Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования   
**Национальный исследовательский университет «МЭИ»**

**Лабораторная работа №4**

по курсу«**Методы и средства защиты информации**»

**«Кодирование, информации методом Шеннона и Хаффмана»**

Выполнил:

студент группы А-07-18

Востриков Р.В.

Вариант 8

Москва  
2020

**1. Задание**

Задание 1

Найти префиксные коды с указанными ниже длинами слов:

а) l1 = l2 = 2; l3 = l4 = 3; l5 =l6 =l7 =4;

б) l1 = 1; l2 = 2; l3 = l4 = l5 = l6 = 4.

Задание 2

Закодировать двоичным кодом Хаффмана:

а) множество сообщений, имеющих вероятности: p1 = 0.25, p2 = 0.20, p3 = p4 = p5 = 0.15, p6 = 0.10;

б)

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Буква | Z1 | Z2 | Z3 | Z4 | Z5 | Z6 | Z7 | Z8 |
| Вероятность | 0.22 | 0.20 | 0.16 | 0.16 | 0.10 | 0.10 | 0.04 | 0.02 |

Задание 3  
Составить программу для кодирования и декодирования слов из букв алфавита кодовой таблицы задания 2.

**2. Описание результатов**

Задание 1

Для указанных длин слов были найдены префиксные коды:

а) x1 = 00, x2 = 01, x3 = 100, x4 = 101, x5 = 1100, x6 = 1110, x7 = 1111

б) x1 = 0, x2 = 10, x3 = 1100, x4 = 1101, x5 = 1110, x6 = 1111

Задание 2

**а)**

0

ABCDEF

ACD

BEF

CD

A

C

D

EF

B

E

F

0

0

0

0

1

1

1

1

1

Пусть алфавит будет X = { A, B, C, D, E, F }, а набор весов W = { 0.25, 0.20, 0.15, 0.15, 0.15, 0.10 }.

В дереве Хаффмана будет 6 узлов:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Узел | A | B | C | D | E | F |
| Вес | 0.25 | 0.20 | 0.15 | 0.15 | 0.15 | 0.10 |

Объединим в один узел два минимальных по весу узла E и F:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Узел | A | EF | B | C | D |
| Вес | 0.25 | 0.25 | 0.20 | 0.15 | 0.15 |

Затем опять объединим в один узел два минимальных по весу узла C и D:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Узел | CD | A | EF | B |
| Вес | 0.30 | 0.25 | 0.25 | 0.20 |

Ещё раз повторим эту же операцию для узлов EF и B:

|  |  |  |  |
| --- | --- | --- | --- |
| Узел | BEF | CD | A |
| Вес | 0.45 | 0.30 | 0.25 |

Ещё раз повторим эту же операцию для узлов CD и A:

|  |  |  |
| --- | --- | --- |
| Узел | ACD | BEF |
| Вес | 0.55 | 0.45 |

На последнем шаге объединм два узла ACD и BEF:

|  |  |
| --- | --- |
| Узел | ABCDEF |
| Вес | 1 |

Остался один узел, значит, мы пришли к корню дерева Хаффмана (смотри рисунок). Теперь для каждого символа выберем кодовое слово (бинарная последовательность, обозначающая путь по дереву к этому символу от корня):

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Узел | A | B | C | D | E | F |
| Код | 01 | 10 | 000 | 001 | 110 | 111 |

**б)**

0

ABCDEFGH

AB

CDEFGH

A

B

EFGH

CD

E

FGH

0

0

0

1

1

1

1

C

D

0

1

F

GH

0

1

G

H

0

1

Пусть алфавит будет X = { A, B, C, D, E, F, G, H },

а набор весов W = { 0.22, 0.20, 0.16, 0.16, 0.10, 0.10, 0.04, 0.02 }.

В дереве Хаффмана будет 8 узлов:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Узел | A | B | C | D | E | F | G | H |
| Вес | 0.22 | 0.20 | 0.16 | 0.16 | 0.10 | 0.10 | 0.04 | 0.02 |

Объединим в один узел два минимальных по весу узла G и H:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Узел | A | B | C | D | E | F | GH |
| Вес | 0.22 | 0.20 | 0.16 | 0.16 | 0.10 | 0.10 | 0.06 |

Затем опять объединим в один узел два минимальных по весу узла F и GH:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Узел | A | B | C | D | FGH | E |
| Вес | 0.22 | 0.20 | 0.16 | 0.16 | 0.16 | 0.10 |

Ещё раз повторим эту же операцию для узлов FGHи E:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Узел | EFGH | A | B | C | D |
| Вес | 0.26 | 0.22 | 0.20 | 0.16 | 0.16 |

Ещё раз повторим эту же операцию для узлов C и D:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Узел | CD | EFGH | A | B |
| Вес | 0.32 | 0.26 | 0.22 | 0.20 |

Ещё раз повторим эту же операцию для узлов A и B:

|  |  |  |  |
| --- | --- | --- | --- |
| Узел | AB | CD | EFGH |
| Вес | 0.42 | 0.32 | 0.26 |

Ещё раз повторим эту же операцию для узлов CD и EFGH:

|  |  |  |
| --- | --- | --- |
| Узел | CDEFGH | AB |
| Вес | 0.58 | 0.42 |

На последнем шаге объединм два узла CDEFGHи AB:

|  |  |
| --- | --- |
| Узел | ABCDEFGH |
| Вес | 1 |

Остался один узел, значит, мы пришли к корню дерева Хаффмана (смотри рисунок). Теперь для каждого символа выберем кодовое слово (бинарная последовательность, обозначающая путь по дереву к этому символу от корня):

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Узел | A | B | C | D | E | F | G | H |
| Код | 00 | 01 | 100 | 101 | 110 | 1110 | 11110 | 11111 |

**3. Код программы (к заданию 3)**

# таблица кодовых слов

codebook = {'A': '00', 'B': '01', 'C': '100', 'D': '101', 'E': '110', 'F': '1110', 'G': '11110', 'H': '11111'}

def encode(msg, codes):

encoded = [codes[l] for l in msg] # составляем набор кодовых слов, соответствующий сообщению

return ''.join(encoded)

def decode(msg, codes):

import re # библиотека для работы с РВ

# приводим таблицу кодовых слов к удобному для декодирования виду

letters = dict(zip(codes.values(), codes.keys()))

p = '|'.join(letters.keys()) # создаем РВ из кодовых слов

result = ''

while msg:

m = re.match(p, msg) # получаем следующее кодовое слово

result += letters[m.group()] # добавляем букву, соответствующую кодовому слову к результату

msg = msg[len(m.group()):]

return result

**4. Вывод**

В лабораторной работе было получен префиксный код с кодововыми словами указанной длины. Продемонстрировано получение префиксного кода методом Хаффмана, а также программное кодирование/декодирование префиксного кода по заданной кодовой таблице.