0	3	6
p	0,2	0,5	0,3

Составить законы распределения случайных величин:

а)XY; б) X+Y. Найти M(X+Y), D(X+Y). Справедливо ли равенство M(X)·M(Y)=M(X·Y)?

Вариант 9. Команда состоит из двух стрелков. Числа очков, выбиваемых каждым из них при одном выстреле, являются случайными величинами X_1 и X_2 , которые характеризуются следующими законами распределения:

X_1	3	4	5
P	0,3	0,4	0,3

и

X_2	2	3	4	5
P	0,2	0,1	0,2	0,5

Результаты стрельбы одного стрелка не влияют на результат стрельбы другого. Составить закон распределения числа очков, выбиваемых командой, если стрелки сделают по одному выстрелу. Убедиться в справедливости равенства $D(X_1+X_2)=D(X_1)+D(X_2)$.

Вариант 10. Производятся выстрелы из орудия с вероятностью попадания в цель 0,9 при каждом выстреле. Стрельба ведётся до первого попадания, но делается не более 4 выстрелов. Составить закон распределения случайной величины X, если: а) X – число произведенных выстрелов; б) X – число промахов; в) X – число попаданий. Найдите математическое ожидание всех найденных случайных величин.

Практическая работа №9. ПУ кодирование

Исходные данные: Раздаточный материал

Содержание и порядок выполнения задания:

Кодирование и декодирование информации: ключевые понятия и методы

Введение

В данной лекции мы рассмотрим основные принципы и методы кодирования и декодирования информации. Кодирование информации является процессом преобразования данных из одной формы в другую, чтобы они могли быть переданы

или сохранены. Декодирование, в свою очередь, представляет собой обратный процесс, в результате которого закодированные данные восстанавливаются в исходную форму. Мы изучим различные методы кодирования и декодирования, а также рассмотрим примеры их применения.

Кодирование информации – это процесс преобразования данных из одной формы в другую, чтобы они могли быть переданы или сохранены с использованием определенных правил и соглашений. В контексте информатики, кодирование информации относится к преобразованию текста, изображений, звука или других типов данных в биты и байты, которые компьютеры могут обрабатывать и передавать.

Кодирование информации играет важную роль в передаче и хранении данных. Оно позволяет нам эффективно использовать ресурсы, уменьшая объем информации, необходимый для передачи или хранения. Кодирование также обеспечивает защиту данных, позволяя шифровать информацию и делать ее непонятной для посторонних лиц.

Кодирование информации может быть использовано для различных целей, включая сжатие данных, передачу данных по сети, хранение данных на носителях, шифрование и дешифрование информации и многое другое.

Зачем нужно кодирование информации

Кодирование информации играет важную роль в передаче и хранении данных. Оно позволяет нам эффективно использовать ресурсы, уменьшая объем информации, необходимый для передачи или хранения. Кодирование также обеспечивает защиту данных, позволяя шифровать информацию и делать ее непонятной для посторонних лиц.

Основная цель кодирования информации – это сократить объем данных, необходимых для передачи или хранения. Когда мы передаем информацию по сети или сохраняем ее на носителе, мы хотим сделать это максимально эффективно. Кодирование позволяет нам уменьшить размер данных, не теряя при этом существенной информации.

Кодирование также играет важную роль в защите данных. Шифрование информации позволяет нам делать ее непонятной для посторонних лиц. Это особенно важно при передаче конфиденциальной информации, такой как пароли, банковские данные или личные данные. Кодирование помогает обеспечить конфиденциальность и безопасность данных.

Кодирование информации также позволяет нам представлять данные в различных форматах. Например, мы можем закодировать текстовую информацию в бинарный формат для передачи по сети или сохранения на носителе. Мы также можем использовать различные методы кодирования для представления изображений, звуков или видео.

В целом, кодирование информации является важным инструментом для эффективной передачи, хранения и защиты данных. Оно позволяет нам использовать ресурсы более эффективно, обеспечивает безопасность данных и позволяет представлять информацию в различных форматах.

Принципы кодирования информации

Кодирование информации – это процесс преобразования данных из одной формы в другую, чтобы они могли быть переданы, сохранены или обработаны. Вот некоторые основные принципы кодирования информации:

Алфавит

Алфавит – это набор символов или знаков, которые используются для представления информации. Например, в текстовом кодировании алфавит может состоять из букв, цифр и специальных символов. В аудио-кодировании алфавит может представлять различные звуки или частоты.

Кодовые слова

Кодовые слова – это комбинации символов или знаков, которые представляют определенные значения или данные. Каждое кодовое слово соответствует определенному символу, числу или другому элементу информации. Например, в бинарном кодировании, где используются только два символа (0 и 1), каждое кодовое слово представляет определенное число или символ.

Кодирование и декодирование

Кодирование – это процесс преобразования исходной информации в кодовые слова, а декодирование – обратный процесс, при котором кодовые слова преобразуются обратно в исходную информацию. Кодирование и декодирование должны быть взаимно обратными операциями, чтобы информация могла быть правильно восстановлена.

Эффективность

Эффективность кодирования означает, что кодирование должно быть компактным и использовать минимальное количество бит или символов для представления информации. Чем более эффективно кодирование, тем меньше пропускной способности или памяти требуется для передачи или хранения данных.

Надежность

Надежность кодирования означает, что кодированная информация должна быть устойчива к ошибкам и искажениям при передаче или хранении. Для этого могут использоваться различные методы обнаружения и исправления ошибок, такие как добавление контрольных сумм или использование кодов с исправлением ошибок.

Это основные принципы кодирования информации, которые помогают нам эффективно представлять, передавать и хранить данные. Различные методы кодирования могут использовать разные алгоритмы и стратегии, но эти принципы остаются общими для всех методов.

Различные методы кодирования информации

Существует множество различных методов кодирования информации, каждый из которых имеет свои особенности и применяется в различных сферах. Рассмотрим некоторые из них:

Бинарное кодирование

Бинарное кодирование – это метод представления информации с помощью двух символов, обычно 0 и 1. Каждый символ называется битом (от англ. binary digit). Бинарное кодирование широко используется в цифровых системах, таких как компьютеры, где информация обрабатывается в виде битовых последовательностей.

Аналоговое кодирование

Аналоговое кодирование – это метод представления информации в виде непрерывных сигналов. Например, звуковые сигналы могут быть представлены аналоговыми волнами, где изменение амплитуды и частоты сигнала отображает звуковую информацию. Аналоговое кодирование используется в аналоговых системах связи и аудио-визуальных устройствах.

Числовое кодирование

Числовое кодирование – это метод представления информации с использованием чисел. Например, целые числа могут быть использованы для представления количественных данных, а дробные числа – для представления десятичных значений. Числовое кодирование широко применяется в математических и научных вычислениях, а также в финансовых и статистических приложениях.

Текстовое кодирование

Текстовое кодирование – это метод представления информации в виде текста, используя символы и символьные коды. Например, ASCII (American Standard Code for Information Interchange) – это одна из самых распространенных систем текстового кодирования, где каждому символу соответствует уникальный числовой код. Текстовое кодирование используется в компьютерных системах для представления и обработки текстовой информации.

Графическое кодирование

Графическое кодирование – это метод представления информации в виде графических изображений или символов. Например, изображения могут быть представлены с помощью пикселей, где каждый пиксель содержит информацию о цвете и яркости. Графическое кодирование широко используется в компьютерной графике, мультимедиа и визуальных приложениях.

Это лишь некоторые из множества методов кодирования информации, которые используются в различных областях. Каждый метод имеет свои преимущества и ограничения, и выбор метода зависит от конкретной задачи и требований.

Примеры кодирования информации

Бинарное кодирование

Бинарное кодирование – это метод представления информации с помощью двух символов, обычно 0 и 1. Каждый символ называется битом (от англ. binary digit). Биты объединяются в байты, которые представляют собой группы из 8 битов. Бинарное кодирование широко используется в компьютерах и цифровых системах.

Текстовое кодирование

Текстовое кодирование – это метод представления текстовой информации с помощью символов и кодов. Наиболее распространенным текстовым кодированием является ASCII (American Standard Code for Information Interchange), где каждый символ представлен числовым кодом от 0 до 127. Другие популярные текстовые кодировки включают Unicode и UTF-8, которые поддерживают широкий набор символов из разных языков и позволяют представлять текст на разных платформах.

Графическое кодирование

Графическое кодирование – это метод представления графической информации, такой как изображения или символы. Графические данные могут быть представлены с помощью пикселей, где каждый пиксель содержит информацию о цвете и яркости. Другие методы графического кодирования включают векторное кодирование, где графические объекты представлены математическими формулами, и компрессия изображений, которая уменьшает размер файла, сохраняя при этом качество изображения.

Звуковое кодирование

Звуковое кодирование – это метод представления звуковой информации, такой как музыка или речь. Звуковые данные могут быть представлены с помощью аналоговых или цифровых сигналов. Цифровое звуковое кодирование использует методы сжатия данных, такие как MP3 или AAC, чтобы уменьшить размер файла, сохраняя при этом качество звука.

Видео кодирование

Видео кодирование – это метод представления видео информации. Видео данные могут быть представлены с помощью последовательности изображений, называемых кадрами. Популярные методы видео кодирования включают форматы MPEG, где видео данные сжимаются с помощью методов сжатия, таких как MPEG-2 или MPEG-4, чтобы уменьшить размер файла, сохраняя при этом качество видео.

Это лишь некоторые из множества методов кодирования информации, которые используются в различных областях. Каждый метод имеет свои преимущества и ограничения, и выбор метода зависит от конкретной задачи и требований.

Практическая работа №10. Адаптивное арифметическое кодирование

Исходные данные: Раздаточный материал

Содержание и порядок выполнения задания:

Что такое декодирование информации

Декодирование информации – это процесс преобразования закодированной информации обратно в исходный вид. Когда информация кодируется, она преобразуется в другой формат или представление, чтобы удовлетворить определенные требования или цели. Декодирование обратно преобразует закодированную информацию в исходный формат, чтобы она стала понятной и доступной для использования.

Декодирование информации широко используется в различных областях, включая коммуникации, компьютерную графику, аудио и видео. В каждой из этих областей информация может быть закодирована для сжатия, защиты или передачи по сети. Декодирование позволяет восстановить исходную информацию из закодированного представления.

Процесс декодирования зависит от метода кодирования, который был использован. В некоторых случаях декодирование может быть простым и обратным процессом кодирования. Например, при использовании метода кодирования ASCII, каждый символ представлен числовым значением, и декодирование заключается в преобразовании числового значения обратно в символ.

Однако в других случаях декодирование может быть более сложным процессом, особенно при использовании сложных методов сжатия данных. Например, при декодировании видеофайла, закодированного с использованием формата MPEG, требуется выполнить ряд шагов для восстановления исходного видео. Это включает в себя декомпрессию данных, восстановление кадров и воспроизведение видео.

Важно отметить, что декодирование информации не всегда является точным процессом восстановления исходной информации. В некоторых случаях при декодировании может происходить потеря данных или качества. Например, при сжатии аудиофайла с потерями, такого как формат MP3, часть аудиоданных может быть удалена для уменьшения размера файла. При декодировании эти данные не могут быть полностью восстановлены, что может привести к потере качества звука.

В целом, декодирование информации играет важную роль в обработке и использовании закодированных данных. Оно позволяет преобразовать закодированную информацию обратно в исходный формат, делая ее доступной для понимания и использования.

Зачем нужно декодирование информации

Декодирование информации является важным процессом в обработке и использовании закодированных данных. Оно позволяет преобразовать закодированную информацию обратно в исходный формат, делая ее доступной для понимания и использования. Вот несколько основных причин, по которым декодирование информации является неотъемлемой частью обработки данных:

Восстановление исходной информации

Одной из основных целей декодирования информации является восстановление исходной информации из закодированного формата. Когда данные кодируются, они могут быть изменены или преобразованы для удобства хранения или передачи. Декодирование позволяет вернуть данные в их исходное состояние, чтобы они могли быть поняты и использованы.

Понимание и анализ данных

Декодирование информации также играет важную роль в понимании и анализе данных. Когда данные закодированы, они могут быть представлены в специальном формате, который не всегда понятен для человека. Декодирование позволяет преобразовать данные в формат, который легче понять и анализировать.

Обработка и использование данных

Декодирование информации также необходимо для обработки и использования данных. Когда данные закодированы, они могут быть недоступны для использования в различных приложениях или системах. Декодирование позволяет преобразовать данные в формат, который может быть использован в различных контекстах, таких как программирование, анализ данных, машинное обучение и другие области.

Восстановление качества данных

В некоторых случаях декодирование информации может быть использовано для восстановления качества данных. Например, при сжатии аудиофайла с потерями, такого как формат MP3, часть аудиоданных может быть удалена для уменьшения размера файла. При декодировании эти данные не могут быть полностью восстановлены, что может привести к потере качества звука.

В целом, декодирование информации играет важную роль в обработке и использовании закодированных данных. Оно позволяет преобразовать закодированную информацию обратно в исходный формат, делая ее доступной для понимания и использования.

Принципы декодирования информации

Декодирование информации – это процесс преобразования закодированной информации обратно в исходный формат или представление. Вот некоторые основные принципы, которые лежат в основе декодирования информации:

Знание кодирования

Для успешного декодирования информации необходимо иметь знание о том, как она была закодирована. Это включает в себя знание используемого кодирования, алгоритмов и методов, которые были применены для преобразования информации. Без этого знания декодирование может быть затруднено или невозможно.

Обратимость кодирования

Декодирование информации возможно только в том случае, если кодирование было обратимым процессом. Это означает, что при кодировании информации никакая информация не теряется или искажается, и она может быть полностью восстановлена при декодировании. Если кодирование не является обратимым, то декодирование может привести к потере части информации или искажению исходных данных.

Использование правильного декодера

Для декодирования информации необходимо использовать правильный декодер или программное обеспечение, способное распознавать и преобразовывать закодированные данные. Различные методы кодирования требуют разных декодеров, поэтому важно выбрать соответствующий декодер для конкретного типа кодирования.

Соответствие параметров декодирования

При декодировании информации необходимо учитывать параметры, используемые при кодировании. Например, если информация была закодирована с использованием определенного алгоритма сжатия данных, то при декодировании необходимо использовать тот же алгоритм и те же параметры сжатия для правильного восстановления данных.

Проверка целостности данных

Важным аспектом декодирования информации является проверка целостности данных. Это означает, что при декодировании необходимо убедиться, что данные были правильно восстановлены и не были повреждены или изменены в процессе декодирования. Для этого могут использоваться различные методы проверки целостности, такие как контрольные суммы или хэш-функции.

Все эти принципы важны для успешного декодирования информации. Они помогают обеспечить правильное восстановление данных и использование закодированной информации в исходном формате.

Различные методы декодирования информации

Декодирование информации – это процесс восстановления исходных данных из закодированного формата. Существует несколько различных методов

декодирования информации, которые могут быть использованы в зависимости от типа кодирования и формата данных. Рассмотрим некоторые из них:

Дешифрование

Дешифрование – это метод декодирования, который используется для восстановления данных, зашифрованных с использованием алгоритма шифрования. Для дешифрования необходимо знать ключ или алгоритм, используемый для шифрования данных. Примерами методов шифрования могут быть шифр Цезаря, шифр Виженера или симметричное шифрование.

Распаковка

Распаковка – это метод декодирования, который используется для восстановления данных, упакованных или сжатых с использованием алгоритма сжатия данных. Распаковка может включать в себя различные алгоритмы сжатия, такие как алгоритм Хаффмана, алгоритм Лемпеля-Зива-Велча или алгоритм DEFLATE. При распаковке данные восстанавливаются в исходном формате.

Декодирование изображений и звука

Декодирование изображений и звука – это методы декодирования, которые используются для восстановления изображений и звуковых данных из их закодированных форматов. Для декодирования изображений могут использоваться алгоритмы сжатия, такие как JPEG или PNG. Для декодирования звука могут использоваться алгоритмы сжатия, такие как MP3 или AAC.

Разархивирование

Разархивирование – это метод декодирования, который используется для восстановления данных из архивов или компрессированных файлов. Архивы могут содержать несколько файлов или папок, которые были упакованы в один файл для удобства хранения или передачи. Для разархивирования может использоваться алгоритм сжатия, такой как ZIP или RAR.

Это лишь некоторые из методов декодирования информации. В зависимости от конкретной ситуации и типа кодирования могут использоваться и другие методы декодирования. Важно выбрать правильный метод декодирования для успешного восстановления и использования данных в исходном формате.

Примеры декодирования информации

Декодирование текста из шифра Цезаря

Шифр Цезаря – это метод шифрования, в котором каждая буква заменяется на другую букву, сдвинутую на определенное количество позиций в алфавите. Для декодирования текста, зашифрованного шифром Цезаря, необходимо знать количество позиций сдвига и применить обратную операцию – сдвиг в обратном направлении.

Декодирование изображений из формата JPEG

JPEG – это формат сжатия изображений, который используется для уменьшения размера файла без значительной потери качества изображения. Для декодирования изображений из формата JPEG необходимо использовать соответствующий алгоритм декомпрессии, который восстанавливает исходное изображение из сжатого файла.

Декодирование аудио из формата MP3

MP3 – это формат сжатия аудио, который позволяет уменьшить размер файла без существенной потери качества звука. Для декодирования аудио из формата MP3 необходимо использовать соответствующий алгоритм декомпрессии, который восстанавливает исходное аудио из сжатого файла.

Декодирование видео из формата MPEG

MPEG – это формат сжатия видео, который используется для уменьшения размера файла без значительной потери качества видео. Для декодирования видео из формата MPEG необходимо использовать соответствующий алгоритм декомпрессии, который восстанавливает исходное видео из сжатого файла.

Декодирование данных из архивов ZIP

ZIP – это формат архивации, который позволяет объединить несколько файлов или папок в один файл для удобства хранения или передачи. Для декодирования данных из архивов ZIP необходимо использовать соответствующий алгоритм разархивации, который восстанавливает исходные файлы или папки из архива.

Это лишь некоторые из примеров декодирования информации. В зависимости от конкретной ситуации и типа кодирования могут использоваться и другие методы декодирования. Важно выбрать правильный метод декодирования для успешного восстановления и использования данных в исходном формате.

Таблица сравнения методов кодирования информации

Метод	Описание	Преимущества	Недостатки
Бинарное кодирование	Кодирование информации с использованием двух символов: 0 и 1	Простота реализации, эффективность использования памяти	Ограниченный набор символов, сложность чтения и понимания
Алфавитное кодирование	Кодирование информации с использованием символов алфавита	Большой набор символов, легкость чтения и понимания	Большое количество символов, требуется больше памяти для хранения
Шифрование	Преобразование информации с использованием специального алгоритма	Высокая степень защиты информации, сложность взлома	Требуется дополнительное время для шифрования и дешифрования

Заключение

Кодирование информации является важным процессом, который позволяет представить данные в определенном формате для их передачи и хранения. Оно позволяет сократить объем информации и обеспечить ее безопасность. Декодирование информации, в свою очередь, позволяет восстановить исходные

данные из закодированного формата. Знание принципов и методов кодирования и декодирования информации является необходимым для работы с различными типами данных и системами передачи информации.

Практическая работа №11. Цифровое кодирование и аналоговое кодирование. Таблично-символьное кодирование

Исходные данные: Раздаточный материал

Содержание и порядок выполнения задания:

Криптография – обеспечивает сокрытие смысла сообщения с помощью шифрования и открытия его расшифровкой, которые выполняются по специальным алгоритмам с помощью ключей.

Ключ – конкретное секретное состояние некоторых параметров алгоритма криптографического преобразования данных, обеспечивающее выбор только одного варианта из всех возможных для данного алгоритма.

Криптоанализ – занимается вскрытием шифра без знания ключа (проверка устойчивости шифра).

Кодирование – (не относится к криптографии) – система условныхобозначений, применяемых при передаче информации. Применяется для увеличения качества передачи информации, сжатия информации и для уменьшения стоимости хранения и передачи.

Криптографические преобразования имеют цель обеспечить недоступность информации для лиц, не имеющих ключа, и поддержание с требуемой надежностью обнаружения несанкционированных искажений.

Большинство средств защиты информации базируется на использовании криптографических шифров и процедур шифрования-расшифровки. В соответствии со стандартом ГОСТ 28147-89 под **шифром** понимают совокупность обратимых преобразований множества открытых данных на множество зашифрованных данных, задаваемых ключом и алгоритмом преобразования.

В криптографии используются следующие основные алгоритмы шифрования:

- алгоритм замены (подстановки) – символы шифруемого текста заменяются символами того же или другого алфавита в соответствии с заранее обусловленной схемой замены;
- алгоритм перестановки – символы шифруемого текста переставляются по определенному правилу в пределах некоторого блока этого текста;
- гаммирование – символы шифруемого текстаскладываются с символами некоторой случайной последовательности;
- аналитическое преобразование – преобразование шифруемого текста по некоторому аналитическому правилу (формуле).

Процессы шифрования и расшифровки осуществляются в рамках некоторой криптосистемы. Для **симметричной** криптосистемы характерно применение одного и того же ключа как при шифровании, так и при расшифровке сообщений. В **асимметричных** криптосистемах для шифрования данныхиспользуется один (общедоступный) ключ, а для расшифровки – другой (секретный) ключ.

Р-12. *Для кодирования некоторой последовательности, состоящей из букв А, Б, В, Г, решили использовать неравномерный двоичный код, удовлетворяющий условию Фано. Для буквы А использовали кодовое слово 0, для буквы Б – кодовое слово 110. Какова наименьшая возможная суммарная длина всех четырёх кодовых слов?*

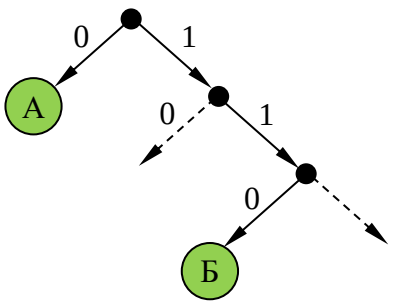
1) 7 2) 8 3) 9 4) 10

Решение (способ 1, исключение вариантов):

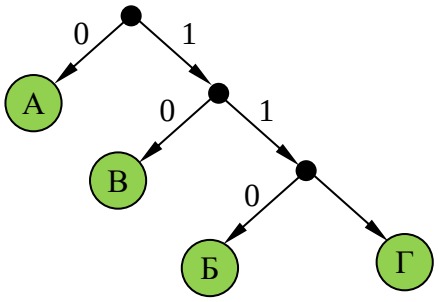
1) условие Фано означает, что ни одно кодовое слово не совпадает с началом другого кодового слова
2) поскольку уже есть кодовое слово 0, ни одно другое кодовое слово не может начинаться с 0
3) поскольку есть код 110, запрещены кодовые слова 1, 11; кроме того, ни одно другое кодовое слово не может начинаться с 110
4) таким образом, нужно выбрать еще два кодовых слова, для которых выполняются эти ограничения
5) есть одно допустимое кодовое слово из двух символов: 10
6) если выбрать кодовое слово 10 для буквы В, то остаётся одно допустимое трёхсимвольное кодовое слово – 111, которое можно выбрать для буквы Г
7) таким образом, выбрав кодовые слова А – 0, Б – 110, В – 10, Г – 111, получаем суммарную длину кодовых слов 9 символов
8) если же не выбрать В – 10, то есть три допустимых трёхсимвольных кодовых слова: 100, 101 и 110; при выборе любых двух их них для букв В и Г получаем суммарную длину кодовых слов 10, что больше 9; поэтому выбираем вариант 3 (9 символов)
9) Ответ: 3.

Решение (способ 2, построение дерева):

1) условие Фано означает, что ни одно кодовое слово не совпадает с началом другого кодового слова; при этом в дереве кода все кодовые слова должны располагаться в листьях дерева, то есть в узлах, которые не имеют потомков;
2) построим дерево для заданных кодовых слов А – 0 и Б – 110:



3) штриховыми линиями отмечены две «пустые» ветви, на которые можно «прикрепить» листья для кодовых слов букв В (10) и Г (111)



4) таким образом, выбрав кодовые слова А – 0, Б – 110, В – 10, Г – 111, получаем суммарную длину кодовых слов 9 символов
5) Ответ: 3.

Ещё пример задания

Р-11. *По каналу связи передаются сообщения, содержащие только 5 букв А, И, К, О, Т. Для кодирования букв используется неравномерный двоичный код с такими кодовыми словами:*
А — 0, И — 00, К — 10, О — 110, Т — 111.
Среди приведённых ниже слов укажите такое, код которого можно декодировать только одним способом. Если таких слов несколько, укажите первое по алфавиту.

1) КАА 2) ИКОТА 3) КОТ 4) ни одно из сообщений не подходит

Решение:

1) прежде всего заметим, что для заданного кода не выполняется ни прямое, ни обратное условие Фано; «виновата» в этом пара А – И: код буквы А совпадает как с началом, так и с окончанием кода буквы И; больше ни для одной пары кодовых слов прямое условие Фано не нарушено
2) это означает, что не все сообщения могут быть декодированы однозначно
3) теперь нужно понять, какие последовательности могут быть декодированы неоднозначно; в данном случае очевидно, что сообщения АА и И кодируются одинаково: 00, поэтому все слова, где есть АА или И, не могут быть декодированы однозначно
4) поэтому варианты 1 (КАА) и 2 (ИКОТА) отпадают
5) на всякий случай проверим вариант 3: КОТ = 10110111; первой буквой может быть только К (по-другому сочетание 10 получить нельзя), аналогично вторая буква – только О, а третья – только Т
6) Ответ: 3.

Ещё пример задания

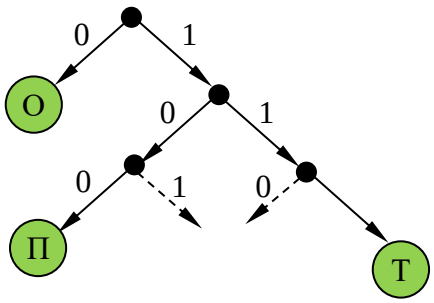
Р-10. *По каналу связи передаются сообщения, содержащие только 4 буквы П, О, С, Т; для передачи используется двоичный код, допускающий однозначное декодирование. Для букв Т, О, П используются такие кодовые слова: Т: 111, О: 0, П: 100.*
Укажите кратчайшее кодовое слово для буквы С, при котором код будет допускать однозначное декодирование. Если таких кодов несколько, укажите код с наименьшим числовым значением.

Решение (способ 1, исключение вариантов):

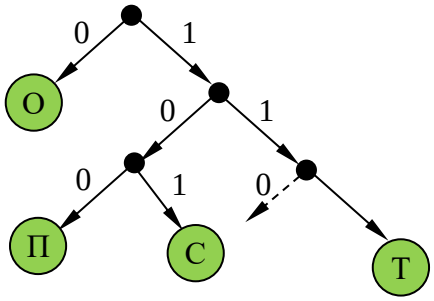
1) код однозначно декодируется, если выполняется условие Фано или обратное условие Фано; в данном случае «прямое» условие Фано выполняется: с кода буквы О (0) не начинается ни один из двух других кодов;
2) новый код не может начинаться с нуля (иначе нарушится условие Фано)
3) начнём проверку с кодов длиной 1; единственный код, не начинающийся с нуля – 1 – не подходит, потому что с 1 начинаются два других кода: Т (111) и П (100
4) кодов длиной 2, начинающихся с 1, всего 2: 10 и 11, но их использовать нельзя, потому что с 10 начинается код буквы П, а с 11 – код буквы Т
5) рассматриваем коды длиной 3, начинающиеся с 1; коды 100 и 111 уже заняты, а ещё два – 101 и 110 – свободны и их можно использовать, причём условие Фано выполняется в обоих случаях;
6) поскольку нужно выбрать код с минимальным значением, выбираем 101
7) Ответ: 101.

Решение (способ 2, построение дерева):

1) условие Фано означает, что ни одно кодовое слово не совпадает с началом другого кодового слова; при этом в дереве кода все кодовые слова должны располагаться в листьях дерева, то есть в узлах, которые не имеют потомков;
2) построим дерево для заданных кодовых слов О – 0, Т – 111 и П – 100:



3) штриховыми линиями отмечены две «пустые» ветви, на которые можно «прикрепить» лист для кодового слова буквы С: 101 или 110; из них минимальное значение имеет код 101



4) таким образом, выбрав кодовые слова А – 0, Б – 110, В – 10, Г – 111, получаем суммарную длину кодовых слов 9 символов
5) Ответ: 101.

Ещё пример задания

Р-09. *Для кодирования некоторой последовательности, состоящей из букв А, Б, В, Г и Д, используется неравномерный двоичный код, позволяющий однозначно декодировать полученную двоичную последовательность. Вот этот код: А – 0; Б – 100; В – 1010; Г – 111; Д – 110. Требуется сократить для одной из букв длину кодового слова так, чтобы код по-прежнему можно было декодировать однозначно. Коды остальных букв меняться не должны. Каким из указанных способов это можно сделать?*

1) для буквы В – 101 2) это невозможно
3) для буквы В – 010 4) для буквы Б – 10

Решение:

1) код однозначно декодируется, если выполняется условие Фано или обратное условие Фано; в данном случае «прямое» условие Фано выполняется: с кода буквы А (0) не начинается ни один другой код, оставшиеся короткие коды (Б, Г и Д) не совпадают с началом длинного кода буквы В; таким образом, при сокращении нужно сохранить выполнение условия Фано
2) вариант 3 не подходит, потому что новый код буквы В начинается с 0 (кода А), поэтому условие Фано нарушено
3) вариант 4 не подходит, потому что код буквы В начинается с 10 (нового кода б), поэтому условие Фано нарушено
4) вариант 1 подходит, условие Фано сохраняется (все трёхбитные коды различны, ни один не начинается с 0)
5) Ответ: 1.

Практическая работа №12. Практическое применение криптографии. Изучение и сравнительный анализ методов шифрования.

Исходные данные: Раздаточный материал

Содержание и порядок выполнения задания:

Асимметричные криптосистемы

Схема шифрования Эль Гамаля

Алгоритм шифрования Эль Гамаля основан на применении больших чисел для генерации открытого и закрытого ключа, криптостойкость же обусловлена сложностью вычисления дискретных логарифмов.

Последовательность действий пользователя:

1. Получатель сообщения выбирает два больших числа P и G, причем $P > G$.
2. Получатель выбирает секретный ключ - случайное целое число $X < P$.
3. Вычисляется открытый ключ $Y = G^x \bmod P$.
4. Получатель выбирает целое число K, $1 < K < P-1$.
5. Шифрование сообщения (M): $a = G^K \bmod P$, $b = Y^K M \bmod P$, где пара чисел (a,b) является шифротекстом.

Криптосистема шифрования данных RSA

Предложена в 1978 году авторами Rivest, Shamir и Aldeman и основана на трудности разложения больших целых чисел на простые сомножители.

Алгоритм создания открытого и секретного ключей:

1. Получатель выбирает 2 больших простых целых числа *p* и *q*, на основе которых вычисляет $n=p*q$ и функцию Эйлера $\varphi(n)=(p-1)(q-1)$.
2. Получатель выбирает целое число e ($1<e<\varphi(n)$), взаимно простое со значением функции $\varphi(n)$.

Пара чисел **(e,n)** публикуется в качестве **открытого ключа**.

3. Получатель вычисляет целое число *d*, которое отвечает условию: $e*d=1 \pmod{\varphi(n)}$.

Пара чисел **(d,n)** является **секретным ключом**. Шифрование сообщения с использованием открытого ключа:
Если *m* – сообщение (сообщениями являются целые числа в интервале от 0до n-1), то зашифровать это сообщение можно как $\mathbf{c=m^e \bmod(n)}$.

Дешифрование сообщения с использованием секретного ключа: Получатель расшифровывает, полученное сообщение *c*: $\mathbf{m=c^d \bmod (n)}$.

3. Задание

Практическая работа состоит из двух частей:

Часть 1 – применение одного из алгоритмов симметричного шифрования;
Часть 2 – шифрование с использованием алгоритма RSA.

Порядок выполнения работы:Часть 1:

1. Используя один из алгоритмов симметричного шифрования (см. вариант), зашифровать свои данные: фамилию, имя, отчество.
2. Выполнить проверку, расшифровав полученное сообщение.

Часть 2:

1. Написать программу, реализующую алгоритм шифрования и дешифрования сообщения RSA. Входные данные:открытый и секретный ключи (значения n, e, d) и сообщение (m).
2. Используя заданные значения p, q, e, d (см. вариант) зашифровать и дешифровать сообщения m_1, m_2, m_3 (см. вариант).

Практическая работа №13. Криптография с симметричным ключом, с открытым ключом. Шифрование с использованием перестановок

Исходные данные: Раздаточный материал

Содержание и порядок выполнения задания:

Симметричные криптосистемы Шифры перестановки

В шифрах средних веков часто использовались таблицы, с помощью которых выполнялись простые процедуры шифрования, основанные на перестановке букв в сообщении. Ключом в данном случае является размеры таблицы. Например, сообщение “Неясное становится еще более непонятным” записывается в таблицу из 5 строк и 7 столбцов по столбцам:

Н	О	Н	С	Б	Н	Я
Е	Е	О	Я	О	Е	Т
Я	С	В	Е	Л	П	Н
С	Т	И	Щ	Е	О	Ы
Н	А	Т	Е	Е	Н	М

Для получения шифрованного сообщения текст считывается по строкам и группируется по 5 букв:

НОНСБ НЯЕЕО ЯОЕТЯ СВЕЛП НСТИЩ ЕОЫНА ТЕЕНМ

Несколько большей стойкостью к раскрытию обладает **метод одиночной перестановки** по ключу. Он отличается от предыдущего тем, что столбцы таблицы переставляются по ключевому слову, фразе или набору чисел длиной в строку таблицы. Используя в качестве ключа слово «ЛУНАТИК», получим следующую таблицу:

Л	У	Н	А	Т	И	К			А	И	К	Л	Н	Т	У
4	7	5	1	6	2	3			1	2	3	4	5	6	7
Н	О	Н	С	Б	Н	Я			С	Н	Я	Н	Н	Б	О
Е	Е	О	Я	О	Е	Т			Я	Е	Т	Е	О	О	Е
Я	С	В	Е	Л	П	Н			Е	П	Н	Я	В	Л	С
С	Т	И	Щ	Е	О	Ы			Щ	О	Ы	С	И	Е	Т
Н	А	Т	Е	Е	Н	М			Е	Н	М	Н	Т	Е	А

До перестановки После перестановки

В верхней строке левой таблицы записан ключ, а номера под буквами ключа определены в соответствии с естественным порядком соответствующих букв

ключа в алфавите. Если в ключе встретились бы одинаковые буквы, они бы нумеровались слева направо. Получается шифровка:

СНЯНН БОЯЕТ ЕООЕЕ ПНЯВЛ СЩОЫС ИЕТЕН МНТЕА

Для обеспечения дополнительной скрытности можно повторно шифровать сообщение, которое уже было зашифровано. Для этого размер второй таблицы подбирают так, чтобы длины ее строк и столбцов отличались от длин строк и столбцов первой таблицы. Лучше всего, если они будут взаимно простыми.

Кроме алгоритмов одиночных перестановок применяются **алгоритмы двойных перестановок**. Сначала в таблицу записывается текст сообщения, а потом поочередно переставляются столбцы, а затем строки. При расшифровке перестановки проводятся в обратном порядке. Например, сообщение “Приезжаю_шестого” можно зашифровать следующим образом:

	2	4	1	3			1	2	3	4			1	2	3	4
4	П	Р	И	Е		4	И	П	Е	Р		1	А	З	Ю	Ж
1	З	Ж	А	Ю		1	А	З	Ю	Ж		2	Е	_	С	Ш
2	_	Ш	Е	С		2	Е.	_	С	Ш		3	Г	Т	О	О
3	Т	О	Г	О		3	Г	Т	О	О		4	И	П	Е	Р

Двойная перестановка столбцов и строк

В результате перестановки получена шифровка АЗЮЖЕ_СШГТООИПЕР. Ключом к шифру служат номера столбцов 2413 и номера строк 4123 исходной таблицы.

Число вариантов двойной перестановки достаточно быстро возрастает с увеличением размера таблицы: для таблицы 3 х 3 их 36, для 4 х 4 их 576, а для 5*5 их 14400.

В средние века для шифрования применялись и **магические квадраты**. Магическими квадратами называются квадратные таблицы с вписанными в их клетки последовательными натуральными числами, начиная с единицы, которые дают в сумме по каждому столбцу, каждой строке и каждой диагонали одно и то же число. Для шифрования необходимо вписать исходный текст по приведенной в квадрате нумерации и затем переписать

содержимое таблицы по строкам. В результате получается шифротекст, сформированный благодаря перестановке букв исходного сообщения.

16	3	2	13			О	И	Р	Т
5	10	11	8			З	Ш	Е	Ю
9	6	7	12			_	Ж	А	С
4	15	14	1			Е	Г	О	П

П	Р	И	Е	З	Ж	А	Ю	_	Ш	Е	С	Т	О	Г	О
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16

Число магических квадратов очень резко возрастает с увеличением размера его сторон: для таблицы 3*3 таких квадратов -1; для таблицы 4*4 - 880; а для таблицы 5*5-250000.

Практическая работа №14. Шифрование с использованием замен

Исходные данные: Раздаточный материал

Содержание и порядок выполнения задания:

Шифры простой замены

Система шифрования Цезаря - частный случай шифра простой замены. Метод основан на замене каждой буквы сообщения на другую букву того же алфавита, путем смещения от исходной буквы на K букв.

Известная фраза Юлия Цезаря VENI VINI VICI – пришел, увидел, победил, зашифрованная с помощью данного метода, преобразуется в SBKF SFAF SFZF (при смещении на 4 символа).

Греческим писателем Полибием за 100 лет до н.э. был изобретен так называемый **полибианский квадрат** размером 5*5, заполненный алфавитомв случайном порядке. Греческий алфавит имеет 24 буквы, а 25-м символом является пробел. Для шифрования на квадрате находили букву текста и записывали в шифротекст букву, расположенную ниже ее в том же столбце. Если буква оказывалась в нижней строке таблицы, то брали верхнюю букву из того же столбца.

Шифры сложной замены

Шифр Гронсфельда состоит в модификации шифра Цезаря числовым ключом. Для этого под буквами сообщения записывают цифры числового ключа. Если ключ короче сообщения, то его запись циклически повторяют. Шифротекст получают примерно также, как в шифре Цезаря, но отсчитывают не третью букву по алфавиту (как в шифре Цезаря), а ту, которая смещена по алфавиту на соответствующую цифру ключа.

Пусть в качестве ключа используется группа из трех цифр – 314, тогда
Сообщение: СОВЕРШЕННО СЕКРЕТНО

Ключ: 3143143143143143143

Шифровка: ФПИСЬИОССАХИЛФИУСС

В **шифрах многоалфавитной замены** для шифрования каждого символа исходного сообщения применяется свой шифр простой замены (свойалфавит):

	АБВГДЕЁЖЗИКЛМНОПРСТУФХЧШЩЪЫЬЭЮЯ_
А	АБВГДЕЁЖЗИКЛМНОПРСТУФХЧШЩЪЫЬЭЮЯ_
Б	_АБВГДЕЁЖЗИКЛМНОПРСТУФХЧШЩЪЫЬЭЮЯ
В	Я_АБВГДЕЁЖЗИКЛМНОПРСТУФХЧШЩЪЫЬЭЮ
Г	ЮЯ_АБВГДЕЁЖЗИКЛМНОПРСТУФХЧШЩЪЫЬЭ
.	……
Я	ВГДЕЁЖЗИКЛМНОПРСТУФХЧШЩЪЫЬЭЮЯ_АБ
_	БВГДЕЁЖЗИКЛМНОПРСТУФХЧШЩЪЫЬЭЮЯ_А

Каждая строка в этой таблице соответствует одному шифру замены аналогично шифру Цезаря для алфавита, дополненного пробелом. При шифровании сообщения его выписывают в строку, а под ним ключ. Если ключ оказался короче сообщения, то его циклически повторяют. Шифротекст получают, находя символ в колонке таблицы по букве текста и строке, соответствующей букве ключа. Например, используя ключ АГАВА, из сообщения ПРИЕЗЖАЮ ШЕСТОГО получаем следующую шифровку:

Сообщение	ПРИЕЗЖАЮ_ШЕСТОГО
Ключ	АГАВААГАВААГАВАА
Шифровка	ПНИГЗЖЮЮЮАЕОТМГО

Гаммирование

Процесс шифрования заключается в генерации гаммы шифра и наложении этой гаммы на исходный открытый текст. Перед шифрованием открытые данные разбиваются на блоки $T(0)_i$ одинаковой длины (по 64 бита). Гамма шифра

вырабатывается в виде последовательности блоков $Г(ш)_i$ аналогичной длины ($Т(ш)_i=Г(ш)_i+Т(0)_i$, где + - побитовое сложение, i =1-m).

Процесс расшифровки сводится к повторной генерации шифра текста и наложение этой гаммы на зашифрованные данные $Т(0)_i=Г(ш)_i+Т(ш)_i$.