

Документ подписан простой электронной подписью
Информация о владельце:
ФИО: Локтионова Оксана Геннадьевна
Должность: проректор по учебной работе
Дата подписания: 16.12.2020 18:58:32
Уникальный программный ключ:
0b817ca911e6668abb13a5d426d39e5f1c11eabbf73e943df4a4851fda56d089

МИНОБРНАУКИ РОССИИ
Федеральное государственное бюджетное образовательное
учреждение высшего образования
«Юго-Западный государственный университет»
(ЮЗГУ)

Кафедра космического приборостроения и систем связи

УТВЕРЖДАЮ
Проректор по учебной работе
О.Г. Локтионова
«15» 12 2017 г.



**РАСЧЁТ КОЛИЧЕСТВА ИНФОРМАЦИИ И ЭНТРОПИИ
ДЛЯ ДИСКРЕТНЫХ СООБЩЕНИЙ**

Методические указания по выполнению лабораторной работы №3
для студентов, обучающихся по направлению подготовки
11.03.02 «Инфокоммуникационные технологии и системы связи»
по дисциплине: «Основы теории информации»

Курск 2017

УДК 621.391

Составители: А.В. Хмелевская, А.Н. Шевцов

Рецензент

Доктор технических наук, старший научный сотрудник,
профессор кафедры *В.Г. Андронов*

Расчёт количества информации и энтропии для дискретных сообщений: методические указания по выполнению лабораторной работы №3 по дисциплине: «Основы теории информации» / Юго-Зап. гос. ун-т; сост.: А.В. Хмелевская, А.Н. Шевцов. – Курск, 2017. – 11 с.: табл. 5. – Библиогр.: с. 11.

Методические указания по выполнению лабораторной работы содержат краткие теоретические сведения о способах расчета энтропии и количества информации для дискретных сообщений, варианты заданий для выполнения работы, а также перечень вопросов для самоконтроля.

Методические указания полностью соответствуют требованиям типовой программы, утвержденной УМО по направлению подготовки 11.03.02 «Инфокоммуникационные технологии и системы связи», а также рабочей программе дисциплины: «Основы теории информации».

Предназначены для студентов, обучающихся по направлению подготовки 11.03.02, очной формы обучения.

Текст печатается в авторской редакции

Подписано в печать *15.12.17*. Формат 60x841/16.
Усл. печ. л. *0,64*. Уч.-изд. л. *0,528* Тираж 100 экз. Заказ *3273* Бесплатно
Юго-Западный государственный университет.
305040, г. Курск, ул. 50 лет Октября, 94

1 Цель работы

- приобрести умение определять количество информации и энтропию для дискретного источника сообщений.

2 Основные теоретические сведения

Источник сообщений, который может в каждый момент времени случайным образом принять одно из конечного множества возможных состояний, называют **дискретным источником сообщений**. Каждому состоянию источника и ставится в соответствие условное обозначение в виде символа из алфавита этого источника: S_1, S_2, \dots, S_m .

Источник сообщений в общем случае характеризуется ансамблем S , то есть полной совокупностью состояний с вероятностями их появления, составляющими в сумме единицу:

$$S = \left(\begin{array}{cccc} s_1 & s_2 & \dots & s_m \\ p(s_1) & p(s_2) & \dots & p(s_m) \end{array} \right); \text{ причём } \sum_{i=1}^m p(s_i) = 1$$

Энтропия $H(S)$ дискретного источника сообщений, задаваемого ансамблем S , может быть определена по формуле Шеннона:

$$H(S) = - \sum_{i=1}^m p(s_i) \log_2 p(s_i) \text{ (бит/символ);}$$

где $p(s_j)$ - вероятность появления символа s_j из алфавита объемом m .

Количество информации в сообщении, состоящем из n символов, может быть найдено следующим образом:

$$I = n \cdot H(S) = -n \sum_{i=1}^m p(s_i) \log_2 p(s_i) \text{ (бит)}$$

При наличии взаимозависимости символов алфавита используется **условная энтропия**. С учетом статистической связи между парами символов условная энтропия может быть найдена из формулы:

$$H_2(S) = - \sum_{i=1}^m \sum_{j=1}^m p(s_i) p(s_j | s_i) \log_2 p(s_j | s_i) \text{ (бит/символ);}$$

где $p(S_j | s_i)$ - условная вероятность появления символа S_j при заданном символе S_i .

Другой важной характеристикой источника сообщений, помимо энтропии, является *информационная избыточность*,

которая показывает относительную недогруженность на символ алфавита источника сообщений:

$$R = \frac{H_{\max} - H}{H_{\max}} = 1 - \frac{H}{H_{\max}};$$

где $H_{\max} = \log_2 m$ - максимально возможная для данного источника сообщений энтропия, которая достигается при равной вероятности появления символов.

Кроме общей информационной избыточности существуют частные избыточности, основными из которых являются:

1) Избыточность, вызванная статистической связью между символами s_1, s_2, \dots, s_m :

$$R_p = 1 - \frac{H}{H'};$$

где $H' = -\sum_{i=1}^m p(s_i) \log_2 p(s_i)$ - энтропия источника сообщений при отсутствии статистической связи между символами.

2) Избыточность, вызванная не экстремальным распределением символов s_1, s_2, s_m (не равновероятным появлением символов в сообщениях):

$$R_\varphi = 1 - \frac{H'}{H_{\max}}.$$

3) Общая информационная избыточность связана с частными избыточностями следующим образом:

$$R = R_p + R_\varphi - R_p R_\varphi.$$

3 Порядок выполнения работы

Данная практическая работа предполагает выполнение следующих этапов:

- 1) Изучить методические указания к практической работе.
- 2) Пройти собеседование с преподавателем и получить задание для выполнения работы.
- 3) Определить энтропию дискретного источника сообщений, алфавит которого состоит из восьми независимых символов S_1, S_2, \dots, S_8 . Известны вероятности появления символов $p(s_1)=p_1, p(s_2)=p_2, \dots, p(s_8)=p_8$ (таблица 1).
- 4) Определить среднее и частное количество информации, содержащейся в сообщении, которое состоит из ваших фамилии, имени и отчества (между словами сообщения допускаются пробелы). Вероятности появления букв русского алфавита взять из табл. П.1. Сравнить полученные результаты.
- 5) В условии задачи, сформулированной в пункте 3, учесть зависимость между парами символов, которая задана матрицей условных вероятностей $p(s_j|s_i)$ (табл. 2 – 5).
- 6) Для задачи, сформулированной в пункте 4, учесть наличие статистической связи между парами и тройками букв.
- 7) Определить частные избыточности, вызванные наличием связи между символами и не экстремальным распределением символов, для полученных при выполнении пунктов 5 и 6 результатов. На основе частных избыточностей найти общую информационную избыточность источника сообщений.
- 8) Оформить и защитить отчет по лабораторной работе.

4 Варианты заданий

Номер варианта студента определяется как порядковый номер студента в журнале преподавателя.

Таблица 1 – Вероятности появления независимых символов s_1, s_2, \dots, s_8

№ вар.	Вероятности для символов s_1, s_2, \dots, s_8							
	P1	P2	P3	P4	P5	P6	P7	P8
1	0,21	0,12	0,05	0,08	0,03	0,14	0,2	0,17
2	0,13	0,04	0,19	0,16	0,18	0,12	0,13	0,05
3	0,2	0,15	0,16	0,04	0,19	0,05	0,12	0,09
4	0,14	0,1	0,04	0,17	0,12	0,16	0,05	0,22
5	0,06	0,21	0,03	0,23	0,13	0,12	0,14	0,08
6	0,08	0,19	0,22	0,2	0,09	0,03	0,17	0,02
7	0,09	0,18	0,06	0,12	0,15	0,1	0,07	0,23
8	0,22	0,13	0,12	0,15	0,07	0,04	0,19	0,08
9	0,1	0,24	0,08	0,07	0,17	0,05	0,16	0,13
10	0,04	0,17	0,15	0,11	0,1	0,09	0,13	0,21
11	0,17	0,05	0,16	0,22	0,08	0,19	0,09	0,04
12	0,12	0,09	0,08	0,14	0,16	0,21	0,04	0,16
13	0,23	0,1	0,07	0,06	0,05	0,12	0,22	0,15
14	0,16	0,08	0,05	0,18	0,11	0,15	0,07	0,2
15	0,24	0,12	0,13	0,09	0,14	0,08	0,15	0,05
16	0,11	0,04	0,19	0,16	0,18	0,12	0,13	0,07
17	0,09	0,23	0,05	0,21	0,15	0,1	0,12	0,05
18	0,2	0,15	0,1	0,14	0,08	0,06	0,17	0,1
19	0,15	0,06	0,16	0,23	0,07	0,19	0,08	0,06
20	0,22	0,12	0,06	0,07	0,04	0,13	0,2	0,16
21	0,13	0,08	0,1	0,09	0,21	0,19	0,05	0,15
22	0,08	0,2	0,07	0,13	0,14	0,09	0,05	0,24

Таблица 2 – Матрица условных вероятностей для вариантов 1, 5, 9, 13, 17, 21

i,j	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8
S_1	0,14	0,22	0	0,27	0,09	0	0,12	0,16
S_2	0,25	0	0,13	0,08	0,2	0,13	0,21	0
S_3	0	0,26	0,16	0,22	0,12	0,09	0	0,15
S_4	0,07	0,18	0,24	0	0,29	0	0,1	0,12
S_5	0,09	0,3	0,08	0,21	0	0,17	0	0,15
S_6	0	0,14	0,15	0	0,28	0,07	0,23	0,13
S_7	0,15	0,06	0	0,27	0,24	0,08	0,2	0
S_8	0,23	0	0,16	0,1	0	0,19	0,07	0,25

Таблица 3 – Матрица условных вероятностей для вариантов 2, 6, 10, 14, 18, 22

h_i	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8
S_1	0	0,17	0,28	0	0,13	0,08	0,23	0,11
S_2	0,13	0,05	0	0,26	0,19	0,13	0	0,24
S_3	0,14	0	0,1	0,15	0,29	0	0,07	0,25
S_4	0,07	0,18	0,24	0	0,28	0	0,1	0,13
S_5	0,26	0	0,19	0,08	0	0,1	0,28	0,09
S_6	0,06	0,24	0	0,17	0,09	0,33	0	0,11
S_7	0	0,08	0,15	0,27	0,22	0,08	0,2	0
S_8	0,09	0,25	0,14	0,16	0	0,21	0,15	0

Таблица 4 – Матрица условных вероятностей для вариантов 3, 7, 11, 15, 19, 23

i,j	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8
S_1	0,24	0,2	0	0,15	0,22	0	0,14	0,05
S_2	0,15	0	0,08	0,14	0,19	0,23	0,21	0
S_3	0	0,14	0,27	0,24	0,15	0,11	0	0,09
S_4	0,09	0,13	0,28	0	0,25	0	0,15	0,1
S_5	0,11	0,22	0,2	0,07	0	0,27	0	0,13
S_6	0	0,06	0,09	0	0,2	0,15	0,29	0,21
S_7	0,1	0,25	0	0,08	0,14	0,17	0,26	0
S_8	0,06	0	0,3	0,14	0	0,23	0,07	0,2

Таблица 5 – Матрица условных вероятностей для вариантов 4, 8, 12, 16, 20, 24

i,j	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8
S_1	0	0,18	0,07	0	0,3	0,12	0,13	0,2
S_2	0,29	0,11	0	0,08	0,21	0,17	0	0,14
S_3	0,06	0	0,15	0,26	0,18	0	0,07	0,28
S_4	0,14	0,09	0,3	0	0,22	0	0,19	0,06
S_5	0,2	0	0,27	0,07	0	0,1	0,21	0,15
S_6	0,15	0,21	0	0,06	0,11	0,29	0	0,18
S_7	0	0,12	0,28	0,15	0,09	0,2	0,16	0
S_8	0,1	0,07	0,18	0,29	0	0,21	0,15	0

5 Требования к оформлению отчёта

Лабораторная работа рассчитана на 4 часа у очной формы обучения направления подготовки 11.03.02. Выполняется во 2й контрольной точке.

Отчёт по лабораторной работе должен содержать:

- цель работы;
- краткие теоретические сведения;
- индивидуальное задание в соответствии с номером варианта;
- порядок выполнения работы;
- подробные расчеты с пояснениями;
- вывод о проделанной работе с анализом полученных результатов;
- ответы на контрольные вопросы.

Результаты различных измерений необходимо представить в виде нескольких самостоятельных таблиц и графиков. Каждая таблица и каждый график должны иметь свой заголовок и исходные данные эксперимента.

При выполнении численных расчетов надо записать формулу определяемой величины, сделать соответствующую численную подстановку и произвести вычисления.

Минимальный балл за лабораторную работу составляет 0.5 балла (выполнил работу, но не защитил). Максимальный балл – 3 (выполнил работу и защитил без замечаний).

Примерные критерии оценки качества отчётов по лабораторной работе:

- оформление отчёта не соответствует предъявляемым требованиям – минус 0,5 балла;
- полученные экспериментальные материалы не обработаны (осциллограммы, спектрограммы и т. п.) – минус 0.5 балла;
- выводы не соответствуют результатам работы – минус 0,5 балла;
- работа защищена не вовремя (после окончания 2й контрольной точки) – минус 0.5 балла.

6 Контрольные вопросы

- 1) Что называют дискретным источником сообщений?
- 2) Что характеризует энтропия сообщения, и каким образом она может быть определена для дискретного источника сообщений?
- 3) В каком случае энтропия дискретного источника сообщений принимает максимальное значение?
- 4) Как можно найти количество информации, содержащееся в дискретном сообщении?
- 5) Как определяется условная энтропия источника сообщений при наличии статистической связи между состояниями этого источника?
- 6) Что характеризует информационная избыточность источника сообщений?
- 7) Какие выделяют частные виды информационной избыточности и как они связаны с общей избыточностью источника сообщений?

7 Список использованных источников

1. Котоусов А.С. Теория информации. Учебное пособие для вузов. – М.: Радио и связь, 2003.
2. Липкин И.А. Статистическая радиотехника. Теория информации и кодирования. – М.: Вузовская книга, 2002.
3. Гоноровский И.С. Радиотехнические цепи и сигналы. – М.: Советское радио, 1986.
4. Темников Ф.Е. и др. Теоретические основы информационной техники. – М.: Энергия, 1979.