

Федеральное государственное бюджетное образовательное учреждение высшего образования
«Петербургский государственный университет путей сообщения
Императора Александра I»
(ФГБОУ ВО ПГУПС)

Петрозаводский филиал ПГУПС

ОДОБРЕНО

на заседании цикловой комиссии
протокол № 11 от 23.06.2017

Председатель цикловой комиссии:

С.И. Ковалева (Ковалева)

УТВЕРЖДАЮ

Начальник УМО

А.В. Калько

А.В. Калько

«23» 06

2017 г.

МЕТОДИЧЕСКИЕ УКАЗАНИЯ

по организации и проведению лабораторных работ

По дисциплине ОП.01. Основы теории информации

Специальность: 09.02.02

Выполнила: преподаватель ПФ ПГУПС Капланова Марианна
Михайловна

2017 г.

ПОЯСНИТЕЛЬНАЯ ЗАПИСКА

Методические указания по организации и проведению лабораторных работ разработаны в соответствии с рабочей программой учебной дисциплины По дисциплине ОП.01. Основы теории информации специальности: модуля и предназначены для выполнения лабораторных работ обучающимися.

Лабораторные работы по учебной дисциплине По дисциплине ОП.01. Основы теории информации направлены на усвоение знаний, освоение умений и формирование элементов общих компетенций, предусмотренных рабочей программой учебной дисциплины.

В результате освоения учебной дисциплины обучающийся должен **уметь:**

- применять закон аддитивности информации;
- применять теорему Котельникова;
- использовать формулу Шеннона;

знать:

- виды и формы представления информации;
- методы и средства определения количества информации;
- принципы кодирования и декодирования информации;
- способы передачи цифровой информации;

В результате освоения учебной дисциплины По дисциплине ОП.01.

Основы теории информации

происходит поэтапное формирование элементов общих и/или профессиональных компетенций:

В результате освоения учебной дисциплины происходит поэтапное формирование элементов общих компетенций:

ОК 1. Понимать сущность и социальную значимость своей будущей профессии, проявлять к ней устойчивый интерес.

ОК 2. Организовывать собственную деятельность, выбирать типовые методы и способы выполнения профессиональных задач, оценивать их эффективность и качество.

ОК 4. Осуществлять поиск и использование информации, необходимой для эффективного выполнения профессиональных задач, профессионального и личностного развития.

ОК 8. Самостоятельно определять задачи профессионального и личностного развития, заниматься самообразованием, осознанно планировать повышение квалификации.

ОК 9. Ориентироваться в условиях частой смены технологий в профессиональной деятельности.

ПК 1.3. Обеспечивать защиту информации в сети с использованием программно-аппаратных средств.

ПК 2.1. Администрировать локальные вычислительные сети и принимать меры по устранению возможных сбоев.

ПК 2.2. Администрировать сетевые ресурсы в информационных системах.

ПК 3.2. Проводить профилактические работы на объектах сетевой инфраструктуры и рабочих станциях.

Рабочей программой предусмотрено выполнение обучающимися лабораторных работ, включая, как обязательный компонент практические задания с использованием персонального компьютера.

Раздел, тема	Контрольно-оценочные мероприятия	Результаты		Поэтапно формируемые элементы общих и профессиональных компетенций
		усвоенные знания	освоенные умения	
Раздел 1. Базовые понятия теории информации				
Тема 1.1. Формальное представление знаний. Виды информации и единицы ее измерения	Тестирование Подготовка докладов в форме презентации	виды и формы представления информации;		ОК 1., ОК 2. ОК 4.
Тема 1.2. Методы и средства определения количества информации	Лабораторная работа №1 Измерение количества информации с помощью меры Хартли и меры Шеннона	виды и формы представления информации; методы и средства определения количества информации;	применение закона аддитивности информации; использование формулы Шеннона;	ОК 2.
Раздел 2. Информация и энтропия. Принципы кодирования и декодирования информации				
Тема 2.1. Энтропия	Тестирование Анализ и решение ситуационных задач	методы и средства определения количества информации;	применение закона аддитивности информации; использование формулы Шеннона;	ОК 2. , ОК 4.
Тема 2.2 Преобразование аналоговой информации в цифровую	Задание по переработке информации различных источников	виды и формы представления информации; методы и средства определения количества информации; способы передачи цифровой информации;	применение теоремы Котельникова; использование формулы Шеннона;	ОК 1., ОК 2. , ОК 8.
Раздел 3. Сжатие и передача информации				

<p>Тема 3.1. Принципы кодирования и декодирования информации Сжатие информации.</p>	<p>Лабораторная работа №2 Избыточность сообщения. Эффективное кодирование по методу Шеннона – Фано Лабораторная работа №3 Эффективное кодирование по методу Хаффмена</p>	<p>методы и средства определения количества информации; принципы кодирования и декодирования информации; способы передачи цифровой информации; основы теории сжатия данных</p>	<p>применение закона аддитивности информации; использование формулы Шеннона;</p>	<p>ОК 1., ОК 2., ОК 4. ПК 1.3. ПК 3.2.</p>
<p>Тема 3.2. Арифметическое кодирование.</p>	<p>Лабораторная работа №4 Восстановление переданного сообщения по методу Хемминга</p>	<p>принципы кодирования и декодирования информации; способы передачи цифровой информации; методы повышения помехозащищенности передачи и приема данных, основы теории сжатия данных</p>	<p>применение закона аддитивности информации;</p>	<p>ОК 1., ОК 2. , ОК 4. ПК 1.3. ПК 2.1. ПК 3.2.</p>
<p>Раздел 4. Основы теории защиты информации</p>				
<p>Тема 4.1. Стандарты шифрования данных. Криптография.</p>	<p>Подготовка презентаций Выполнение исследовательской работы «Вскрытие шифра»</p>	<p>методы повышения помехозащищенности передачи и приема данных, основы теории сжатия данных</p>	<p>применение закона аддитивности информации; применение формулы Шеннона;</p>	<p>ОК 1., ОК 2. , ОК 4. ПК 1.3. ПК 2.1. ПК 2.2. ПК 3.2.</p>

КРИТЕРИИ ОЦЕНИВАНИЯ ЛАБОРАТОРНЫХ РАБОТ

При оценке освоенных умений при выполнении лабораторных работ применяется пятибалльная шкала оценивания

Оценивание лабораторных работ производится в соответствии со следующими нормативными актами:

- Положение о текущем контроле успеваемости и промежуточной аттестации обучающихся;
- Положение о планировании, организации и проведении лабораторных работ и практических занятий.

Лабораторная работа № 1
Тема Измерение количества информации. Мера Шеннона

Цель: исследовать влияние вероятности на количество информации,

Постановка задачи

Задача 1

бросаются две кости

Определить количество информации, содержащееся в сообщении о том, что выпавшие числа будут:

1. Только четные
2. Одно четное, другое нечетное
3. Сумма которых четна;
4. Сумма которых нечетна
5. Равны
6. Сумма которых больше, чем их произведение
7. Сумма которых меньше чем их произведение
8. Сумма которых меньше шести
9. Сумма которых больше восьми
10. Различны
11. Сумма которых делится на три
12. Произведение которых делится на четыре
13. Разность между которыми два
14. Только нечетные
15. Только одно делится на три

Задача 2

Варианты	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Количество вопросов N	50	60	50	60	50	60	50	60	50	60	50	60	50	60	50
Сколько знает вопросов K	30	30	35	35	40	40	25	25	20	20	35	35	40	45	45
Сколько в билете R	3	2	4	2	2	3	3	4	4	2	3	3	4	2	3

Студент идет сдавать экзамен, зная M вопросов из N . Чтобы сдать экзамен студент должен знать хотя бы один вопрос из R выпавших. Студент сдал экзамен. Какое количество информации содержится в этом сообщении

Задача 3

Имеется два ящика. В каждом из которых по n шаров. В первом a белых, b красных, c синих. Во втором – по d шаров каждого цвета. Опыт состоит в вытаскивании по одному шару из каждого ящика. Каково количество информации содержащееся в сообщении об исходе опыта? Сравнить количество информации в первом и во втором опыте и сделать вывод

вариант	N	a	b	c	d
0	24	11	7	6	8
1	63	10	50	3	21
2	78	11	48	19	26
3	78	12	46	20	26
4	69	13	44	12	23
5	72	14	42	16	24
6	75	15	40	20	25
7	72	16	38	18	24
8	66	17	36	13	22
9	75	18	34	23	25
10	81	19	32	30	27
11	84	20	30	34	28
12	66	21	28	17	22
13	87	22	26	39	29
14	90	23	24	43	30
15	60	24	22	14	20

Образец выполнения задания

Задача 1

Бросаются две кости

Определить количество информации, содержащееся в сообщении о том, что:

Произведение числа выпавших очков четное

Пусть A - событие, состоящее в том, что при бросании двух костей, произведение числа выпавших очков- четное

Тогда, $P(A)$ - вероятность наступления этого события равна $P(A) = \frac{m}{n}$

Число n - число возможных комбинаций составит $6^2 = 36$ (так как кубиков 2, а граней -6)

Пусть A - событие, состоящее в том, что при бросании двух костей, произведение числа выпавших очков- четное

Тогда, $P(A)$ - вероятность наступления этого события равна

$$P(A) = \frac{m}{n}$$

Число n - число возможных комбинаций составит $6^2 = 36$ (так как кубиков 2, а граней -6)

$$P(A) = \frac{6}{36} = 0.167$$

Число m благоприятствующих событию A комбинаций составит:

1 кость	2 кость
2	2
2	4
2	6
4	2
4	4
4	6
6	2
6	4
6	6

$$m=6 \quad P(A) = \frac{6}{36} = 0.167$$

Количество информации, содержащееся в этом сообщении, вычисленное по формуле Шеннона: $I = -\log_2(0.167) = 2.585$ (бит)

Задача 2

Студент идет сдавать экзамен, зная 6 вопросов из 100. Чтобы сдать экзамен студент должен знать хотя бы один вопрос из 5 выпавших. Студент сдал экзамен. Какое количество информации содержится в этом сообщении

Пусть В- событие, состоящее в том, что студент сдал экзамен
Тогда \bar{B} - событие, состоящее в том, что студент не сдал экзамен
Всего существует C_{100}^5 способов создания билетов по 5 вопросов из 100
возможных

Событию \bar{B} -студент не сдал экзамен благоприятствуют C_{100-6}^5 случаев, что
означает, что он не знал ни одного вопроса из 5 в билете

Вероятность того, что он не сдал экзамен, $P(\bar{B}) = \frac{C_{100}^5}{C_{94}^5}$

Вероятност того, что он знал хотя бы один вопрос в билете, и поэтому
сдал экзамен равн: $P(B) = 1 - P(\bar{B})$

Количество информации в сообщении о том, что студент сдал экзамен:
 $I = -\log_2 P(B)$

Количество вопросов	100	Количество вопросов	100
Сколько знает	6	Сколько знает	6
Сколько в 1 билете	5	Сколько в 1 билете	5
Всего исходов	=ЧИСЛКОМБ(В1;В3)	Всего исходов	7528752000
благоприя тствующих событию	=ЧИСЛКОМБ(В1-В2;В3)	благоприя тствующих событию В	54891018
противоположному В		вероятность события В	0,27
вероятность события В	=1-В5/В4	Количество информации	1,88
Количество информации	=-LOG(В6;2)		

Ответ:

Количество информации в сообщении о том, что студент сдал экзамен
составит 1,88 бит

Задача 3

Имеется два ящика. В каждом из которых по n шаров. В первом a белых, b красных, c синих. Во втором – по d шаров каждого цвета. Опыт состоит в вытаскивании по одному шару из каждого ящика. Каково количество информации содержащееся в сообщении об исходе опыта? Сравнить количество информации в первом и во втором опыте и сделать вывод

вариант	N	a	b	c	d
0	24	11	7	6	8

Пусть событие $A1$ -возможность достать белый шар из 1 ящика
 событие $B1$ -возможность достать красный шар из 1 ящика
 событие $C1$ -возможность достать синий шар из 1 ящика

$$P(A1) = \frac{11}{24} \quad P(B1) = \frac{7}{24} \quad P(C1) = \frac{6}{24}$$

Количество информации об извлечении какого-либо шара равно:

$$I1 = -(P(A1) \cdot \log_2 P(A1) + P(B1) \cdot \log_2 P(B1) + P(C1) \cdot \log_2 P(C1))$$

В то же время.:

Пусть событие $A2$ -возможность достать белый шар из 2 ящика
 событие $B2$ -возможность достать красный шар из 2 ящика
 событие $C2$ -возможность достать синий шар из 2 ящика

Но все возможности во втором ящике одинаково ожидаемы. Или
 равновероятны $P(A2) = P(B2) = P(C2) = \frac{8}{24}$

Поэтому Количество информации об извлечении какого-либо шара
 определяется фактически с помощью меры Хартли:

$$I2 = -3 \cdot \frac{8}{24} \log_2 P(A2) = -\log_2 P(A2)$$

	A	B	C	D	E
1		1 ящик		2 ящик	
2	$P(A1)=$	$=11/24$	$=\text{LOG}(B2;2)$	$I2=$	$=-\text{LOG}(1/3;2)$
3	$P(B1)=$	$=7/24$	$=\text{LOG}(B3;2)$		
4	$P(C1)=$	$=6/24$	$=\text{LOG}(B4;2)$		
5	$I1=$	$=\text{СУММПРОИЗВ}(B2:B4;C2:C4)$			

1 ящик			2 ящик	
$P(A1)=$	0,458	-1,126	$I2$	= 1,584963
$P(B1)=$	0,292	-1,778		
$P(C1)=$	0,250	-2,000		
$I1=$	1,5343372			

Вывод : Я научился(лась) вычислять количество информации в сообщении с помощью меры Шеннона и закона аддитивности информации и пришел к выводу, что мера Хартли всегда больше меры Шеннона

Лабораторная работа № 2

Тема: Избыточность сообщения. Эффективное кодирование по методу Шеннона – Фано

Цель: Уметь вычислять избыточность и эффективность кодов Сообщения

1	Я был разбужен спозаранку Щелчком оконного стекла.
2	Размокшей каменной баранкой В воде Венеция плыла.
3	Все было тихо, и, однако, Во сне я слышал крик, и он
4	Подобьем смолкнувшего знака Еще тревожил небосклон.
5	Он вис трезубцем Скорпиона Над гладью стихших мандолин
6	И женщиною оскорбленной, Быть может, издан был вдали.
7	Теперь он стих и черной вилкой Торчал по черенок.
8	Большой канал с косою ухмылкой Оглядывался, как беглец
9	Туда, голодные, противясь, Шли волны, шлендрая с тоски,
10	И гондолы рубили привязь, Точа о пристань тесаки.
11	Тихо льется тихий Дон, Желтый месяц входит в дом.
12	Входит в шапке набекрень - Видит желтый месяц тень.
13	Любить иных - тяжелый крест, А ты прекрасна без извилин,
14	И прелести твоей секрет Разгадке жизни равносильен.
15	Весною слышен шорох снов И шелест новостей и истин.
16	Ты из семьи таких основ. Твой смысл, как воздух, бескорыстен.
17	Легко проснуться и прозреть, Словесный сор из сердца вон.
18	И жить, не засоряясь впредь, Все это - не большая хитрость.
19	На меня наставлен сумрак ночи Тысячью биноклей на оси.
20	Если только можно, Авва Отче, Чашу эту мимо пронеси.
21	Я люблю твой замысел упрямый И играть согласен эту роль.
22	Но сейчас идет другая драма, И на этот раз меня уволь.
23	В ушах с утра какой-то шум. Он в памяти иль грезит?
24	С порога смотрит человек, Не узнавая дома.

Задание 1

Оценить избыточность сообщения.

Задание 2

Провести кодирование, сообщения. используя метод Шеннона – Фано.
Сравнить эффективности кодов

Образец выполнения задания для Варианта № 0

Задание 1

Вычислить длину сообщения:

Неясное становится еще более непонятным – 39 символов.

Выписать частоту появления каждого символа, а затем вычислить вероятность появления символа как частное от деления частоты на длину сообщения.

Частную энтропию вычислить по

формуле: $-p_i \log p_i$

Суммарное значение частных энтропий вычисляются по формуле:

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

Нулевая энтропия :

$H_0 = \log_2 m$, где m - количество символов в алфавите, используемом в данном сообщении

Избыточность:

$$r = 1 - \frac{H}{H_0}$$

Символы	частота	Вероятность P_i	Частная Энтропия
е	7	0,17949	0,4447777
н	6	0,15385	0,4154523
о	4	0,10256	0,3369643
пробел	4	0,10256	0,3369643
я	3	0,07692	0,2846492
с	3	0,07692	0,2846492
т	3	0,07692	0,2846492
а	1	0,02564	0,1355231
в	1	0,02564	0,1355231
и	1	0,02564	0,1355231
щ	1	0,02564	0,1355231
б	1	0,02564	0,1355231
л	1	0,02564	0,1355231
п	1	0,02564	0,1355231
ы	1	0,02564	0,1355231
м	1	0,02564	0,1355231
Суммарное значение частных энтропий H			3,6078145
Длина сообщения			39
Мощность алфавита m			16
"Нулевая" энтропия H_0			4
Избыточность r			9,80%

Задание 2

Провести кодирование, сообщения. используя метод Шеннона – Фано:
Неясное становится еще более непонятным

Кодирование сообщение осуществляется согласно алгоритму:

1. Буквы или символы исходного алфавита записывают в порядке убывающей вероятности.
2. Упорядоченное таким образом множество букв разбивают на подмножества так, чтобы суммарные вероятности этих подмножеств были примерно равны.
3. Всем знакам (буквам) верхней половины в качестве первого символа присваивают кодовый элемент 1, а всем нижним 0.
4. Затем каждое подмножество снова разбивается на два подмножества с соблюдением того же условия равенства вероятностей и с тем же условием присваивания кодовых элементов в качестве второго символа.
5. Такое разбиение продолжается до тех пор, пока в подмножестве не окажется только по одной букве кодируемого алфавита.
6. При каждом разбиении буквам верхнего подмножества присваивается кодовый элемент 1, а буквам нижнего подмножества - 0.

Среднее количество информации неравномерного кода вычисляется по

формуле: $Q_{ср} = \sum_{i=1}^m p_i q_i$,

где q_i – длина кодовой комбинации соответствующей i -той букве, а p_i – вероятность появления i - той буквы

Эффективность кода оценивается по формуле: $R = \frac{Q_{ср} - H}{H}$

$Q_{ср}$ принадлежит промежутку $[H_0, H]$ а эффективность показывает, насколько созданный неравномерный код близок к идеальному и имеющему энтропию H

Если эффективность превышает 10%, то код- неэффективный. Возможно разбиение сделано неверно

Символы	частота	Вероятность P _i	Кодовые последовательности													Код	Длина q _i	P _i Q _i	P _i LogP _i
			Номер разбиения																
			1		2		3		4		5		6		7				
е	7	0,1795	1	0,17949	1	0,17949	1									111	3	0,54	0,44478
н	6	0,1538	1	0,15385	1	0,15385	0									110	3	0,46	0,41545
о	4	0,1026	1	0,10256	0	0,10256	1									101	3	0,31	0,33696
пробел	4	0,1026	1	0,10256	0	0,10256	0									100	3	0,31	0,33696
я	3	0,0769	0	0,07692	1	0,07692	1									011	3	0,23	0,28465
с	3	0,0769	0	0,07692	1	0,07692	0	0,07692	1							0101	4	0,31	0,28465
т	3	0,0769	0	0,07692	1	0,07692	0	0,07692	0							0100	4	0,31	0,28465
а	1	0,0256	0	0,02564	0	0,02564	1	0,02564	1	0,02564	1	0,02564	1			001111	6	0,15	0,13552
в	1	0,0256	0	0,02564	0	0,02564	1	0,02564	1	0,02564	1	0,02564	0			001110	6	0,15	0,13552
и	1	0,0256	0	0,02564	0	0,02564	1	0,02564	1	0,02564	0	0,02564	1			001100	6	0,15	0,13552
щ	1	0,0256	0	0,02564	0	0,02564	1	0,02564	1	0,02564	0	0,02564	0			001100	6	0,15	0,13552
б	1	0,0256	0	0,02564	0	0,02564	0	0,02564	0	0,02564	1	0,02564	1			000011	6	0,15	0,13552
л	1	0,0256	0	0,02564	0	0,02564	0	0,02564	0	0,02564	1	0,02564	0			000010	6	0,15	0,13552
п	1	0,0256	0	0,02564	0	0,02564	0	0,02564	0	0,02564	0	0,02564	1			000001	6	0,15	0,13552
ы	1	0,0256	0	0,02564	0	0,02564	0	0,02564	0	0,02564	0	0,02564	0	0,02564	1	0000001	7	0,18	0,13552
м	1	0,0256	0	0,02564	0	0,02564	0	0,02564	0	0,02564	0	0,02564	0	0,02564	0	0000000	7	0,18	0,13552

39

I= 156 бит

Q _{ср}	3,90	3,61
Эффект. кодa R=	8%	
длина кодa	152	

Вывод: Полученный код является эффективным, так как средняя длина
кода на 8 % превышает энтропию
Приступаем к кодированию

		Н	е	я	с	н	о	е	пробел				
е	111	110	111	011	0101	110	101	111	100				
н	110	1101110110101110101111100											
о	101												
пр об ел	100												
я	011	с	т	а	н	о	в	и	т	с	я	пробел	
		0101	0100	001111	110	101	001110	001100	0100	0101	011	100	
с	0101	0101010000111111010100111000110001000101011100											
т	0100												
а	0011 11	е	щ	е	пробел	б	о	л	е	е	пробел		
в	0011 10	111	001100	111	100	000011	101	000010	111	111	100		
и	0011 00	111001100111100000011101000010111111100											
щ	0011 00												
б	0000 11	н	е	п	о	н	я	т	н	ы	м		
л	0000 10	110	111	000001	101	110	011	0100	110	0000001	0000000		
п	0000 01	110111000001101110011010011000000010000000											
ы	0000 001	Итоговый код											
м	0000 000	1101110110101110101111100010101000011111101010011100011000100010101110011100110011110000001110100001011111001101110000011011001101001100000010000000											

Лабораторная работа №3

Тема: Эффективное кодирование по методу Хаффмена.

Цель: Изучить методы эффективного статистического кодирования на примере эффективного кода Хаффмена.

Задание

Провести кодирование, сообщения. используя метод Хаффмена

Сообщения

1	Я был разбужен спозаранку Щелчком.	34
2	каменной баранкой В воде Венеция .	34
3	И, однако, Во сне он слышал крики.	34
4	С небес подобьем смолкнувшего знака	35
5	Он вис Над гладью стихших мандолин	34
6	Вскрик, быть может, издан был вдали	35
7	И черной вилкой Торчал по черенок	34
8	Большой канал - с косою ухмылкой.	33
9	Туда, голодные, Шли волны от тоски	34
10	И гондолы рубили привязь о тесаки.	34
11	Льетса Дон тихо месяц входит в дом	34
12	Входит в шапке набекрень Видит тень	35
13	Любить тяжелый крест а ты прекрасна	35
14	Секрет - разгадке жизни равносильен	34
15	Весною слышен шорох и снов и истин.	35
16	Твой смысл, как воздух, бескорыстен	35
17	Легко проснуться и прозреть от снов	35
18	Не засоряясь, есть большая хитрость	35
19	сумрак ночи Тысячью биноклей на оси	35
20	Авва Отче, Чашу эту мимо пронеси.	33
21	Люблю Твой замысел играть эту роль.	35
22	Но, на этот раз идет другая драма -	35
23	В ушах с утра какой-то шум и греза?	35
24	И смотрит человек, Не узнавая дома,	35

Сделать вывод

Образец выполнения задания для Варианта № 0

Провести кодирование, сообщения. используя метод Хаффмена

Оценить эффективность кодирования

Вычислить длину сообщения: Неясное становится еще более не понятным – 39 символов. Мощность алфавита : 16 символов. Считая все символы равновероятными количество информации, приходящееся на один символ сообщения: $i = \log_2 16 = 4$ бит. Длина сообщения при таком кодировании : 156 бит.

Приступаем к сжатию сообщения.

Выписать частоту появления каждого символа, а затем вычислить вероятность появления символа как частное от деления частоты на длину сообщения.

Частную энтропию вычислить по формуле: $-p_i \log p_i$

Суммарное значение частных энтропий вычисляют по формуле:

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

Алгоритм сжатия Хаффмена

1. Буквы алфавита сообщений выписывают в основной столбец таблицы в порядке убывания вероятностей.
2. Две последние буквы объединяют в одну вспомогательную букву, которой приписывают суммарную вероятность.
3. Вероятности букв, не участвующих в объединении и полученная суммарная вероятность снова располагаются в порядке убывания вероятностей в вспомогательном столбце, а две последние объединяются.
4. Процесс продолжается до тех пор, пока не получат единственную вспомогательную букву с вероятностью, равной единице.
5. Далее для создания кода нужно построить **кодвое дерево**
6. Из точки, соответствующей вероятности 1, направляем две ветви, соответствующие вероятностям из которых была получена эта сумма, причем ветви с большей вероятностью направляем влево и присваиваем символ 1, а с меньшей направляем вправо и присваиваем символ 0.
7. Такое последовательное ветвление продолжаем до тех пор, пока не дойдем до вероятности каждой буквы.
8. Двигаясь по кодовому дереву сверху вниз, можно записать для каждой буквы соответствующую ей кодовую комбинацию.

энтропия.xlsx - Microsoft Excel

Файл Главная Вставка Разметка страницы Формулы Данные Рецензирование Вид Разработчик

Вставить Шрифт Выравнивание Число

Условное форматирование Форматировать как таблицу Стиль ячеек Вставить Удалить Формат Ячейки

Сортировка и фильтр Найти и выделить Редактирование

P14

1	Неясное становится еще более непонятным															Кодовая комбинация	qi	qi*pi	Hi					
2	Символы	частота	Вероятность Pi	Вспомогательные столбцы																				
3				1	2	3	4	5	6	7	8	9	10	11	12	13	14	15						
5	е	7	0,1795	0,1795	0,1795	0,1795	0,1795	0,1795	0,1795	0,1795	0,1795	0,1795	0,1795	0,2051	0,2821	0,3333	0,3846	0,6154	1,0000	00	2	0,3590	0,4448	
6	н	6	0,1538	0,1538	0,1538	0,1538	0,1538	0,1538	0,1538	0,1538	0,1538	0,1538	0,1795	0,1795	0,2051	0,2821	0,3333	0,3846		110	3	0,4615	0,4155	
7	о	4	0,1026	0,1026	0,1026	0,1026	0,1026	0,1026	0,1026	0,1026	0,1026	0,1026	0,1282	0,1538	0,1538	0,1795	0,1795	0,2051	0,2821		011	3	0,3077	0,3370
8	пробел	4	0,1026	0,1026	0,1026	0,1026	0,1026	0,1026	0,1026	0,1026	0,1026	0,1026	0,1026	0,1282	0,1538	0,1538	0,1795	0,1795			010	3	0,3077	0,3370
9	я	3	0,0769	0,0769	0,0769	0,0769	0,0769	0,0769	0,0769	0,0769	0,0769	0,0769	0,1026	0,1026	0,1026	0,1282	0,1538	0,1538			1110	4	0,3077	0,2846
10	с	3	0,0769	0,0769	0,0769	0,0769	0,0769	0,0769	0,0769	0,0769	0,0769	0,0769	0,1026	0,1026	0,1026	0,1282					1011	4	0,3077	0,2846
11	т	3	0,0769	0,0769	0,0769	0,0769	0,0769	0,0769	0,0769	0,0769	0,0769	0,0769	0,1026	0,1026	0,1026	0,1282					1010	4	0,3077	0,2846
12	а	1	0,0256	0,0513	0,0513	0,0513	0,0513	0,0513	0,0769	0,0769	0,0769	0,0769									10010	5	0,1282	0,1355
13	в	1	0,0256	0,0256	0,0513	0,0513	0,0513	0,0513	0,0769	0,0769	0,0769										100111	6	0,1538	0,1355
14	и	1	0,0256	0,0256	0,0256	0,0513	0,0513	0,0513	0,0513												100110	6	0,1538	0,1355
15	щ	1	0,0256	0,0256	0,0256	0,0256	0,0513	0,0513													111101	6	0,1538	0,1355
16	б	1	0,0256	0,0256	0,0256	0,0256	0,0256	0,0256													111100	6	0,1538	0,1355
17	л	1	0,0256	0,0256	0,0256	0,0256															111111	6	0,1538	0,1355
18	п	1	0,0256	0,0256	0,0256																111110	6	0,1538	0,1355
19	ы	1	0,0256	0,0256																	10001	5	0,1282	0,1355
20	м	1	0,0256																		10000	5	0,1282	0,1355
21		39	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	средняя длина кода	3,667	3,608	
22																					эффективность	1,63%		
23																					Длина сообщения	143		

метод Хаффмена (пример_лаб) (3) неясное Хаффмен неясное Ш_Ф метод Хаффмена_1 неясное пустой неясное метод Хаффмена (пример_лаб) (2)

Готово Вычислить 87%

е	00		Н	е	я	с	н	о	е	пробел		длина кода	143		
н	110		110	00	1110	1011	110	011	00	010					
о	011		110001110101111001100010												
прб л	010														
я	1110		с	т	а	н	о	в	и	т	с	я	пробел		
с	1011		1011	1010	10010	110	011	100111	100110	1010	1011	1110	011		
т	1010		1011101010010110011100111100110101010111110011												
а	10010														
в	100111		е	щ	е	пробел	б	о	л	е	е	пробел			
и	100110		00	111101	00	010	111100	011	111111	00	00	011			
щ	111101		00111101000101111000111111110000011												
б	111100														
л	111111		н	е	п	о	н	я	т	н	ы	м			
п	111110		110	00	111110	011	110	1110	1010	110	10001	10000			
ы	10001		11000111110011110111010101101000110000												
м	10000														
		143	1100011101011110011000101011101010010110011100111100110101010111110011001111010001011110001111111100000111100011111001111011101010101000110000												

Вывод В результате кодирования сообщения по методу Хаффмена был создан неравномерный префиксный код однозначно кодирующий-декодированный сообщение с эффективностью 1,63%, при этом длина сообщения была снижена с 156 бит до 143 бита.

Лабораторная работа №4

Тема: Код Хемминга

Цель:

Освоить компетенции

ПК 1.3 Обеспечивать защиту информации в сети с использованием программно-аппаратных средств.

ПК 2.1. Администрировать локальные вычислительные сети и принимать меры по устранению возможных сбоев.

ПК 3.2. Проводить профилактические работы на объектах сетевой инфраструктуры и рабочих станциях

Задание

Передать сообщение без ошибок, приняв длину информационного слова равной 16

0	habr	8	home	15	play	23	look
1	more	9	best	16	nail	24	like
2	many	10	what	17	zoom	25	side
3	then	11	nice	18	page	26	from
4	auto	12	ctrl	19	next	27	lamp
5	disk	13	lock	20	mute	28	plus
6	meat	14	wake	21	chair	29	moon
7	past			22	cash		

Образец выполнения задания для Варианта № 0

Представить сообщение в бинарном виде (выписать коды, соответствующие строч

символ	10-Б код	2-Б код	символ	10-Б код	2-Б код	символ	10-Б код	2-Б код	символ	10-Б код	2-Б код
	32	00100000	8	56	00111000	P	80	01010000	h	104	01101000
!	33	00100001	9	57	00111001	Q	81	01010001	i	105	01101001
"	34	00100010	:	58	00111010	R	82	01010010	j	106	01101010
#	35	00100011	;	59	00111011	S	83	01010011	k	107	01101011
\$	36	00100100	<	60	00111100	T	84	01010100	l	108	01101100
%	37	00100101	=	61	00111101	U	85	01010101	m	109	01101101
&	38	00100110	>	62	00111110	V	86	01010110	n	110	01101110
'	39	00100111	?	63	00111111	W	87	01010111	o	111	01101111
(40	00101000	@	64	01000000	X	88	01011000	p	112	01110000
)	41	00101001	A	65	01000001	Y	89	01011001	q	113	01110001
*	42	00101010	B	66	01000010	Z	90	01011010	r	114	01110010
+	43	00101011	C	67	01000011	[91	01011011	s	115	01110011
,	44	00101100	D	68	01000100	\	92	01011100	t	116	01110100
-	45	00101101	E	69	01000101]	93	01011101	u	117	01110101
.	46	00101110	F	70	01000110	^	94	01011110	v	118	01110110
/	47	00101111	G	71	01000111	_	95	01011111	w	119	01110111
0	48	00110000	H	72	01001000	`	96	01100000	x	120	01111000
1	49	00110001	I	73	01001001	a	97	01100001	y	121	01111001
2	50	00110010	J	74	01001010	b	98	01100010	z	122	01111010
3	51	00110011	K	75	01001011	c	99	01100011	{	123	01111011
4	52	00110100	L	76	01001100	d	100	01100100		124	01111100
5	53	00110101	M	77	01001101	e	101	01100101	}	125	01111101
6	54	00110110	N	78	01001110	f	102	01100110	~	126	01111110
7	55	00110111	O	79	01001111	g	103	01100111	□	127	01111111

НЫМ

буквам из таблицы ASCII)

Символ	ASCII код	Бинарное представление
h	104	01101000
a	97	01100001
b	98	01100010
r	114	01110010

Разбить сообщение на блоки по 16 бит:

h	a
01101000	01100001

и

b	r
01100010	01110010

(Здесь рассматривается алгоритм работы на примере II блока – **br**)

Часть 1

1. необходимо вставить контрольные биты. Они вставляются в строго определённых местах — это позиции с номерами, равными степеням двойки. При длине информационного слова в 16 бит это будут позиции 1, 2, 4, 8, 16.

Следовательно, получится 5 контрольных бит. Таким образом, длина всего сообщения увеличилась на 5 бит. До вычисления самих контрольных бит, им присваивают значение «0».

было

b	r
01100010	01110010

стало

b	r
000011000010	01110010

символ	10-Б код	2-Б код	символ	10-Б код	2-Б код
P	80	01010000	h	104	01101000
Q	81	01010001	i	105	01101001
R	82	01010010	j	106	01101010
S	83	01010011	k	107	01101011
T	84	01010100	l	108	01101100
U	85	01010101	m	109	01101101
V	86	01010110	n	110	01101110
W	87	01010111	o	111	01101111
X	88	01011000	p	112	01110000
Y	89	01011001	q	113	01110001
Z	90	01011010	r	114	01110010
[91	01011011	s	115	01110011
\	92	01011100	t	116	01110100
]	93	01011101	u	117	01110101
^	94	01011110	v	118	01110110
_	95	01011111	w	119	01110111
`	96	01100000	x	120	01111000
a	97	01100001	y	121	01111001
b	98	01100010	z	122	01111010
c	99	01100011	{	123	01111011
d	100	01100100		124	01111100
e	101	01100101	}	125	01111101
f	102	01100110	~	126	01111110
g	103	01100111	□	127	01111111

2. Начать заполнение таблицы размерности 22×8 (Первая строка- номера битов 1-21.Вторая строка - сами биты. Разместить в ней 21 бит, таким образом, чтобы 1,2,4,8, 16 ячейки заполняли контрольные биты, а между ними распределить последовательно информационные биты)
3. Отметить цветом и знаком «X» те биты, которые контролирует контрольный бит, номер которого справа, подчиняясь следующему правилу: контрольный бит с номером N контролирует все последующие N бит через

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	
0	0	0	0	1	1	0	0	0	0	1	0	0	1	1	0	1	0	0	1	0	
x		x		x		x		x		x		x		x		x		x		x	1
	x	x			x	x			x	x			x	x			x	x			2
			x	x	x	x					x	x	x	x					x	x	4
							x	x	x	x	x	x	x	x							8
															x	x	x	x	x	x	16
0	0	0	1	1	1	0	1	0	0	1	0	0	1	1	0	1	0	0	1	0	

каждые N бит, начиная с позиции N (см таблицу):

4. Вычислить контрольные биты. Для определения значения каждого контрольного бита необходимо :
 1. Последовательно зафиксировать каждый контрольный бит и сосчитать среди контролируемых им битов количество единиц
 2. Если сосчитанное число – четное контрольный бит равен -0, если нечетное – контрольный бит равен 1
5. Записать вычисленные значения в 8 строку таблицы и выписать **отдельно** полученный код

Часть II Декодирование и исправление ошибок

1. Ввести 1 ошибку в код – например – в 17 (выбирать любой не контрольный, кроме 17!!!)бите вместо 1 получен при передаче сообщения - 0 (01100010 01100010) фактически это значит, что вместо br будет получено bb (значение ошибочного символа подобрать из таблицы)

- Вновь создать таблицу размерности 22×11 и начать ее заполнение вновь приняв значения контрольных бит равными 0. А ошибочный бит дважды выделить.
- Отметить контролируемые биты и вновь вычислить значения контрольных бит (см таблицу 8 строку)
- В 9 строку ввести последовательно исходный код, полученный в первой части
- Сложить по модулю 2 оба кода.
- Выписать и сложить номера информационных бит, в которых получены 1. В данном случае 1+16=17. Полученное значение указывает на бит, в котором произошла ошибка

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	
0	0	0	0	1	1	0	0	0	0	1	0	0	1	1	0	0	0	0	1	0	0
x		x		x		x		x		x		x		x		x		x		x	1
	x	x			x	x			x	x			x	x			x	x			2
			x	x	x	x					x	x	x	x					x	x	4
							x	x	x	x	x	x	x	x							8
															x	x	x	x	x	x	16
1	0	0	1	1	1	0	1	0	0	1	0	0	1	1	1	0	0	0	1	0	
0	0	0	1	1	1	0	1	0	0	1	0	0	1	1	0	1	0	0	1	0	
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	Сложение по модулю 2
		0		1	1	0		0	0	1	0	0	1	1		1	0	0	1	0	

01100010

b

01110010

r

17

Позиция ошибки

- Заключительный этап: Инвертировать ошибочный бит и отбросить контрольные биты. Будет получено исходное сообщение в первоизданном виде -br. Подписать под кодами соответствующие значения символов из таблицы.
- Аналогично кодировать – декодировать 2 блок.

9.

Вывод:

передаваемое сообщение – $habr$, в каждом блоке получило одиночную ошибку и было преобразовано в сообщение $fabb$.

После преобразования по методу Хемминга была обнаружена и исправлена ошибка. Полученное сообщение- $habr$.