Министерство образования и науки Российской Федерации

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ «САРАТОВСКИЙ НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ Н.Г. ЧЕРНЫШЕВСКОГО»

Кафедра геометрии

Порождающий многочлен циклического кода

АВТОРЕФЕРАТ БАКАЛАВРСКОЙ РАБОТЫ

студента (ки) 4 курса 421 группы

направления 02.03.01 – Математика и компьютерные науки,
код и наименование направления
профиль подготовки: Математические основы компьютерных наук

Механико-математический факультет
наименование факультета, института, колледжа

РИДЕЛЯ АРТЁМА ВЛАДИМИРОВИЧА
фамилия, имя, отчество

Научный руководитель
доцент, к.ф.-м.н., доцент _____ В.Е.НОВИКОВ
должность, уч. степень, уч. звание подпись, дата инициалы, фамилия

Зав. кафедрой
доктор физ-мат. наук, профессор _____ В.В.РОЗЕН
должность, уч. степень, уч. звание подпись, дата инициалы, фамилия

Саратов 2016

ВВЕДЕНИЕ

Теория кодирования — одна из тех областей математики, которые заметно повлияли на развитие компьютерных наук. Ее область действия распространяется на передачу данных по реальным (или зашумленным) каналам, а предметом является обеспечение корректности переданной информации. Иными словами, она изучает, как лучше упаковать данные, чтобы после передачи сигнала из данных можно было надежно и просто выделить полезную информацию.

Первым теоретическое решение проблемы передачи данных по зашумленным каналам предложил Клод Шеннон, основоположник статистической теории информации. Шеннон написал работу «Математическая теория передачи сообщений» (1948), где показал, что если энтропия равна единице, то информация передается со скоростью, равной пропускной способности канала, а если энтропия меньше единицы, то информация будет передана с временными задержками.

Труды Шеннона дали пищу для множества дальнейших исследований в области теории информации, но практического инженерного приложения они не имели. Переход от теории к практике стал возможен благодаря усилиям Ричарда Хэмминга, коллеги Шеннона по Bell Labs, получившего известность за открытие класса кодов, которые так и стали называть «кодами Хэмминга». Именно Хэмминг первым предложил «коды с исправлением ошибок». Современные модификации этих кодов используются во всех системах хранения данных и для обмена между процессором и оперативной памятью. Один из их вариантов, коды Рида-Соломона применяются в компакт-дисках, позволяя воспроизводить записи без скрипов и шумов, которые могли бы вызвать царапины и пылинки. Существует множество версий кодов, построенных «по мотивам» Хэмминга, они различаются алгоритмами кодирования и количеством проверочных битов. Особое значение подобные коды приобрели в связи с развитием дальней космической связи с межпланетными станциями.

В настоящее время теория кодирования имеет важное широкое практическое применение как средство экономной, удобной, быстрой, а также надежной передачи сообщений по линиям связи с различного вида шумами

(телефон, телеграф, радио, телевидение, компьютерная, космическая связи и т. д.).

Всего в работе 4 раздела. Для начала, рассматривается теория конечных групп и полей, как необходимая база для дальнейшего восприятия информации по теории кодирования. Далее следуют линейные коды и методы их задания. В качестве примера будут рассмотрены Коды Хэмминга. Основными в работе являются циклические коды вместе с методами их кодирования и декодирования. Циклические коды незаменимы при необходимости передавать информацию в каналах связи, в которых отсутствует возможность повторной передачи данных. Циклические коды применяются при записи и считывании на HDD, CD и DVD, при использовании USB-портов для обмена информацией, при передаче аудио и видео информации.

Краткое содержание работы:

В **Разделе 1** рассматривается множество с бинарной операцией, называемое группой. В качестве примера приводится множество целых чисел с операцией + (обычным сложением). Отмечено, в каком случае группа называется конечной, а в каком бесконечной. Дано определение подгруппы, а также фигурирует группа, в которой все элементы являются степенями какого-либо одного (такая группа называется циклической). Приводится описание отношения R_H, которое разбивает группу на классы эквивалентности. Основным результатом данного раздела является теорема Лагранжа:

Теорема 1.3. Порядок конечной группы G равен произведению порядка любой подгруппы H на её индекс, т.е. $|G| = |H| \cdot (G : H)$.

Раздел 2 посвящен конечным полям. Для начала дается определение кольца, как вспомогательной алгебраической структуры. Дается понятие поля, в качестве примера которого приводится множество действительных чисел, которое является полем относительно операций сложения и умножения. Определяется крайне важная вещь, называемая характеристикой поля, которой является наименьшее натуральное n, для которого выполняется равенство $n \cdot r = 0 \quad \forall r \in F$, где F — поле. Приводится определение подполя, а также простого поля, как поля, не имеющего собственных подполей. Далее следует теорема 2.2, которая говорит о том, что если поле F имеет характеристику 0, то в него естественно вложено подполе $\mathbb{Q}$ рациональных чисел, а если поле F имеет характеристику p, то в него естественно вложено подполе $\mathbb{F}_p$.

В качестве подраздела приводится описание мультипликативной группы поля, где в теореме 2.7 указывается, что мультипликативная группа любого конечного поля — циклическая.

Завершается раздел 2 описанием многочленов над конечными полями. Определяется, что такое многочлены, а также показывается как складывать и умножать их. Приводятся различные свойства степеней многочленов. Вводится понятие делимости, результатом которого становится теорема о делении многочленов с остатком (теорема 2.9). После отмечается тот факт, что многочлены бывают приводимыми и неприводимыми. Упоминается, что нас в данной работе будут в основном интересовать именно неприводимые мно-

гочлены, в силу своей значимости в теории циклических кодов. Основным результатом в данном разделе является теорема 2.14, а также следствие из этой теоремы (следствие 2.15):

Теорема 2.14. Для каждого простого числа p и каждого натурального числа n существует конечное поле из p^n элементов. Любое конечное поле из $q = p^n$ элементов изоморфно полю разложения многочлена $x^q - x$ над полем F_p.

Следствие 2.15. Для каждого натурального числа n и для любого простого p над полем $GF(p)$ существует неприводимый многочлен степени n.

В **Разделе 3** представлены линейные коды. Для начала вводятся такие понятия, как вес вектора, расстояние Хэмминга, а также минимальное расстояние линейного кода.

Приводится матричное описание линейного кода, который в данном случае задается как пространство строк порождающей матрицы. Чтобы закодировать информационное слово с помощью порождающей матрицы, необходимо воспользоваться формулой (3.1). После примера кодирования приводится интересное замечание, суть которого в том, что одно и тоже информационное слово можно кодировать в разные кодовые слова, так как для одного и того же кода существуют различные порождающие матрицы. Происходит это вследствие того, что множество базисных векторов для одного и того же линейного пространства можно выбрать различными способами.

Следующим этапом в данном разделе становится определение так называемой проверочной матрицы линейного кода. Из условия, что линейный код является подпространством, следует, что для него существует ортогональное дополнение. На основе этого говорится, что проверочной матрицей линейного кода называется порождающая матрица для ортогонального дополнения этого кода. В качестве метода построения проверочной матрицы указывается нахождение фундаментальной системы решений для системы уравнений $G \cdot x^T = 0$, где G — порождающая матрица. Указывается теорема 3.2, устанавливающая связь между кодовым расстоянием линейного кода и проверочной матрицей. Следствие данной теоремы является главным результатом раздела 3:

Следствие 3.3. Если любые $d-1$ столбцов проверочной матрицы H линейно независимы, то минимальное расстояние кода меньше либо равно d. Если при этом найдутся d линейно зависимых столбцов, то минимальное расстояние кода в точности равно d.

В качестве примера приводятся коды Хэмминга, как самые известные из существующих кодов. Описывается построение кода $(7, 4)$ с соответствующей ему проверочной матрицей

$$H = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}$$

Указывается, как происходит кодирование и декодирование. На основе предложений 3.4 и 3.5 делается вывод, что код Хэмминга способен обнаруживать и исправлять одиночные ошибки.

В завершение данного раздела приводится так называемый расширенный код Хэмминга длины 7. Он получается из обычного кода Хэмминга путем добавления к проверочной матрице нулевого столбца, а затем строки из единиц. В результате получается

$$H' = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

Главным отличием от обычного кода Хэмминга является то, что расширенный код способен обнаруживать двойные ошибки.

Раздел 4 полностью посвящен циклическим кодам. Для начала указывается, что циклические коды являются разновидностью линейных кодов и поэтому обладают всеми их свойствами. Отмечается тот факт, что представление двоичных комбинаций будет осуществляться с помощью многочленов, а не в виде последовательностей нулей и единиц. Приводится определение циклического кода, которое говорит о том, что циклический код это тот же

самый линейный код, с одним лишь условием, что вместе с каждым кодовым вектором коду также принадлежит циклически сдвинутый кодовый вектор.

Образование циклического кода происходит с помощью многочлена, который называется порождающим. Основными в данном разделе являются следующие теоремы:

Теорема 4.3. Пусть C — циклический (n, k)-код и $p(x)$ — его порождающий многочлен. Тогда степень $p(x)$ равна $n - k$ и каждое слово из C может быть единственным образом представлено в виде

$$c(x) = m(x)p(x).$$

Теорема 4.4. Порождающий многочлен $p(x)$ циклического (n, k)-кода является делителем двучлена $x^n - 1$.

Выполнение теоремы 4.3 приводит к тому, что все рабочие кодовые комбинации циклического кода приобретают свойство делимости на порождающий многочлен без остатка.

Матричное описание циклического кода строится на утверждении 4.5, в котором говорится, что многочлены $p(x), xp(x), \ldots, x^{k-1}p(x)$ линейно независимы. Из этого факта делается вывод, что они могут быть выбраны в качестве базиса (n, k)-кода. Порождающая матрица циклического кода представляется в символьной форме как:

$$G = \begin{pmatrix} p_0 & p_1 & \ldots & p_{r-1} & p_r & 0 & \ldots & 0 \\ 0 & p_0 & \ldots & p_{r-2} & p_{r-1} & p_r & \ldots & 0 \\ \vdots & \vdots & \ldots & \vdots & \vdots & \vdots & \ldots & \vdots \\ 0 & 0 & \ldots & 0 & p_0 & p_1 & \ldots & p_r \end{pmatrix}$$

Кодирование и декодирование с использованием циклических кодов происходит довольно просто. Кодирование заключается в отображении многочлена $m(x)$ в кодовое слово $c(x)$, которое сопоставляется этому информационному многочлену. Утверждается, что кодирование бывает двух видов: систематическое и несистематическое. В примере 11 рассматривается как кодируется информационное сообщение несистематическим способом, а в примере 12 кодирование систематическим способом:

Пример 12. Пусть $k = 4$, и дан порождающий многочлен третьей степени $p(x) = x^3 + x + 1$. Следовательно, кодовые комбинации циклического кода будут иметь по семь разрядов. Возьмем последовательность (1101), т.е. $m(x) = x^3 + x^2 + 1$, и вычислим ее кодовую комбинацию.
Вычислим произведение $m(x) \cdot x^3 = x^6 + x^5 + x^3$. Разделим на порождающий многочлен $p(x)$:

$$\frac{m(x) \cdot x^r}{p(x)} = \frac{x^6 + x^5 + x^3}{x^3 + x + 1} = x^3 + x^2 + x + 1 + \frac{1}{p(x)}.$$

Видим, что остаток $R(x) = 1$. Следовательно, можем найти кодовую комбинацию, принадлежащую циклическому $(7, 4)$-коду. $c(x) = m(x) \cdot x^r + R(x) = x^6 + x^5 + x^3 + 1$, что в двоичной форме означает (1101001).

Указывается, что в работе нас будет интересовать именно систематическое кодирование, для которого написана программа (приложение А), автоматизирующая процесс кодирования. В качестве метода декодирования указывается, что существует также два способа. Это очевидно, так как выбор способа декодирования полностью зависит от того, каким способом сообщение было закодировано.

Что касается коррекции ошибок циклическими кодами, то в подразделе 4.4 приводится алгоритм коррекции одиночных ошибок. Он заключается в том, что вес остатка стараются привести в рамки корректирующей способности кода. В примере 14 можно наглядно увидеть процесс коррекции ошибки. Также указывается, что в работе (приложение Б) присутствует программа-декодер, способная корректировать одиночные ошибки.

ЗАКЛЮЧЕНИЕ

В ходе данной работы были описаны основные понятия и важные теоремы теории конечных групп и конечных полей. На основе этого были рассмотрены линейные коды, методы их задания, а также коды Хэмминга, как основные коды в теории кодирования. Циклические коды, как основные в данной работе, были рассмотрены с точки зрения образования с помощью порождающего многочлена. Был подробно описан способ кодирования, а также декодирования с возможностью коррекции одиночных ошибок. На основе всего изложенного были написаны две программы на языке Python, реализующие операции кодирования и декодирования с использованием циклических кодов.