7.10. Новая теория кодирования

Что такое теория кодирования?

«Кодирование» - одно из наиболее распространенных понятий в современной информатике. Что же такое «кодирование»? Под «кодированием» понимается операция отождествления символов или групп символов одного кода с символами или группами символов другого кода. Необходимость кодирования возникает прежде всего из потребности приспособить форму сообщения к данному «каналу связи» или какому-либо устройству, предназначенному для преобразования или хранения информации.

«Теория кодирования» имеет длительную историю. *Системы счисления*, о которых я рассказывал раньше, и были первыми кодами, предназначенными для представления чисел. Следующее древнее направление в теории кодирования – это *криптография* или *секретное кодирование*. Криптография берет свое начало в Египетской науке и ее возникновение относится к тому периоду (2-е тысячелетие до н.э.), когда Египтяне использовали иероглифический код для надписей на могилах.

Развитие современной теории кодирования стимулировалось прогрессом систем связи. В середине 20-го столетия американский ученый Клод Шеннон разработал *теорию эффективных кодов,* основанных на понятии энтропии. Эта теория нашла широкое применение в современных компьютерах для сжатия информации.

Потребность защитить информацию и системы связи от вредного влияния шумов и помех, разрушающих информацию, способствовали развитию *теории избыточных кодов*. Код Хемминга, циклические коды – это хорошо известные специалистам примеры избыточных кодов.

Таким образом, современная теория кодирования – это совокупность, по крайней мере, четырех различных направлений:

(1) *Теория систем счисления;*
(2) Теория криптографии;
(3) Теория эффективных кодов;
(4) Теория избыточных кодов.

Для решения задач кодирования широко используется различный математический аппарат. Например, в теории алгебраических кодов широко используется аппарат теории групп, в современной криптографии основным математическим аппаратом является теория чисел.

Как я пришел к новой теории кодирования?

С сентября 1998 года я начал работать на кафедре математики информатики Университета Эдуардо Мондлане (Мапуту, Мозамбик). Очень важным достоинством кафедры было наличие в ней прекрасной библиотеки, в которой была собран хорошая коллекция книг по математике и информатике. Библиотека стала одним из мест моего частого времяпровождения в период работы в этом знаменитом африканском университете. И там меня однажды сфотографировали для кафедрального сайта.



Фото в библиотеке кафедры математики и информатики Университета Эдуардо Мондлане (Мапуту, Мозамбик, 1999 г.)

Однажды мне попалась на глаза книга “Computer security management” (автор Dennis van Tassel), опубликованная издательством “Prentice-Hall”, New Jersey, 1972. В этой книге я обнаружил описание так называемого «матричного» метода кодирования информации. Суть кодирования состояла в представлении исходного сообщения в виде квадратной матрицы с последующим умножении такой матрицы на специальную кодирующую матрицу того же порядка. Декодирование состояло в умножении «кодовой матрицы» на «инверсную» матрицу. Прочитав описанный метод, у меня мгновенно созрела идея использования разработанных мною Q_p-матриц для кодирования информации.

Суть метода кодирования состоит в представлении исходного сообщения в виде квадратной матрицы M и последующем ее умножением на кодирующую матрицу, в качестве которой выбирается матрица типа Q_p^n. Ясно, что размер исходной матрицы M должен быть таким же, как и размер кодирующей матрицы Q_p^n. В результате такой операции образуется «кодовая матрица» E, элементы которой направляются в «канал связи». Декодирование состоит в умножении кодовой матрицы E на матрицу Q_p^{-n}, инверсную к матрице Q_p^n, то есть, такую, что $Q_p^n \times Q_p^{-n} = I$. Таблица ниже демонстрирует общую идею «фибоначчиевого» алгоритма кодирования-декодирования.

Кодирование	**Декодирование**
$M \times Q_p^n = E$	$E \times Q_p^{-n} = M$

А теперь установим некоторые свойства сформулированного выше метода «фибоначчиевого» кодирования-декодирования, основанного на применении Q_p-матриц. Прежде всего установим, как связаны между собой детерминанты исходной матрицы M и кодовой матрицы E. Поскольку кодовая матрица E вычисляется на путем умножения исходной матрицы M на кодирующую матрицу Q_p^n, то есть,

$$E = M \times Q_p^n \ , \tag{1}$$

то детерминант матрицы (1) может быть вычислен следующим образом:

$$\text{Det } E = (\text{Det } M) \times (\text{Det } Q_p^n). \tag{2}$$

Но выше мы установили, что

$$\text{Det } Q_p^n = (-1)^{pn}. \tag{3}$$

Подставляя выражение (3) в (4), мы получим очень важное соотношение, связывающее детерминанты матриц E и M:

$$\text{Det } E = (\text{Det } M)\times(-1)^{pn}. \quad (4)$$

Обнаружение ошибок

Как упоминалось, одна из целей теории «избыточного кодирования» состоит в обнаружении и коррекции ошибок, возникающих в кодовом сообщении *E* под влиянием «помех» в «канале». Покажем сейчас, как обнаруживать ошибки в кодовом сообщении *E*, используя «фибоначчиевое» кодирование. Главная идея состоит в использовании свойства (4), связывающего детерминанты исходной матрицы *M* и кодовой матрицы *E*. Для того, чтобы обнаруживать ошибки в элементах кодовой матрицы *E,* достаточно передать по «каналу связи» значение детерминанта исходного сообщения *M*. И тогда, получив элементы кодовой матрицы *E,* мы можем проверить выполнение фундаментального соотношения (4), связывающего детерминанты матриц *M* и *E*. Нарушение равенства (4) и является признаком ошибки в элементах матрицы *E*.

Исправление ошибок

Но обнаружить ошибку в сообщении – это только первый шаг в повышении достоверности передаваемого сообщения. Надо теперь эту ошибку исправить. И оказалось, что рассмотренное выше фундаментальное свойство (4) позволяет не только обнаруживать, но и исправлять ошибки в элементах кодовой матрицы *E*. Продемонстрируем эту возможность на простейшем примере, когда мы используем для кодирования простейшую «фибоначчиевую» *Q*-матрицу. Для этого случая исходное сообщение может быть представлено в виде квадратной матрицы 2-го порядка:

$$M = \begin{pmatrix} m_1 & m_2 \\ m_3 & m_4 \end{pmatrix}. \quad (5)$$

Детерминант такой матрицы равен:

$$\text{Det } M = m_1 m_4 - m_2 m_3. \quad (6)$$

Пусть в качестве кодирующей выбрана «фибоначчинвая» матрица $Q^5 = \begin{pmatrix} 8 & 5 \\ 5 & 3 \end{pmatrix}$. Тогда в соответствии с правилами матричного умножения кодовая матрица *E* принимает следующий вид:

$$E = M \times Q^5 = \begin{pmatrix} m_1 & m_2 \\ m_3 & m_4 \end{pmatrix} \times \begin{pmatrix} 8 & 5 \\ 5 & 3 \end{pmatrix} = \begin{pmatrix} 8m_1 + 5m_2 & 5m_1 + 3m_2 \\ 8m_3 + 5m_4 & 5m_3 + 3m_4 \end{pmatrix} = \begin{pmatrix} m_1' & m_2' \\ m_3' & m_4' \end{pmatrix}. \quad (7)$$

В «канал связи» посылаются элементы кодовой матрицы (7) и вслед за ними детерминант исходной матрицы, вычисленный по формуле (6). Предположим, что в процессе передачи произошло искажение первого элемента кодового сообщения m_1', то есть $m_1' = x$.

Тогда мы можем представить «разрушенное» кодовое сообщение в следующем виде:

$$E' = \begin{pmatrix} x & m_2' \\ m_3' & m_4' \end{pmatrix} \quad (8)$$

где *x* есть «разрушенный» элемент кодового сообщения *E*, а остальные элементы матрицы должны быть корректными и равными следующим значениям:

$$\begin{aligned} m_2' &= 5m_1 + 3m_2; \\ m_3' &= 8m_3 + 5m_4; \\ m_4' &= 5m_3 + 3m_4 \end{aligned} \quad (9)$$

Тогда согласно свойствам «фибоначчиевого» кодирования для случая кодирующей матрицы Q^5 детерминант матрицы M' должен быть равным детерминанту (6), взятым с противоположным знаком. Тогда мы можем записать следующее уравнение для вычисления «разрушенного» элемента x:

$$xm_4^{'} - m_2^{'}m_3^{'} = x(5m_3 + 3m_4) - (5m_1 + 3m_2)(8m_3 + 5m_4) = - m_1m_4 + m_2m_3 \qquad (10)$$

И после выполнения элементарных математических преобразований над уравнением (10) мы можем получить решение уравнения (10) в виде:

$$x = 8m_1 + 5m_2 \qquad (11)$$

Сравнивая вычисленное значение (11) с элементом $m_1^{'}$ кодовой матрицы E, задаваемой (7), мы можем сделать следующее важное заключение:

$$x = m_1^{'}.$$

Таким образом, мы восстановили кодовое сообщение E, используя свойства детерминанта «фибоначчиевой» матрицы Q!

Я не буду утомлять моего читателя дополнительными довольно сложными рассуждениями и отсылаю наиболее любопытных к моей книге «Introduction into Fibonacci Coding and Cryptography", опубликованной в 1999 г.

«Фибоначчиевая» криптография

В упомянутой выше книге изложена идея применения рассмотренного выше «фибоначчиевого» метода кодирования-декодирования для криптографической шифрации-дешифрации информации.

Что такое «криптографическое кодирование»? Это – очень важное направление современной информатики. С помощью «криптографического кодирования» мы осуществляем преобразование исходного сообщения M в некоторое зашифрованное сообщение E; при этом его дешифрация может быть осуществлена только пользователем, знающим «криптографический ключ».

Рассмотрим теперь «фибоначчиевое» кодирование-декодирование, основанное на применении «фибоначчиевой» матрицы Q_p^n. В данном случае числа p и n, задающие матрицу Q_p^n, могут принимать значение из множеств {p=0, 1, 2, 3, …}и {n = 1, 2, 3, …}. Это означает, что теоретически существует бесконечное количество способов «фибоначчиевого» кодирования-декодирования. И это обстоятельство мы можем использовать для «криптографического кодирования». Действительно, каждый пользователь Internet может получить в свое распоряжение только свои, известные только ему «ключевые» числа p и n и тогда его общение через Intenet с другими пользователями может осуществляться с помощью «фибоначчинвого» кодирования, основанного на «своей» матрице типа Q_p^n. С помощью такого метода можно легко зашифровать некоторую важную информацию в собственном компьютере или некоторой «базе данных». То есть, можно найти много различных приложений предложенного метода, если только захотеть. При этом не надо забывать, что наш «криптографический метод» позволяет еще и обнаруживать и исправлять ошибки в засекреченном сообщении!

И сейчас мы подведем итог нашему рассмотрению нового метода кодирования, основанного на «фибоначчиевых» матрицах типа Q_p^n. Во-первых, мы рассмотрели весьма необычный метод избыточного кодирования и «криптографической шифрации» информации! В нашем методе объектами обнаружения и исправления ошибок являются элементы матрицы E! Мы корректируем не отдельные биты или их комбинации в двоичных кодах, как это принято в классической теории кодирования, а числа, являющиеся элементами матриц. И наконец, мы используем необычный математический аппарат для нашей теории кодирования, а именно теорию матриц и теорию p-чисел Фибоначчи, введенных мною еще в моей докторской диссертации.