Оглавление

[02.09.2024 Лекция 1. Введение 2](#_Toc176789613)

[Постановка задачи помехоустойчивого кодирования 2](#_Toc176789614)

[Основные определения 2](#_Toc176789615)

[Декодирование с мягкими решениями 5](#_Toc176789616)

[Основные классы корректирующих кодов 6](#_Toc176789617)

[Выводы 6](#_Toc176789618)

[04.09.2024 Лекция 2. 7](#_Toc176789619)

[Кодирование текста 7](#_Toc176789620)

[Кодирование изображений 8](#_Toc176789621)

[Кодирование звука и видео 8](#_Toc176789622)

[Кодирование вершин куба 8](#_Toc176789623)

[Характеристический вектор множества 9](#_Toc176789624)

[09.09.2024 Лекция 3. Линейные коды 10](#_Toc176789625)

[Арифметика пространства двоичных последовательностей 10](#_Toc176789626)

[Основные свойства базиса и порождённого им пространства 11](#_Toc176789627)

[Порождающая и проверочная матрица 12](#_Toc176789628)

[Выводы 13](#_Toc176789629)

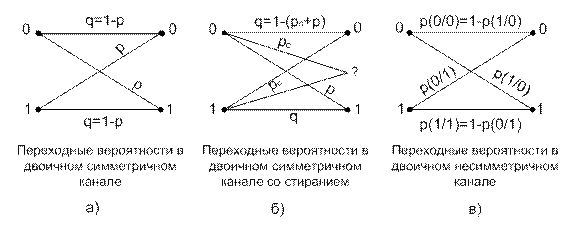
# 02.09.2024 Лекция 1. Введение

Теорию кодирования можно рассматривать как раздел теории информации, посвящённый проблеме достижения пределов скорости передачи информации, предугаданных Клодом Шенноном.

В 1948 г., когда, благодаря опубликованию работы Клода Шеннона появилась на свет теория информации, большая часть информации передавалась в аналоговом виде.

Предположим, что имеется канал, на вход которого можно подавать символы. Такой канал называют двоичным симметричным каналом (ДСК).

## Постановка задачи помехоустойчивого кодирования



## Основные определения

**Код** – список передаваемых последовательностей

**Кодовые слова** – сами последовательности

**Кодирование** – отображение передаваемых данных в кодовые слова

Вместо некоторого числа k по каналу передаётся количество сигналов, достаточное для передачи n бит, k < n. Отношение:

Называется скоростью передачи данных. То, что из общего числа возможных сигналов для передачи информации используется лишь , позволяет исправлять либо обнаруживать ошибки. Задача декодера – выбрать из числа кодовых слов одно, предположительно – то, что передавалось.

Таким образом, последовательности (кодовые слова), используемые для передачи сообщений, должны как можно сильнее отличаться друг от друга, иными словами, находиться на как можно большем расстоянии.

Для ДСК с переходной вероятностью пропускная способность C канала вычисляется по формуле:

Где

В реальной системе связи на вход канала поступают не нули и единицы, а сигналы, выбираеме в соответствии с физической природо канала(электрические импульсы, электпомагнитные волны, акустические колеюания и т.п). В этой более общей ситуации вход канала x(t) и выход канала y(t) могут быть связаны, например, соотношениями вида:

,

Где случайный процесс n(t) представляет собой шум в канале связи. Такой канал называют **непрерывным** (до этого мы рассматривали дискретный канал) и, если шум n(t) не зависит от сигналов (от x(t)), то шум называют **аддитивным.**

Энергия одного сигнала длительности T равна:

Предположим, что n(t) – стационарный случайный процесс, одномерное распредлеление которого – гауссово с нулевым математическим ожиданием и с дисперсией Спектральная плотность мощности предполагается равномерной равной (вт/Гц) в диапазоне частот [-W, W]. Такой гум называется аддитивным белым гауссовским шумом (АГБШ) и соответствующий канал – каналом АГБШ. В этих условиях пропускная способность канала, измеренная в битах, передаваемых в единицу времени, вычисляется по формуле:

бит/с.

При отсутствии ограничений на полосу, переходя к пределу при , получаем совсем простую формулу:

бит/с.

Сопоставим между собой характеристики системы связи без кодирования и системы с таким кодированием, которое по своим параметрам близко к теоретическим пределам

За время при T затратах энергии P T будет передано бит информации. Введем характеристику:

Отражающую затраты энергии на бит передаваемой информации.

Величину называют отношением сигнал/шум на бит.

Согласное формуле , надёжная передача возможна при

Или

дБ.

Итак, из теоремы кодирования для канала АГБШ следует: надёжная передача информации возможна только тогда, когда отношение сигнал/шум на бит не меньше величины -1,59 дБ.

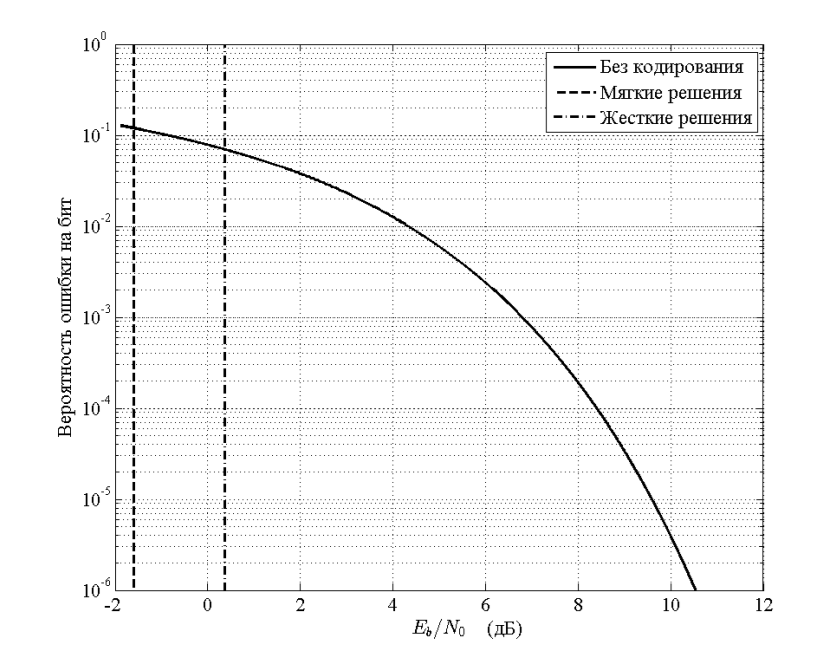


Рисунок 1 График зависимости вероятности ошибки от отношения сигнал/шум при передаче без кодирования

Энергетический выигрыш кодирования составляет 1.59 + 10.53 = 12.0 дБ, или в 16 раз!

Обозначим через вероятность ошибочного решения о сигнале энергии . Максимальная достижимая скорость кодирования в битах на символ не превышает пропускной способности ДСК. Тогда, с переходной вероятностью и согласно последней формуле, имеем:

Величину называют отношением сигнал/шум на сигнал

Принятое выше условие о бесконечной полосе означает, что при заданной скорости передачи в битах/с.

В итоге получаем

декодированию с жесткими решениями

## Декодирование с мягкими решениями

Спектральной эффективностью кодирования называется величина:

Граница снизу на отношение сигнал/шум при заданной спектральной эффективности:

Получаем ещё одну границу снизу на отношение сигнал/шум при заданной скорости используемого корректирующего кода:

Опираясь на теорию кодирования, при заданном критерии качества, можно уточнить оценку Шеннона для малых отношений сигнал/шум на бит. С учётом, что допустима некоторая вероятность ошибки. Эта поправка к пределу Шеннона впервые была использована Мак-Элисом.

В реальных системах используется в относительно широкополосных каналах связи (некоторых каналах мобильной связи, спутниковой связи, космической связи). соответствует узкополосным каналам (телефонная связь).

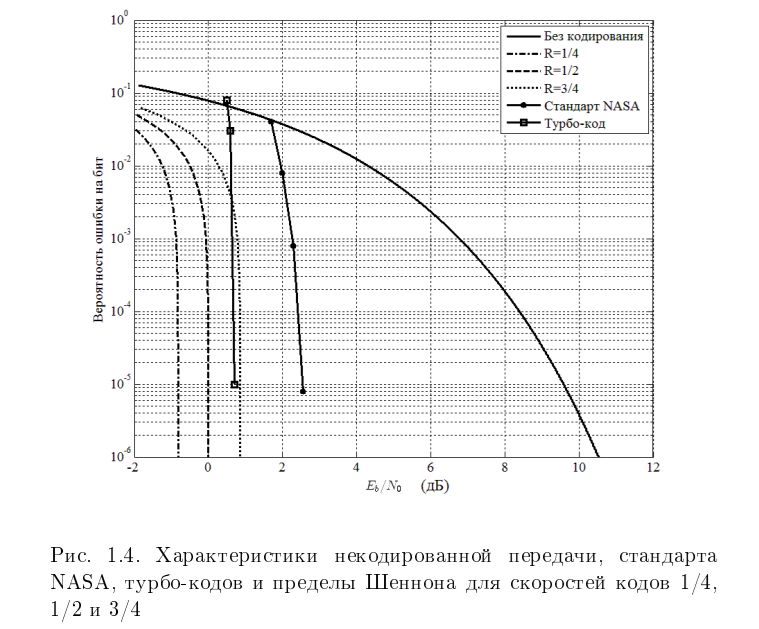


Рисунок 2 Характеристики некодированой передачи, стандарта NASA, турбо-кодов и пределы Шеннона для скоростей кодов ¼ , ½ и ¾ .

Кривая «Турбо-код» отражает характеристики турбо-кода длинной …

Код сложен, но имеет применение в международном сандарте спутникового цифрового телевидения DVBS2.

Таким образом, реально достижимый эффект от кодирования достигает 9-10 дБ.

## Обзор кодов для защиты информации от ошибок

### Основные классы корректирующих кодов

* Общая теория кодирования (построение кодов, границы)
* - Линейные блоковые коды
* - Коды Рида-Маллера
* - Нелинейные блоковые коды
* - Коды для криптографии
* - Каскадные конструкции
* Алгоритмы декодирования линейных блоковых кодов

- Мажоритарное декодирование

- Алгоритмы для каналов с мягкими решениями

* Циклические коды, декодирование, обобщения

- Методы декодирования

- Альтернативные коды, коды Гоппы, алгебро-геометрические коды

- Декодирование по алгоритму Гурусвами-Судана.

* Свёрточные коды и их декодирование
* Коды для итеративного декодирования

- Свёрточные низкоплотностные коды

- Полярные коды

* Коды для специальных условий применения

- Модуляционные коды

* - Коды для магнитной записи
* - Квантовые коды
* - коды для MIMO
* - коды для сетей связи

## Выводы

- Потенциально достижимая скорость передачи информации в канале связи в битах в секунду определяется ограничениями на мощность сигнала и полосу частот. Последнее ограничение может быть пересчитано в минимальную допустимую скорость кодирования в битах на сигнал.

- Применение кодов, исправляющих ошибки, обеспечивает большой энергетический выигрыш по сравнению с передачей без кодирования. Для обеспечения высокой эффективности использования ресурсов канала коды должны быть длинными и допускать декодирование с мягкими решениями.

- Важной характеристикой кода является его минимальное расстояние. Задача построения хороших кодов может быть сформулирована как задача построения кодов с наибольшим возможным минимальным расстоянием при заданной скорости.

- Минимальное расстояние — важный параметр кода, но не только оно принимается во внимание при выборе способа защиты от ошибок. С учетом ограничений на сложность реализации кодера и декодера при решении многих практических задач лучшими оказываются коды с относительно небольшим минимальным расстоянием, но зато допускающие простое декодирование.

# 04.09.2024 Лекция 2.

Теория кодирования – раздел теории информации, изучающий способы отождествления сообщений с отображающими их сигналами.

Задачей является согласование источника информации с каналом связи.

Объектом кодирования служит как дискретная, так и непрерывная информация, а понятия кодирования означает преобразование информации в форму, удобную для передачи по определённому каналу связи.

Обратная операция – декодирование (восстановление принятого сообщения в доступный для потребителя формат).

В теории кодирования существует ряд направлений:

* Статическое (эффективное) кодирование;
* Помехоустойчивое кодирование;
* Корректирующие коды;
* Циклические коды;
* Арифметические коды.

## Кодирование текста

Текст – последовательность символов, входящих в некоторый алфавит. Кодирование текста сводится к кодированию двоичного алфавита. Тогда, мощность алфавита составляет 256 символов. Пример кода – таблица ASCII (1963 г.).

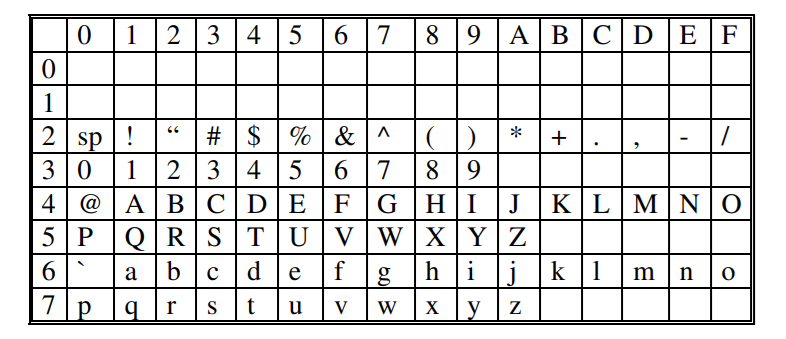


Рисунок 3 Таблица кодов ASCII (American Standard Code for Information Interchange) представлены код от 00 до 7F (часть).

В 1991 г. На смену пришла универсальная система 16-разрядного кодирования символов Unicode. Мощность алфавита: возможных символов (65 536 различных кодов), из которых 63 484 соответствует символам большинства алфавитов, а оставшиеся 2048 кодов разделены пополам и образуют таблицу размером 1024х1024. Это место оставлено для мёртвых языков и дополнительных символов.

Для представления необходима пара 16-разрядных слов.

## Кодирование изображений



Рисунок 4 Пример двухцветная картинка (черно-белой) представленных совокупность отдельных точек, расставленных на прямоугольной решетке.

Чёрно-белые изображения принято представлять в градациях серого цвета. Для этого используют модель GreyScale. Яркость точки кодируется 1 байтом. Для цветных используют модель RGB.

## Кодирование звука и видео

Аналитически метод кодирования, применимый к любым звуковым сигналам, основан на аналогово-цифровом преобразовании. Исходный аналоговый сигнал представляют, как последовательность цифровых сигналов в двоичном коде. Разрядность преобразования определяет объём данных, соответствующих отдельному цифровому сигналу. При воспроизведении звука выполняет обратное – цифро-аналоговое преобразование. Метод не лишён погрешности.

Метод кодирования на основе табличного синтеза. Применяется к музыкальным произведениям. В таблицах хранятся образцы (сэмплы) звуков различных музыкальных инструментов. Числовые коды определяют инструмент, ноту и продолжительность звучания.

## Кодирование вершин куба

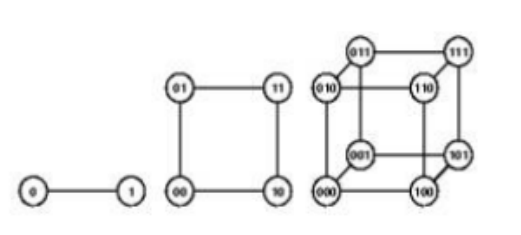


Рисунок 5 Представление кубов с пронумерованными вершинами для k=1, 2, 3

Последовательностью из k нулей можно обозначить вершины k-мерного куба.

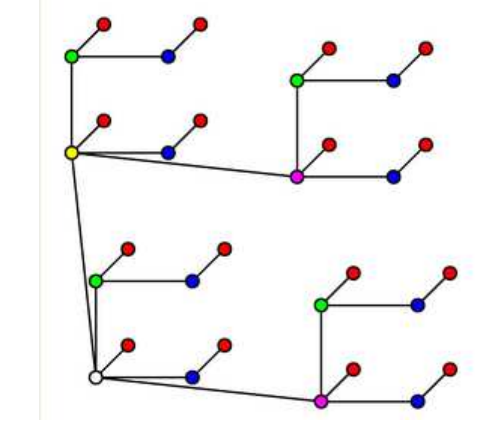


Рисунок 6 Изображение 5-мерного куба, в котором проведены только ребра, необходимые для минимальной связи между вершинами

Каждый процессор кодируется набором из 5 символов от 0 до 1. Процессор, в ходе которого последний символ 1, связывается с процессором, у которого код отличается только в последнем символе. Процессор, код которого заканчивается на 1 0, прикрепляется к процессору с кодом, отличающимся во втором с конца символе и т.д.

## Характеристический вектор множества

Пусть , тогда множество состоит из последовательностей 0 и 1 длины n. Они называются строкой бит и применяются для моделирования операций на конечных множествах.

Пусть . Если , то поставим подмножеству А строку , где , тогда строка бит называется характеристическим вектором подмножества А.

Пример:

Характеристическим вектором , а характеристическим вектором .

# 09.09.2024 Лекция 3. Линейные коды

**Коды** – это наборы кодовых слов.

**Эффективность кода** определяется его длинной и большим минимальным расстоянием при заданной скорости, рациональной сложностью кодирования и декодирования

Списки кодовых слов слишком велики, чтобы можно было сохранять их в памяти кодера и декодера.

В этой связи, первый шаг к сокращению сложности описания кодов – построение линейных кодов, т.е. кодов, слова которых образуют линейное пространство. Для описания линейного кода достаточно задать базис пространства, а кодирование сводится к умножению на матрицу.

## Арифметика пространства двоичных последовательностей

Множество чисел {0, 1} образует простейшее поле из двух элементов.

Таблицы сложения и умножения в поле :

|  |  |  |
| --- | --- | --- |
| 4 | 0 | 1 |
| 0 | 0 | 1 |
| 1 | 1 | 0 |

|  |  |  |
| --- | --- | --- |
| \* | 0 | 1 |
| 0 | 0 | 0 |
| 1 | 0 | 1 |

Для произвольного множества A обозначим множество последовательностей длинны n через , составленных из элементов множества А.

Определим сумму векторов как покомпонентно:

Умножение вектора x на скалярную величину определим, как:

Множества векторов, замкнутое относительно операций сложения и умножения на скаляр, называется **линейным векторным пространством**.

Примеры двоичных линейных пространств:

Множество всех двоичных последовательностей длинны n являются линейным пространством. Такие пространства содержат, в общем случае, векторов.

Пространство примера содержится в множестве , но не совпадает с ним, т.е. является его линейным подпространством.

Каждое пространство содержит нулевой вектор – это свойство линейного пространства.

**Базисом** линейного пространства называется наибольшее множество линейно независимых векторов пространства.

Число базисных вектором называется **размерностью пространства**.

Для линейных пространств примера имеем:

А) Размерность равна 2, пример базиса:

Б) Размерность равна 3, примеры базисов: .

В) Размерность равна 2, пример базиса:

## Основные свойства базиса и порождённого им пространства

* Любой вектор пространства единственным образом может быть представлен в виде линейной комбинации базисных векторов:
* Число элементов векторного пространства размерности k равно

**Линейным q-ичным (n, k) кодом C** называется любое k-мерное подпространство всевозможных векторов длинны n.

**Скоростью линейного (n, k) кода**  называется отношение:

. В двоичном случае скорость измеряется в битах на символ канала.

Мы уже замечали что хорошим будет код, слова которого «далеки» друг от друга в смысле расстояния Хэмминга, которое для последовательностей x, y длинны n определяется как

Где:

Расстояние Хэмминга удовлетворяет аксиомам расстояния:

* если и только если ;
* (симметричность)
* (неравенство треугольника)

Таким образом, имеем метрическое пространство.

**Минимальным расстоянием кода С** называется наименьшее из попарных расстояний между кодовыми словами:

**Теорема.** Код с минимальным расстоянием d исправляет любые комбинации ошибок кратности .

**Теорема.**

Где – число ненулевых элементов в последовательности x или вес вектора x в метрике Хэмминга.

Экономию памяти возможно реализовать в процессе кодирвания, запоминая базис, а не всё пространство.

Пример:

Для кода С вместо слов достаточно хранить 500 базисных слов длины 1000. Для этого нужно 500 Кбит памяти. Цифра завышена. Можно использовать циклические коды, в которых слова являются циклическим сдвигом друг друга, тогда можно будет хранить всего 1 слово, т.е. всего 1000 бит. Но и эта цифра достаточно высокая, её тоже можно сократить.

## Порождающая и проверочная матрица

**Порождающей матрицей линейного (n, k) кода** называется матрица размера , строки которой – его базисные векторы.

Например, матрица является порождающим для кода В) из примера выше.

Сообщение записывается в виде вектора и соответствующее сообщению кодовое слово вычисляется по формуле:

Слова могут быть получены умножением вектора на матрицу. Вектор из k бит превращается в последовательность из n двоичных кодовых символов, передаваемых по каналу или записываемых в памяти.

Предположим, что можно упростить декодирование линейного кода.

Предположим, что для некоторого двоичного вектора все кодовые слова (n, k) кода удовлетворяют тождеству , в котором запись (x, y) обозначает скалярное произведение векторов x и y.

Вектор h ортогонален коду, назовём его **проверкой** по отношению к коду.

справедливо для всех кодовых слов, если оно справедливо для базисных векторов, т.е. если , где верхний индекс обозначает транспонирование.

**Пространство проверок** назовём пространством, ортогональным линейному коду или проверочным пространством.

Список номеров k линейно независимых столбцов матрицы G – такой набор индексов назовём **информационной совокупностью**.

Совокупность индексов остальных позиций назовём **проверочной совокупностью**, а сами позиции и записанные на них символы будем называть **информационными и проверочными** соответственно. Выберем произвольно и зафиксируем значения вектора h на позициях проверочной совокупности. Т.к. значения вектора h на проверочных позициях можно выбрать способами, то размерность пространства будет равна r.

**Теорема.** Размерность проверочного пространства линейного (n, k) кода равна .

Число r называется **избыточностью** кода.

Базис проверочного пространства запишем в виде матрицы:

Называемой **проверочной матрицей** кода.

Проверочная и порождающая матрицы связаны соотношением:

Из этого соотношения мы видим, что для любого кодового слова имеет место:

Это тождество можно использовать как критерий принадлежности произвольной последовательности коду, т.е. для обнаружения ошибок.

Код можно задать разными порождающими матрицами, заменив в матрице G любую строку на любую линейную комбинацию этой строки с другими строками. Получаем новую порождающую матрицу того же кода.

Перестановка столбцов матрицы G приводит к другому коду, который по своим характеристикам не отличается от исходного. Коды, различающиеся только нумерацией позиций, называются **эквивалентными**.

Для каждого кода найдётся такой эквивалентный ему код, что первые k позиций кодовых слов образуют информационную совокупность, иными словами, первые k столбцов порождающей матрицы образуют невырожденную подматрицу размера . Заменяя строки линейными комбинациями строк (метод Гаусса), полученную матрицу можно принести к виду

Где – единичная матрица порядка k, а P – некоторая матрица размера .

Матрица вида (2.7) называется порождающей матрицей, **приведённой к систематическому виду**, а соответствующий код называется **систематическим.** Кодирование для систематического кода проще, чем для кода общего вида:

т.е. в кодовом словае первые k позиций – просто копия информационной последовательности M, а остальные r – проверочных.

## Выводы

- линейный (n, k) – код может быть задан люой из двух матриц: порождающией G размера либо проверочной H размера . По G легко найти H приведением кода к систематической форме. Аналогично по H находится G.

- матрица G используется при кодировании матрицей H можно воспользоваться при декодировании, по меньшей мере, для обнаружения ошибок. Выполнение тождества свидетельствует о том, что данная последовательность принадлежит коду.

# 10.09.2024 Лекция 3. Вычисление минимального расстояния по проверочной матрице

Формула нахождения минимального расстояния в терминах порождающей матрицы записывается как

Вычисление по этой формуле требует перебора по ненулевым словам кода. Отметим, что на практике мы используем коды, минимальное расстояние которых никому неизвестно. Заметим, что при , т.е. когда скорость кода больше проверочная матрица имеет меньший размер, чем порождающая. Воспользуемся этим обстоятельством для упрощения поиска минимального расстояния кода.

Согласно формуле кодовому слову с соответствует вес и некоторый набор из столбцов проверочной матрицы , сумма которых – это нулевой вектор. Следовательно, для минимального расстояния кода нужно найти минимальный набор линейно независимых столбцов.

**Дуальным** к данному коду называется код, порождающая матрица которого является проверочной матрицей данного кода.

**Теорема 2.4.** Минимальное расстояние линейного кода равно в том и только в том случае, когда любые столбцов проверочной матрицы линейно независимы и существует набор из линейно независимых столбцов.

**Теорема 2.5.** Граница Синглтона: минимальное расстояние линейного кода удовлетворяет неравенству:

Доказательство. Ранг проверочной матрицы равен , следовательно, столбцов всегда линейно независимы. Граница Синглтона следует из Теоремы 2.4.

## Синдромное декодирование

Пусть и обозначают порождающую и проверочную матрицу (n, k) кода и y – двоичную послеовательность, подлежащую дкодированию. Если передавалось кодовое слово , то y можно представить в виде

Где вектор называют **вектором ошибок.** Число передаваемых комбинаций ошибок не зависит от кода и равно .

Для линейных кодов объём множества последовательностей , декодируемых в слово , одинаков для всех слов . Объём можно посчитать отношением:

Обозначим через множество исправляемых комбинаций ошибок, тогда модуль этого множества равен . Тогда, хороший код отличается от плохого множеством исправляемых комбинаций, по своей форме близкой к Хэмминговой сфере. При декодировании по минимум расстоянии необходимо обеспечивать исправление всех комбинаций ошибок множества . Т.к. для кодового слова выполняется равенство:

, то декодирование со сложностью состоит в переборе по множеству ошибок комбинаций, всякий раз прибавляя предполагаемый вектор ошибок к и вычисляя произведение . Вектор , для которого будет получен нулевой результат – результат декодирования.

Недостаток: хранение множества и умножение на проверочную матрицу для каждой последовательности множества .

Усовершенствованный способ декодирования:

Рассмотрим произведение

Называемое **синдромом** вектора . Неравенство синдрома нулевому вектору указывает на наличие ошибок в принятой последовательности .

Поскольку

Имеет место взаимно однозначное соответствие между различными синдромами и исправляемыми комбинациями ошибок.

Это означает, что умножение на матрицу достаточно выполнить 1 раз – при вычислении синдрома . В памяти декодера хранится соответствие между синдромами и и векторами ошибок : для каждого хранит вектор наименьшего веса , для которого синдром *.*

Декодер вычисляет и по таблице находит соответствующий вектор ошибок e, а результатом декодирования служит слово . Процедура декодирования представлена алгоритмом:

* Input: вход канала y;
* Output: информационные символы;
* Вычислить синдром ;
* Извлечь из ячейки памяти с адресом вектор ошибок ;
* Вычислить оценку кодового слова ;
* Вывод: Информационные символы, соответствующие кодовому слову .

## Радиус покрытия и декодирование по минимуму расстояния Хемминга

Два способа декодирования для произвольно линейного кода:

* Перебор по кодовым словам;
* Перебор по векторам ошибок;

Удобно характеризовать алгоритмы этого класса асимптотическим показателем экспоненты сложности

Описанные выше алгоритмы приводят нас к оценке

для любого линейного кода со скоростью .

## Радиус покрытия

**Радиус покрытия** – наибольшее число ошибок, которые может исправить код.

Для произвольного множества последовательностей длины n введём обозначение для расстояния между последовательнотсью и A:

**Радиусом покрытия** кода C назыается величина

Иными словами, для линейного (n, k) кода представляет собой наибольший вес среди весов всех последовательностей, декодируемых в нулевое кодовое слово.

Практическое применение метода связано с кодированием двоичного источника с заданным уровнем ошибки в метрике Хэмминга. Если радиус покрытия равен , то максимальная ошибка приближения любой последовательности словами кода не превысит .

Задачи построения кода с заданной кратностью гарантированно исправляемых ошибок называют задачами ещё упаковки пространства шарами радиуса с центрами, задаваемыми словами кода. Построение кодов с заданным радиусом покрытия – это задача покрытия пространства шарами радиуса.

В первом случае строится пространство с максимальным числом шаров (кодовых слов). Во втором случае строятся покрытия минимальным числом этих слов.

Коды, удовлетворяющие границе сферической упаковки, являются оптимальными для такого рода задач. К ним относятся коды Хэмминга и Голея.

Запишем некоторые свойства радиуса покрытия.

**Свойство 2.1.** Декодирование в ДСК с исправлением всех ошибок кратности включительно эквивалентно декодированию по минимуму расстояния Хэмминга.

**Свойство 2.2.** Пусть обозначает минимальный вес вектора такого, что ему соответствует синдром . Тогда

**Свойство 2.3.** Радиус покрытия линейного кода не превышает числа проверочных символов кода .

**Свойство 2.4.** Если радиус покрытия линейного кода , то размерность кода может быть увеличена без уменьшения минимального расстояния кода.

**Свойство 2.5.** Если (n, k) код имеет наибольшую возможную размерность k для заданных длины n и минимального расстояния d, то имеет место неравенство .

**Теорема.** При достаточно большой длине (n, k) кода со скоростью для любого в ансамбле равновероятных случайных кодов существуют коды с радиусом покрытия таким, что

Иными словами, существуют коды, радиус покрытия которых удовлетворяет границе Варшамова-Гилберта.

## Декодирование по соседям нулевого слова

Этот остроумный алгоритм декодирования предложен в

Levitin L., Hartmann C. A new approach to the general minimum distance decoding problem: The zero-neighbors algorithm. IEEE Transactions on Information Theory, 31(3):378–384, 1985.

Выигрыш метода декодирования по соседям нулевого слова по экспоненте сложности в сравнении с непосредственным перебором очень большой. Тем не менее, декодирование в СК может быть выполнено ещё проще. Оценка сложности интересна тем, что точно такое е поведение сложности наблюдается для декодирования кодов, представленных решётками, но уже не с жёсткими решениями, а с мягкими решениями в канале с АГБШ.

## Декодирование по информационным совокупностям

Идея ещё одного метода декодирования основана на том, что по значениям кодовых симолов на любой информационной совокупности однозначно восстанавливается кодовое слово. Согласно свойству, радиус покрытия не превышает числа проверочных символов. Поэто , если число ошибок в канале не превышает вличины радиуса покрытия кода, то, вероятно, найдётся свободная от ошибок информационная совокупность, по которой должно восстановить кодовое слово. Вывод формулы асимптотической сложности декодирования по информационным совокупностям был впервые опубликован в работах.

Coffey J. T., Goodman R. M. The complexity of information set decoding. IEEE Transactions on Information Theory, 36(5):1031–1037, 1990.

Крук Е. А. Граница для сложности декодирования линейных блоковых кодов. Проблемы передачи информации, 25(3):103– 107, 1989.

## Вывод

* Линейный код – линейное подпространство линейного пространства
* Порождающая матрица кода – базис подпространства, образующий код.
* Проверочная матрица – базис ортогонального пространства.
* Любой код либо может быть представлен в систематической форме, либо эквивалентен систематическому коду.
* Преобразование порождающей матрицы в проверочную и наоборот выполняется с помощью приведения кода к систематической форме.
* Минимальное расстояние кода равно минимальному весу ненулевых кодовых слов.
* Минимальное расстояние кода равно минимальному числу линейно зависимых столбцов.
* Параметры линейного кода удовлетворяют границе Синглтона
* Декодирование линейного (n, k) кода по минимуму расстояния Хэмминга при скорости кода может быть выполнено перебором по кодовым словам со сложностью. Пропорциональной . При при помощи синдромного декодирования оно может быть выполнено со сложностью пропорциональной .

# 11.09.2024 Лекция 4. Код, длина кода, основание кода

Кодер – это правило (совокупность правил), в соответствие с которыми производится отображение дискретных сообщений сигналами в виде определенных сочетаний символов вторичного алфавита.

Источник выдает некоторое дискретное сообщение a, которое представляет собой последовательность элементарных сообщений , где

Элементарные сообщения – это символы сообщений, а их совокупность – алфавит источника.

Кодирование заключается в том, что последовательность символов источника a заменяется последовательностью кодовых символов, т.е. кодовых комбинаций (кодовых слов)

Общее число символов, составляющих кодовую комбинацию, называют значностью или длиной кода n.

Количество значений кодовых признаков, используемых в кодовых комбинациях, называют основанием кода r.

Пример: на рисунке приведена кодовая комбинация с длиной кода n = 5. В качестве импульсного признака используется величина импульсов , тогда основание кода m = 3

Задание: определить длину кода n и основание кода m. Если кодирование представляет собой перевод десятичного число 20 в двоичное число.

Обратимое и необратимое кодирование, условие обратимости кодирования

Декодирование можно рассматривать как функцию обратную функции F (кодирование). Данные операции называются обратимыми, если их последовательное применение не приводит к потери информации.

Различают:

* принципиально необратимое кодирование (хэширование);
* обратимое с помощью дополнительной информации (ключа шифрования);
* безусловное обратимое кодирование: хранение паролей.

Условие обратимости кодирования в формализованном виде , где – это количество информации исходного сообщения, состоящего из первичного алфавита A, – количество информации в том же сообщении, записанном в символах вторичного алфавита B, то есть количество информации после кодирования.

Правила при построении обратимого кода:

разным символам первичного алфавита A должны быть сопоставлены разные кодовые комбинации;

никакая кодовая комбинация не должна составлять начальную часть другой кодовой комбинации.

Определите возможный код символа «Г» в первичном алфавите , если в качестве вторичного алфавита используется двоичный алфавит и известны следующие коды: А – 00, Б – 010, В – 101: Г = {01; 10; 000; …}.

### Асимптотический оптимальный код, теорема Шеннона № 1

Для некоторого первичного алфавита A и некоторого вторичного B асимптотическим оптимальным кодом будет называться такой способ кодирования, при котором избыточность кодирования стремится к 0, если длина сообщения стремится к ∞.

Относительная избыточность кода: .

### Теорема Шеннона № 1 «Основная теорема кодирования при отсутствии шумов»

при отсутствии шумов всегда возможен такой вариант кодирования сообщения, при котором относительная избыточность кода будет сколь угодно близка к 0.

### Классификация способов кодирования

Критерии классификации:

* условия построения кодовых комбинаций:
* равномерные коды: все сообщения передаются кодовыми группами с одинаковым числом элементов и длинна кода неизменна, обладает большими возможностями с точки зрения помехозащищённости передачи (телеграфный код Бодо);
* неравномерные коды: сообщения передаются кодовыми группами, содержащими неодинаковое число элементов, обеспечивают наибольшую экономичность построения и быстродействие передачи сообщений, используются при статистическом кодировании, при передаче необходимо наличие разделяющего символа (код Морзе).
* число различных символов в кодовых комбинациях:
* единичные код (числоимпульсные): использует одинаковые символы, а комбинации отличаются лишь их количеством, простой неравномерный код – помехозащищённость низкая, не используется для передач информации по каналам связи, а только для преобразований сигнала (машина Поста, машина Тьюринга, цифровые электронные счётчики);
* двоичные коды;
* многопозиционные коды: алфавит состоит из большого числа символов.
* форма представления в канале передачи данных:
* последовательные коды: элементарные сигналы (кодовая комбинация) посылаются в канал передачи последовательно во времени и разделены между собой определённым временным интервалом;
* параллельные коды: элементарные сигналы посылаются одновременно по разным электрическим сетям, число которых соответствует элементов кода.
* возможность обнаружения и исправления ошибок:
* простые коды: всевозможные кодовые комбинации и ошибка в одном элементе кода приводит к неверной регистрации всего передаваемого сообщения;
* корректирующие коды.
* число одновременно кодируемых символов первичного алфавита:
* алфавитное кодирование;
* блочный метод кодирования: понижает избыточность.