УДК 681.32
Ю. А. МИЛОВА, аспирантка;
А. А. ДУДАРЕВА, аспирантка,
Государственный университет телекоммуникаций, Киев

СУММАРНЫЕ КОДЫ

В статье рассматриваются числовые коды, кодовые слова которых формируются суммированием натуральных чисел. Такие коды относятся к классу полипараметрических кодов. Основные параметры кода проявляются после нормирования его любым простым числом, не менее числа 5. Сама структура суммарных кодовых слов позволяет обнаруживать канальные ошибки. Параметры последовательности суммарных кодов обеспечивают исправление ошибок подбором таких кодовых слов, которые характеризуются этими параметрами. Нормирование позволяет получить ряд разновидностей суммарных кодов. При этом весьма важно, что нормирование можно выполнить на приемном конце.

Ключевые слова: натуральный ряд чисел; параметры кода; кодовое слово; суммарное кодовое слово; двоичные суммарные коды; полипараметричность; остаток от деления; дуальная кратность.

ВВЕДЕНИЕ

Коды, обнаруживающие и исправляющие ошибки

Если взять достаточно длинное предложение из любого текста на естественном языке и исказить его, заменив в некоторых местах одни буквы на другие, исключив ряд букв или добавив новые, то знание структуры отдельных слов и предложения в целом, а также предыдущего и последующего текстов во многих случаях позволит полностью восстановить исходное предложение или по крайней мере безошибочно уловить его смысл. Это показывает, что естественные языки обладают очень большой избыточностью, благодаря которой возможно обнаружение ошибок и повышение надежности используемых технических устройств.

Однако проблема состоит не в том, чтобы просто повысить надежность за счет введения очень большой избыточности, а в том, как с помощью по возможности меньшей специальным образом вводимой избыточности достичь нужной степени надежности системы связи.

Как известно, избыточность английского языка достигает 70%, однако нельзя сказать, что избыточность вводится в английский текст оптимально с точки зрения возможности исправления и обнаружения ошибок. Основная задача теории кодирования — повышение надежности систем связи и вычислительных систем с помощью целенаправленного эффективного введения избыточности в процессе представления информации (в процессе преобразования информации). Введение избыточности приводит к снижению количества сообщений, которые могут быть переданы или обработаны за определенный период времени, а кроме того, предполагает использование в системе дополнительных устройств для целенаправленного введения избыточности (кодеров), устройств для обнаружения и исправления возникающих ошибок (декодеров) и ряда других дополнительных устройств.

Поступающая от источника информация вначале с помощью преобразователя «источник информации – двоичная последовательность» преобразуется в последовательность двоичных символов, которая подается на вход кодера, где в нее вводится избыточность. Символы с выхода кодера с помощью модулятора преобразуются в сигналы, которые могут быть переданы по каналу. Эти сигналы поступают в канал, где они обычно искажаются шумами. На приемном конце искаженные сигналы канала с помощью демодулятора преобразуются в последовательность двоичных символов, содержащую избыточные символы. Используя эту избыточность, декодер обнаруживает и исправляет возникшие ошибки. Двоичная последовательность символов на выходе декодера уже не является избыточной. Если воздействие шума в канале не слишком сильное и декодер может исправить все возникшие ошибки, то двоичная последовательность на выходе декодера будет совпадать с двоичной последовательностью на выходе преобразователя «источник информации – двоичная последовательность». И, наконец, двоичная последовательность, получающаяся на выходе декодера, преобразуется в сигналы, приемлемые для получателя информации, и передается последнему.

Рассмотрим пример, раскрывающий идею суммарных кодов.

Предположим, что имеется 16 сообщений, каждому из которых взаимно однозначно поставлена в соответствие двоичная последовательность (x_3, x_5, x_6, x_7) длины 4. Вместо последовательности (x_3, x_5, x_6, x_7) для представления сообщений будем использовать двоичную последовательность (x_1, x_2, ..., x_7) длины 7, символы x_1, x_2 и x_4 которой определены следующим образом:

$$\begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{bmatrix} = 0, \text{ т. е. } \begin{cases} x_1 = x_3 + x_5 + x_7, \\ x_2 = x_3 + x_6 + x_7, \\ x_4 = x_5 + x_6 + x_7. \end{cases} \quad (1)$$

Здесь знак «+» означает сложение по модулю 2, а остальные обозначения имеют тот же смысл, что и в обычных матричных соотношениях. Поскольку в процессе передачи или хранения последовательности $(x_1, x_2, ..., x_7)$ могут возникнуть ошибки, то принимаемая последовательность $(x'_1, x'_2, ..., x'_7)$, вообще говоря, будет отлична от исходной. Рассмотрим задачу восстановления исходной последовательности $(x_1, x_2, ..., x_7)$ по известной последовательности $(x'_1, x'_2, ..., x'_7)$. Предположим, что вероятность возникновения двух и более ошибок в двоичных символах одной последовательности длины 7 настолько мала, что этим событием можно пренебречь. Пусть

$$x_i + x'_i = e_i,$$

где сложение выполняется по модулю 2. Если $e_i = 1$, то это означает, что в i-м двоичном символе произошла ошибка; если $e_i = 0$, то i-й двоичный символ передан без ошибок. Используя матрицу из левой части (1), определяем r_1, r_2 и r_3 следующим образом:

$$\begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x'_1 \\ x'_2 \\ x'_3 \\ x'_4 \\ x'_5 \\ x'_6 \\ x'_7 \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix}. \quad (2)$$

Из формул (1) и (2) получаем

$$\begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \\ e_7 \end{bmatrix} = \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix}. \quad (3)$$

По предположению только один из символов $e_1, e_2, ..., e_7$ может быть равен 1. При отсутствии ошибок, когда $e_1 = e_2 = ... = e_7 = 0$, имеем $r_1 = r_2 = r_3 = 0$.

Далее рассмотрим случай, когда $e_i = 1$, а остальные шесть символов e_j, $j \neq i$, равны 0. Заметим, во-первых, что все столбцы матрицы

$$P = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}$$

из левой части (3) различны, а во-вторых, что k-й столбец этой матрицы представляет собой двоичную запись номера k (самый младший разряд двоичного представления числа k является верхним элементом столбца). При сделанных предположениях вектор-столбец $\begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix}$ совпадает с i-м столбцом матрицы P.

Следовательно, номер i искаженного двоичного символа может быть найден следующим образом: $i = r_1 + 2r_2 + 4r_3$. С помощью равенства $x_i = x'_i + e_i$ находим правильные двоичные символы $x_1, ..., x_7$.

ОСНОВНАЯ ЧАСТЬ

Получение суммарных кодов

Известно, что любая двоичная кодовая комбинация может быть представлена натуральным числом.

Суммарные коды являются кодами числовыми. По своей структуре каждое суммарное кодовое слово порождается натуральным рядом чисел и может быть представлено в виде двоичного числа.

Любое кодовое слово имеет свой порядковый номер. Условимся счетный набор суммарных кодов нумеровать в порядке следования натуральных чисел: 1, 2, 3, ... и формировать сложением этих чисел по такому правилу: код с порядковым номером n получается как двоичное представление суммы десятичных чисел вида $1 + 2 + \ldots + n$. Например, код с порядковым номером $n = 5$ получается как двоичное представление суммы $S_{(10)}$ десятичных чисел вида $1 + 2 + 3 + 4 + 5 = 15$, или суммы $S_{(2)} = 1111$ двоичных чисел. При $n = 10$ $S_{(10)} = 1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 + 10 = 55$, или $S_{(2)} = 110111$.

Особенностью суммарных кодов является их полипараметричность. Каждое кодовое слово имеет обычно несколько параметров, по которым его можно выделить из множества подобных кодовых слов и на фоне внешних помех, искажающих структуру кодового слова, восстановить его. Эти параметры рассмотрены далее.

Основные параметры суммарных кодов

Приведем параметры, позволяющие проверить правильность кодового слова суммарного кода и во многих случаях восстановить его в первоначальном виде.

♦ Принадлежность кодового слова к группе суммарных кодов. Если принятое кодовое слово согласно своему значению не принадлежит к группе суммарных кодов, построенных по указанному правилу, оно заведомо ошибочно.

Суммарная структура кодового слова является его главным параметром.

♦ Если остаток от деления кодового слова на некоторое выбранное заранее простое число имеет остаток, отличный от ожидаемого остатка, оно ошибочно.

Вид и значение остатка от деления суммарного кодового слова на заданное простое число является вторым его параметром.

Отметим два характерных свойства суммарных кодов, которые легко проверяются экспериментально.

Первое свойство. Множество кодовых слов при некоторых значениях их длин n, соотнесенных тому или иному простому числу K, обязательно дают одну или несколько дуальных кратностей D (попарно расположенные результаты деления $S_{(n)}$ на K без остатка).

Варианты дуальных кратностей D для суммарных кодов с порядковыми номерами n от 1 до 105 для четырех простых $K = 7$, 11, 13 и 17 приведены в таблице.

Дуальные кратности суммарных кодов

Параметры кода													
$K = 7$	n	6	7	13	14	20	21	48	49	83	84	97	98
	D	3	4	13	15	30	33	168	175	498	510	679	693
$K = 11$	n	10	11	21	22	32	33	54	55	76	77	98	99
	D	5	6	21	23	48	51	135	140	266	273	441	450
$K = 13$	n	12	13	25	26	38	39	51	52	77	78	90	91
	D	6	7	25	27	57	60	102	106	231	237	315	322
$K = 17$	n	16	17	33	34	50	51	67	68	84	85	101	102
	D	8	9	33	35	75	78	134	138	210	215	303	309

Из таблицы следует, что дуальные кратности встречаются достаточно часто с постоянным интервалом, равным K. При этом первая кратность D возникает при $n = K - 1$.

Второе свойство. Если одно целое число разделить на второе, в результате получим целую часть и остаток, который всегда меньше делителя. Для суммарных кодов существует следующая закономерность: начиная с любой дуальной кратности, остатки от деления целых чисел $S(n)$ на K вверх и вниз по множеству порядковых номеров n являются симметричными и попарно равны между собой.

Значения остатков сохраняются по всему множеству номеров n кодовых слов.

Таким образом, для каждого суммарного кодового слова можно определить три его идентификационных параметра:

1) порядковый номер n;

2) значение суммарного кодового слова $S(n)$;

3) остаток от деления суммарного кодового слова на выбранное число K.

Отметим, что *два соседних суммарных кодовых слова с дуальной кратностью имеют нулевые остатки.*

ВЫВОДЫ

1. Суммарные коды являются полипараметрическими, что может служить важным фактором для их идентификации.

2. Посредством выбора различных значений делителей K получаются различные версии нормированных по этому делителю суммарных кодов.

3. Совместное одновременное использование нескольких параметров суммарных кодов увеличивает верность обмена данными.

4. Один отдельный параметр суммарного кода позволяет оценить верность суммарного кодового слова. Дополнительное использование других параметров позволяет восстановить его правильность.

5. Возможны различные методы реализации кодовых параметров.

6. Полипараметричность суммарных кодов удобно использовать при реализации защиты информации от несанкционированного доступа.

7. Конкретная реализация каждого параметра определяется пользователями.

8. Суммарные коды могут найти применение для адресации новых процессов.

Список использованной литературы

*1. **Дикарев А. В.** Фрактальная структура сжатого отрезка натурального ряда // Зв'язок. 2017. №3(127). С. 34–38.*

*2. **Дикарев А. В.** Сжатие двоичных блочных кодов // Зв'язок. 2017. №1(125). С. 40–42.*

*3. **Алгоритми** створення проріджувальних кодів / В. Г. Сайко, О. В. Дікарєв, Л. М. Грищенко [та ін.] // Зв'язок. 2017. №2(126). С. 33–38.*

*4. **Фракталы** в физике / под ред. Л. Пьетронеро, Э. Тозатти. Москва: Мир, 1988. С. 94.*

*5. **Берлекэмп Э.** Алгебраическая теория кодирования / пер. с англ. Э. Берлекэмп. Москва: Мир, 1971. 477 с.*

Рецензент: доктор техн. наук, доцент **С. И. Отрох**, Государственный университет телекоммуникаций, Киев.

Ю. О. Мілова, Г. О. Дударєва

СУМАРНІ КОДИ

У статті розглядаються числові коди, кодові слова яких формуються підсумовуванням натуральних чисел. Такі коди належать до класу поліпараметричних кодів. Основні параметри коду визначаються після нормування його будь-яким простим числом, не меншим за число 5. Сама структура сумарних кодових слів дозволяє знаходити канальні помилки. Параметри послідовності сумарних кодів забезпечують виправлення помилок добором таких кодових слів, які відповідають цим параметрам. Нормування дозволяє отримати низку різновидів сумарних кодів. При цьому важливо, що нормування можна виконати на приймальному кінці.

Ключові слова: натуральний ряд чисел; параметри коду; кодове слово; сумарне кодове слово; двійкові сумарні коди; поліпараметричність; остача від ділення; дуальна кратність.

Y. O. Milova, A. A. Dudareva

TOTAL CODES

This article presents numerical cumulative codes. Every cumulative code word is created by way of addition of natural sequence elements. Furthermore every cumulative code word can be represented in the form of binary number. This code is polyparametric code, where each code word has several parameters. The main parameters of the code are determined after it's normalizing with any prime number upwards of 5. The structure of the cumulative codes detect channel errors. The parameters of the cumulative codes sequence provide errors correction by selecting such code words that satisfy these parameters. One parameter can be used to estimate the accuracy of the cumulative code word. The parameter itself is determined directly by the user. Additional parameters make it possible to restore the correctness of the code word. The properties are demonstrated which is that beginning with any dual multiplicity, the residuals from dividing the integers of the cumulative codes by a given divisor up and down according to the set of sequence numbers are symmetric and pairwise equal to each other, starting with any dual multiplicity. This method makes it possible to implement various methods of code parameters. Rationing allows you to get a number of varieties of codes, and this valuation can be performed at the receiving end. The polyparametricity of cumulative codes allows them to be identified. These cumulative codes can be used to compress information during its transmission, to realize the protection of information from errors, and also to realize the protection of confidential information from unauthorized access. They can also be used to address new processes.

Keywords: numeric natural series; code parameters; code sequence number; code word; digital cumulative codes; code settings; cumulative code word; the value of the cumulative code word; the remainder of the division of the cumulative code word; binary code combination; binary cumulative codes; a countable set of aggregate codes; polyparametrical; remainder; dual multiplicity; external interference; external interference that distorts the structure of the code word.