БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИНФОРМАТИКИ И РАДИОЭЛЕКТРОНИКИ

кафедра РЭС

реферат на тему:

**«Основы экономного кодирования»**

МИНСК, 2009

Сообщения, передаваемые с использованием *РТС ПИ* (речь, музыка, телевизионные изображения и т.д.) в большинстве своем предназначены для непосредственного восприятия органами чувств человека и обычно плохо приспособлены для их эффективной передачи по каналам связи. Поэтому они в процессе передачи, как правило, подвергаются кодированию.

Что такое кодирование и зачем оно используется?

Под *кодированием* в общем случае понимают *преобразование алфавита* сообщения A***{ λi }***, ( ***i = 1,2…K***) в алфавит некоторым образом выбранных кодовых символов ***ℜ{ xj },*** ( ***j = 1,2…N***)*.* Обычно (но не обязательно) размер алфавита кодовых символов ***dim ℜ{ xj }*** меньше или намного меньше размера алфавита источника ***dim***A***{ λi }.***

Кодирование сообщений может преследовать различные цели. Например, это кодирование с целью засекречивания передаваемой информации. При этом элементарным сообщениям ***λi*** из алфавита A***{ λi }*** ставятся в соответствие последовательности, к примеру, цифр или букв из специальных кодовых таблиц, известных лишь отправителю и получателю информации.

Другим примером кодирования может служить преобразование дискретных сообщений ***λi*** из одних систем счисления в другие (из десятичной в двоичную, восьмеричную и т. п., из непозиционной в позиционную, преобразование буквенного алфавита в цифровой и т. д.).

Кодирование в канале, или помехоустойчивое кодирование информации, может быть использовано для уменьшения количества ошибок, возникающих при передаче по каналу с помехами.

Наконец, кодирование сообщений может производиться с целью *сокращения объема информации и повышения скорости ее передачи или сокращения полосы частот, требуемых для передачи.* Такое *к*одирование называют *экономным, безызбыточным,* или *эффективным* *кодированием,* а также *сжатием данных.* В данном разделе будет идти речь именно о такого рода кодировании. Процедуре кодирования обычно предшествуют (и включаются в нее) дискретизация и квантование непрерывного сообщения ***λ(t)***, то есть его преобразование в последовательность элементарных дискретных сообщений ***{ λiq }.***

Прежде чем перейти к вопросу *экономного* *кодирования,* кратко поясним суть самой процедуры кодирования.

Любое дискретное *сообщение* ***λi*** *из алфавита источника* A***{ λi }*** объемом в ***K*** символов можно закодировать последовательностью соответствующим образом выбранных *кодовых символов* ***xj*** из алфавита ***ℜ{xj}.***

Например, любое число (а ***λi*** можно считать числом) можно записать в заданной позиционной системе счисления следующим образом:

λi = M = xn−1⋅m n-1 + xn-2⋅m n-2 +… + x0⋅m 0, (1)

где ***m*** - основание системы счисления; ***x0 … xn-1***- коэффициенты при степенях ***m***; ***x ⊂ 0, m - 1***.

Пусть, к примеру, значение ***λi*** = ***M = 59 ,*** тогда его код по основанию ***m*** ***= 8***, будет иметь вид

***M = 59 = 7·81 + 3·80 =738 .***

Код того же числа, но по основанию ***m = 4*** будет выглядеть следующим образом:

***M = 59 = 3⋅42 + 2⋅41+ 3⋅40 = 3234 .***

Наконец, если основание кода ***m = 2***, то

***M = 59 = 1⋅25 + 1⋅24 + 1⋅23 + 0⋅22 + 1⋅21 + 1⋅20 = 1110112*** .

Таким образом, числа ***73***, ***323*** и ***111011*** можно считать, соответственно, восьмеричным, четверичным и двоичным кодами числа ***M = 59***.

В принципе основание кода может быть любым, однако наибольшее распространение получили *двоичные коды*, или коды с основанием ***2,*** длякоторых размер алфавита кодовых символов ***ℜ{ xj }*** равен двум, ***xj ⊂ 0,1.*** Двоичные коды, то есть коды, содержащие только *нули и единицы,* очень просто формируются и передаются по каналам связи и, главное, являются внутренним языком цифровых *ЭВМ*, то есть без всяких преобразований могут обрабатываться цифровыми средствами. Поэтому, когда речь идет о кодировании и кодах, чаще всего имеют в виду именно *двоичные коды.* В дальнейшем будем рассматривать в основном двоичное кодирование.

Самым простым способом представления или задания кодов являются кодовые таблицы, ставящие в соответствие сообщениям ***λi*** соответствующие им коды (табл. 1).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Буква λi | Число λi | Кодс основанием 10 | Кодс основанием 4 | Кодс основанием 2 |
| А | 0 | 0 | 00 | 000 |
| Б | 1 | 1 | 01 | 001 |
| В | 2 | 2 | 02 | 010 |
| Г | 3 | 3 | 03 | 011 |
| Д | 4 | 4 | 10 | 100 |
| Е | 5 | 5 | 11 | 101 |
| Ж | 6 | 6 | 12 | 110 |
| З | 7 | 7 | 13 | 111 |

Таблица 1

*Другим наглядным и удобным способом описания кодов является их представление в виде кодового дерева (рис..1).*

 Корень Узлы

 **0 1**

 Вершина

 **0 1 0 1**

 **0 1 0 1 0 1 0 1**

 **А Б В Г Д Е Ж З**

 **1 2 3 4 5 6 7 8**

Рис. 4.

Рис. 1

Для того, чтобы построить кодовое дерево для данного кода, начиная с некоторой точки - *корня* кодового дерева - проводятся ветви - 0 или 1. На *вершинах* кодового дерева находятся буквы алфавита источника, причем каждой букве соответствуют своя вершина и *свой путь от корня к вершине*. К примеру, букве А соответствует код 000, букве В – 010, букве Е – 101 и т.д.

Код, полученный с использованием кодового дерева, изображенного на рис. 1, является *равномерным трехразрядным кодом*.

Равномерные коды очень широко используются в силу своей простоты и удобства процедур кодирования-декодирования: каждой букве – одинаковое число бит; принял заданное число бит – ищи в кодовой таблице соответствующую букву.

Наряду с равномерными кодами могут применяться и неравномерные коды, когда каждая буква из алфавита источника кодируется различным числом символов, к примеру, А - 10, Б – 110, В – 1110 и т.д.

Кодовое дерево для неравномерного кодирования может выглядеть, например, так, как показано на рис. 2.

А

З

0

0

0

0

1

1

1

1

Б

В

Г

Д

Е

Ж

Рис. 2

При использовании этого кода буква А будет кодироваться, как 1, Б - как 0, В – как 11 и т.д. Однако можно заметить, что, закодировав, к примеру, текст АББА = 1001 , мы не сможем его однозначно декодировать, поскольку такой же код дают фразы: ЖА = 1001, АЕА = 1001 и ГД = 1001. Такие коды, не обеспечивающие однозначного декодирования, называются *приводимыми, или непрефиксными,* кодами и не могут на практике применяться без специальных разделяющих символов. Примером применения такого типа кодов может служить азбука Морзе, в которой кроме точек и тире есть специальные символы разделения букв и слов. Но это уже не двоичный код.

Однако можно построить неравномерные неприводимые коды, допускающие однозначное декодирование. Для этого необходимо, чтобы всем буквам алфавита соответствовали лишь вершины кодового дерева, например, такого, как показано на рис. 3. Здесь ни одна кодовая комбинация не является началом другой, более длинной, поэтому неоднозначности декодирования не будет. Такие неравномерные коды называются префиксными.

1

1

0

0

А

Б

В

1

0

Г

0

Д

1

*Рис..3*

Прием и декодирование неравномерных кодов - процедура гораздо более сложная, нежели для равномерных. При этом усложняется аппаратура декодирования и синхронизации, поскольку поступление элементов сообщения (букв) становится нерегулярным. Так, к примеру, приняв первый 0, декодер должен посмотреть в кодовую таблицу и выяснить, какой букве соответствует принятая последовательность. Поскольку такой буквы нет, он должен ждать прихода следующего символа. Если следующим символом будет 1, тогда декодирование первой буквы завершится – это будет Б, если же вторым принятым символом снова будет 0, придется ждать третьего символа и т.д.

Зачем же используются неравномерные коды, если они столь неудобны?

Рассмотрим пример кодирования сообщений ***λi***из алфавита объемом ***K = 8*** с помощью произвольного ***n***-разрядного двоичного кода.

*Пусть источник сообщения выдает некоторый текст с алфавитом от* **А** *до* **З** *и одинаковой вероятностью букв* **Р(λi ) =** *1/8.*

Кодирующее устройство кодирует эти буквы *равномерным трехразрядным кодом* (см. табл. 1).

Определим основные информационные характеристики источника с таким алфавитом:

- энтропия источника , ;

- избыточность источника ;

- среднее число символов в коде ;

- избыточность кода .

*Таким образом, при кодировании сообщений с равновероятными буквами избыточность выбранного (равномерного) кода оказалась равной нулю.*

Пусть теперь вероятности появления в тексте различных букв будут разными (табл. 2).

Таблица 2

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| А | Б | В | Г | Д | Е | Ж | З |
| ***Ра***=0.6 | ***Рб***=0.2 | ***Рв***=0.1 | ***Рг***=0.04 | ***Рд***=0.025 | ***Ре***=0.015 | ***Рж***=0.01 | ***Рз***=0.01 |

Энтропия источника  в этом случае, естественно, будет меньшей и составит  = 1.781.

*Среднее число символов на одно сообщение при использовании того же равномерного трехразрядного кода*



Избыточность кода в этом случае будет

**,**

или довольно значительной величиной (в среднем 4 символа из 10 не несут никакой информации).

В связи с тем, что *при кодировании неравновероятных сообщений равномерные коды обладают большой избыточностью*, было предложено использовать *неравномерные коды, длительность кодовых комбинаций которых была бы согласована с вероятностью выпадения различных букв*.

Такое кодирование называется статистическим.

Неравномерный код при статистическом кодировании выбирают так, чтобы *более вероятные буквы передавались с помощью более коротких комбинаций кода, менее вероятные - с помощью более длинных*. В результате уменьшается средняя длина кодовой группы в сравнении со случаем равномерного кодирования.

Один из способов такого кодирования предложен Хаффменом. Построение кодового дерева неравномерного кода Хаффмена для передачи одного из восьми сообщений ***λi*** с различными вероятностями иллюстрируется табл. 3.

Таблица 3

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Буква | Вероятность***Рi*** | Кодовое дерево | Код | ***ni*** | ***ni ⋅ Pi*** |
| ***А******Б******В******Г******Д******Е******Ж*** ***З*** | 0.60.20.10.040.0250.0150.010.01 |  | 101001000100001000001000000100000001 | 12345678 | 0.60.40.30.160.1250.080.070.08 |

Среднее число символов для такого кода составит



а избыточность кода



т.е. на порядок меньше, чем при равномерном кодировании.

Другим простейшим способом статистического кодирования является кодирование по методу Шеннона-Фано. Кодирование в соответствии с этим алгоритмом производится так:

* сначала все буквы из алфавита сообщения записывают в порядке убывания их вероятностей;
* затем всю совокупность букв разбивают на две примерно равные по сумме вероятностей группы; одной из них (в группе может быть любое число символов, в том числе – один) присваивают символ “1”, другой - “0”;
* каждую из этих групп снова разбивают (если это возможно) на две части и каждой из частей присваивают “1” и “0” и т.д.

Процедура кодирования по методу Шеннона-Фано иллюстрируется табл. 4.

Таблица 4

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Буква | ***Р(λi )*** | I | II | III | IV | V | Kод | ***ni ⋅ Pi*** |
| **А** | 0.6 | 1 |  |  |  |  | 1 | 0.6 |
| **Б** | 0.2 | 0 | 1 | 1 |  |  | 011 | 0.6 |
| **В** | 0.1 | 0 |  |  | 010 | 0.3 |
| **Г** | 0.04 | 0 | 1 |  |  | 001 | 0.12 |
| **Д** | 0.025 | 0 | 1 |  | 0001 | 0.1 |
| **Е** | 0.015 | 0 |  | 00001 | 0.075 |
| **Ж** | 0.01 | 1 | 000001 | 0.06 |
| **З** | 0.01 | 0 | 000000 | 0.06 |

Для полученного таким образом кода среднее число двоичных символов, приходящихся на одну букву, равно

,

а избыточность кода составит



то есть также существенно меньшую величину, нежели для равномерного кода.

Обратим внимание на тот факт, что как для кода Хаффмена, так и для кода Шеннона-Фано среднее количество двоичных символов, приходящееся на символ источника, приближается к энтропии источника, но не равно ей. Данный результат представляет собой следствие *теоремы кодирования без шума для источника (первой теоремы Шеннона),* которая утверждает:

*Любой источник можно закодировать двоичной последовательностью при среднем количестве двоичных символов на символ источника* ***λi****, сколь угодно близком к энтропии, и невозможно добиться средней длины кода, меньшей, чем энтропия* ***H(λ).***

Значение этой теоремы для современной радиотехники трудно переоценить – она устанавливает те границы в компактности представления информации, которых можно достичь при правильном ее кодировании.

**ЛИТЕРАТУРА**

1. Лидовский В.И. Теория информации. - М., «Высшая школа», 2002г. – 120с.
2. Метрология и радиоизмерения в телекоммуникационных системах. Учебник для вузов. / В.И. Нефедов, В.И. Халкин, Е.В. Федоров и др. – М.: высшая школа, 2001 г. – 383с.
3. Цапенко М.П. измерительные информационные системы.– М.: Энергоатом издат, 2005. - 440с.
4. Зюко А.Г., Кловский Д.Д., Назаров М.В., Финк Л.М. Теория передачи сигналов. М: Радио и связь, 2001 г. –368 с.
5. Б. Скляр. Цифровая связь. Теоретические основы и практическое применение. Изд. 2-е, испр.: Пер. с англ. – М.: Издательский дом «Вильямс», 2003 г. – 1104 с.