

Теория кодирования

Михайлов Максим

21 октября 2022 г.

Оглавление

Лекция 1	2 сентября	4
1	Введение	4
1.1	Способы снижения вероятности ошибки	6
1.2	Понятие кода	7
1.3	Теоремы кодирования	7
1.4	Пропускная способность некоторых каналов	8
Лекция 2	9 сентября	9
1.5	Энергетический выигрыш кодирования	9
2	Жесткое и мягкое декодирование	9
2.1	Критерии декодирования	10
Лекция 3	16 сентября	11
2.2	Покрытия	12
2.3	Дуальные коды	13
2.4	Граница Хэмминга	13
2.4.1	Асимптотическая оценка	13
2.5	Граница Варшамова–Гилберта	13
2.6	Граница Варшамова–Гилберта для линейных кодов	14
2.7	Граница Грайсмера	14
Лекция 4	23 сентября	15
3	Универсальные методы декодирования линейных кодов	15
3.1	Метод порядковых статистик	16
3.2	Декодирование по решеткам	16
4	Декодирование с мягким выходом	18
4.1	Алгоритм Бала–Коке–Елинека–Равива	18
Лекция 5	30 сентября	19
4.2	Выводы	19
5	Сверточные коды	20
Лекция 6	7 октября	21
6	Методы модификации и комбинирования кодов	21
6.1	Конструкция Плоткина	21
6.1.1	Коды Рида–Маллера	22
6.2	Модификации кодов	22
6.2.1	Укорочение	22
6.2.2	Выкалывание	22

6.2.3	Расширение	22
6.3	Каскадные коды	23
6.4	Прямое произведение кодов	23
6.4.1	Пример: код Рао-Редди (48, 31, 8)	23
6.5	Лестничные коды	24
6.6	Граница Зяблова	24
6.7	Обобщенные каскадные коды (Блох, Зяблов, Зиновьев)	24
6.8	Турбо-коды	24
6.9	Построение перемежителей	25
Лекция 7	14 октября	26
7	Полярные коды	26
7.1	Некоторые определения	26
7.1.1	Функция переходных вероятностей канала	26
7.1.2	Параметр Бхаттачарьи	27
7.2	Поляризация канала	28

Лекция 1

2 сентября

1 Введение

Определение. Передаваемый сигнал это

$$x(t) = \sum_i S_{x_i}(t - iT)$$

, где x_i — передаваемые символы, T — продолжительность передачи одного символа.

Пример (M -ичная амплитудно-импульсная модуляция).

$$S_i(t) = \alpha(2i + 1 - M)g(t) \sin(2\pi ft)$$

, где:

- $g(t)$ — сигнальный импульс
- f — несущая частота
- α — коэффициент энергии передаваемого сигнала

Модели канала:

1. В непрерывном времени: $y(t) = x(t) + \eta(t)$
2. В дискретном времени: $y_i = \alpha(2x_i + 1 - M) + \eta_i$

η — белый шум, обычно Гауссов $\mathcal{N}(0, \sigma^2)$.

У передаваемого сигнала обычно не должно быть постоянной компоненты (по причинам физики), поэтому сигнал симметричен. С точки зрения теории кодирования это несущественно.

Приемник наблюдает на выходе канала вектор $y = (y_0 \dots y_{n-1})$. Канал характеризуется условным распределением $P_{Y|X}(y | x)$, где X, Y — случайные величины, соответствующие векторам переданных и принятых символов.

Приемник разбивает векторное пространство на решающие области $R_x : y \in R_x \implies \hat{x} = x$, т.е. если сигнал попал в область R_x , то мы ему сопоставляем кодовый символ x . Тогда вероятность ошибки:

$$\begin{aligned} P_e &= \int_{\mathbb{R}^N} P_e(y) p_Y(y) dy \\ &= \sum_x \int_{R_x} p_e(y) p_Y(y) dy \\ &= \sum_x \int_{R_x} (1 - p_{X|Y}\{x | y\}) p_Y(y) dy \\ &= 1 - \sum_x \int_{R_x} p_{X|Y}\{x | y\} p_Y(y) dy \end{aligned}$$

Приемник должен найти оптимальные решающие области. Оптимальность определяется критерием, например:

1. Критерий максимума апостериорной вероятности (критерий идеального наблюдателя)

$$\begin{aligned} R_x &= \{y | p_{X|Y}(x | y) > p_{X|Y}(x' | y), x' \neq x\} \\ &= \{y | P_X(x) p_{Y|X}(y | x) > P_X(x') p_{Y|X}(y | x'), x' \neq x\} \end{aligned}$$

2. Критерий максимума правдоподобия

$$R_x = \{y | p_{Y|X}(y | x) > p_{Y|X}(y | x'), x' \neq x\}$$

В случае равновероятных символов этот критерий совпадает с критерием идеального наблюдателя.

Пример (2-ичная амплитудно-импульсная модуляция (2-AM)). Пусть $y_i = \alpha(2x_i - 1) + \eta_i$, $\eta_i \sim \mathcal{N}(0, \sigma^2)$, $x_i \in \{0, 1\}$. Тогда:

$$p_{Y|X}(y | x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y - \alpha(2x-1))^2}{2\sigma^2}}$$

Применим критерий максимального правдоподобия:

$$R_0 = \{y | y < 0\}, R_1 = \{y | y \geq 0\}$$

Вычислим вероятность ошибки:

$$P_e = P_X(0)P\{Y \geq 0 | X = 0\} + P_X(1)P\{Y < 0 | X = 1\}$$

$$\begin{aligned}
&= 0.5 \int_0^\infty p_{Y|X}(y | 0) dy + 0.5 \int_{-\infty}^0 p_{Y|X}(y | 1) dy \\
&= \int_0^\infty \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y+\alpha)^2}{2\sigma^2}} dy \\
&= \int_\alpha^\infty \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{y^2}{2\sigma^2}} dy \\
&= \int_{\frac{\alpha}{\sigma}}^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy \\
&=: Q\left(\frac{\alpha}{\sigma}\right) \\
&= \frac{1}{2} \operatorname{erfc}\left(\frac{\alpha}{\sqrt{2}\sigma}\right)
\end{aligned}$$

Значение сигнала это обычно уровень напряжения. Как мы знаем из школьной физики, мощность $P = \frac{U^2}{R}$. Мы хотим минимизировать мощность, чтобы экономить электроэнергию. Мощность сигнала суть случайная величина с матожиданием, пропорциональным $E_S = \alpha^2$. Мощность белого шума не зависит от частоты и пропорциональна $\sigma^2 = \frac{N_0}{2}$. Если же шум зависит от частоты, то он называется розовым или голубым.

Соотношение мощностей сигнал/шум на символ это $\frac{E_S}{N_0}$, обычно измеряемое в децибелах, т.е. $10 \log_{10} \frac{E_S}{N_0}$. Однако нас интересуют не символы, а биты и тогда соотношение сигнал/шум на бит это $\frac{E_S}{RN_0}$, где R — количество бит информации, представленных одним символом.

1.1 Способы снижения вероятности ошибки

1. Посимвольное повторение: будем передавать вместо каждого символа m копий того же символа.

$$y_{mi+j} = \alpha(2x_i - 1) + \eta_{mi+j}, \quad 0 \leq j < m$$

Рассмотрим разные способы работы приемника:

- Для каждого y_{mi+j} проведем голосование. Тогда вероятность ошибки:

$$P_v(m) = \sum_{j=\lceil \frac{m}{2} \rceil}^{m-1} C_m^j P_e^j (1 - P_e)^{m-j}$$

- Примем решение по

$$\sum_{j=0}^{m-1} y_{mi+j} = m\alpha(2x_i - 1) + \sum_{j=0}^{m-1} \eta_{mi+j}$$

, т.е сложим наблюдения в одном блоке. Тогда вероятность ошибки:

$$P_a(m) = Q\left(\frac{m\alpha}{\sqrt{m}\sigma}\right) = Q\left(\sqrt{2\frac{mE_s}{N_0}}\right) = Q\left(\sqrt{2\frac{E_b}{N_0}}\right)$$

Выигрыша не будет, т.к. на один символ передается в m раз меньше бит (см. последний переход).

Второй метод лучше первого, т.к. мы не теряем информацию о нашей уверенности в каждом принятом символе.

Избыточность на уровне битов не дала улучшения, поэтому введем избыточность на уровне блоков.

1.2 Понятие кода

Определение. Код — множество допустимых последовательностей символов алфавита X .

Последовательности могут быть конечными или бесконечными, но на практике только конечными. Не каждая последовательность символов алфавита является кодовой.

Определение. Кодер — устройство, отображающее информационные последовательности символов алфавита \mathcal{B} в кодовые.

Различным последовательностям алфавита \mathcal{B} сопоставляются различные последовательности алфавита X для однозначности кодирования.

Определение. Скорость кода — отношение длин информационной и кодовой последовательностей.

Определение. Декодер — устройство, восстанавливающее по принятой последовательности символов наиболее вероятную кодовую последовательность.

1.3 Теоремы кодирования

Пусть для передачи используется код $\mathcal{C} \subset X^n$ длины n , состоящий из M кодовых слов, выбираемых с одинаковой вероятностью.

Теорема 1 (обратная). Для дискретного постоянного канала с пропускной способностью C для любого $\delta > 0$ существует $\varepsilon > 0$ такое, что для любого кода со скоростью $R > C + \delta$ средняя вероятность ошибки $\overline{P}_e \geq \varepsilon$

Теорема 2 (прямая). Для дискретного постоянного канала с пропускной способностью C для любых $\varepsilon, \delta > 0$ существует достаточно большое число $n_0 > 0$, такое что для всех натуральных $n \geq n_0$ существует код длиной n со скоростью $R \geq C - \delta$, средняя вероятность ошибки которого $\overline{P}_e \leq \varepsilon$

1.4 Пропускная способность некоторых каналов

1. Двоичный симметричный канал: $X, Y \in \{0, 1\}$, $p_{Y|X}(y | x) = \begin{cases} p, & y \neq x \\ 1 - p, & y = x \end{cases}$. Из теории информации:

$$C_{BSC} = 1 + p \log_2 p + (1 - p) \log_2 (1 - p)$$

2. Идеальный частотно ограниченный гауссовский канал: $y(t) = x(t) + \eta(t)$, $\eta(t)$ — гауссовский случайный процесс, спектральная плотность мощности которого равна $S(f) = \begin{cases} \frac{N_0}{2}, & -W < f < W \\ 0, & \text{иначе} \end{cases}$ Из теории случайных процессов:

$$C_{AWGN} = W \log_2 \left(1 + \frac{E_S}{W N_0} \right)$$

$$\lim_{W \rightarrow \infty} C_{AWGN} = \frac{E_S}{N_0 \ln 2}$$

Лекция 2

9 сентября

1.5 Энергетический выигрыш кодирования

Энергетический выигрыш кодирования показывает, во сколько раз использование кодирования позволяет уменьшить отношение сигнал/шум, необходимое для достижения заданной вероятности ошибки по сравнению со случаем отсутствия кодирования.

Выигрыш от кодирования неидеален, т.к.:

1. Оптимальный код сложно придумать (прямая теорема кодирования не конструктивна)
2. Используются коды конечной (и обычно малой) длины
3. Алгоритмы декодирования часто декодируют не по максимуму правдоподобия, а как получится
4. Дискретизация выхода канала

Теория кодирования занимается увеличением энергетического выигрыша кодирования.

2 Жесткое и мягкое декодирование

- При **мягком декодировании** декодер непосредственно использует y_i , т.е. есть информация о надежности символов.
- При **жестком декодировании** декодер использует только оценки \hat{x}_i .

Пример. Канал с аддитивным белым гауссовским шумом (АБГШ): $y_i = (2x_i - 1) + \eta_i$, $x_i \in \{0, 1\}$

В случае жесткого декодирования канал превращается в двойной симметричный канал с вероятностью ошибки $p = Q\left(\sqrt{2\frac{E_s}{N_0}}\right)$

Примечание. При малых x $Q(x) \approx \frac{1}{2} - \frac{x}{\sqrt{2\pi}}$

Достижимая скорость передачи данных:

$$R < \frac{2}{\ln 2} \left(\sqrt{\frac{E_s}{N_0 \pi}} \right)^2$$

$$\frac{E_b}{N_0} = \frac{E_s}{RN_0} > \frac{\pi}{2} \ln 2 = 1.09 = 0.37$$

Как мы уже выяснили, для мягкого декодирования это значение -1.59 , жесткое декодирование хуже.

Пример. Рассмотрим идеальный частотно ограниченный гауссовский канал, т.е. спектральная плотность мощности отличается от нуля только в некотором диапазоне частот.

Спектральная эффективность кодирования — $\beta = \frac{R}{W}$, скорость передачи данных, деленная на ширину канала.

$$\frac{R}{W} < \log_2 \left(1 + \frac{R}{W} \frac{E_b}{N_0} \right)$$

$$\frac{E_b}{N_0} > \frac{2^\beta - 1}{\beta}$$

2.1 Критерии декодирования

Оптимальные критерии декодирования сложны в реализации, сведем их к чему-нибудь попроще.

1. **Критерий минимального расстояния:** кодовое слово $c = \operatorname{argmin}_{c \in C} d(c, y)$
2. **Списочное декодирование:** ищем все кодовые слова, находящиеся в сфере заданного радиуса вокруг полученного вектора.
3. **Побитовое декодирование:** используется критерий идеального наблюдателя для отдельных символов кодового слова. Часто также дополняется вычислением логарифма отношений правдоподобия для отдельных символов.

$$L_i = \ln \frac{\sum_{c_i=0}^{c \in C} P\{c | y\}}{\sum_{c_i=1}^{c \in C} P\{c | y\}}$$

aaa

Не дописано

Лекция 3

16 сентября

Определение. Для группы $\mathcal{G} = (G, +)$ с подгруппой $\mathcal{H} = (H, +)$ назовём **смежным классом** для $x \in G$ множество:

$$x + H = \{x + h \mid h \in H\}$$

Пример. Для группы \mathbb{Z} группа чётных чисел $2\mathbb{Z}$ является подгруппой. Для 1 смежный класс — все нечётные числа, для 2 — все чётные.

Не дописано

Иногда демодулятор может сообщить, что некоторый полученный символ ненадежен. Это называется **стиранием**. Также стиранием считаются потери пакетов в компьютерных сетях (UDP).

Стирания исправлять проще, чем ошибки, т.к. мы знаем, где они произошли.

Определение. **Весовой спектр кода** это $A_i = |\{c \in C \mid \text{wt}(c) = i\}|$

Для двоичного симметричного канала с переходной вероятностью p вероятность необнаружения ошибки это:

$$P\{S = 0\} = \sum_{i=d}^n A_i p^i (1-p)^{n-i} \leq \sum_{i=d}^n C_n^i p^i (1-p)^{n-i}$$

Не дописано

Для аддитивного гауссовского канала с двоичной модуляцией вероятность ошибки мягкого декодирования по максимуму правдоподобия линейного блочного кода:

$$P \leq \sum_{i=d}^n A_i Q\left(\sqrt{2i \frac{E_s}{N_0}}\right) = \sum_{i=d}^n A_i Q\left(\sqrt{2iR \frac{E_b}{N_0}}\right) = \frac{1}{2} \sum_{i=d}^n A_i \text{erfc}\left(\sqrt{iR \frac{E_s}{N_0}}\right)$$

Определение. Полное декодирование по минимальному расстоянию — нахождение по y ближайшего кодового слова $c = \operatorname{argmin}_{c \in C} d(c, y)$

Определение. Информационная совокупность (ИС) — множество из k позиций в кодовом слове, значения которых однозначно определяют значения символов на остальных позициях кодового слова.

Определение. Если $\gamma = \{j_1 \dots j_k\}$ — информационная совокупность, то остальные позиции кодового слова называются **проверочной совокупностью**.

Если γ — информационная совокупность, то матрица из столбцов порождающей матрицы с номерами $j_1 \dots j_k$, обратима. Почему? **Не дописано** Будем обозначать такую матрицу $M(\gamma)$.

Примечание. Эта матрица не единственна.

Пусть $G(\gamma) = M(\gamma)G$ — порождающая матрица, содержащая единичную подматрицу на столбцах $j_1 \dots j_k$.

Информационные совокупности позволяют более эффективно проверять коды, т.к. стандартный метод требует таблицу размера 2^k .

Теорема 3. Алгоритм декодирования по информационной совокупности обеспечивает полное декодирование по минимальному расстоянию.

Доказательство. Необходимо доказать, что для всякого исправимого вектора ошибки r существует информационная совокупность, свободная от ошибок.

Пусть c — единственное решение задачи декодирования по минимальному расстоянию. Тогда $e = r - c$ — вектор ошибки, $E = \operatorname{supp}(e)$ — множество позиций ненулевых элементов e и $|E| \leq n - k$.

Пусть $N = \{1 \dots n\}$. Предположим, что $N \setminus E$ не содержит информационных совокупностей, тогда существуют различные кодовые слова, отличающиеся от r в позициях E . Но тогда c не единственно, противоречие. \square

Сложность декодирования для $(n, k)_q$ кода экспоненциальная.

2.2 Покрывтия

Определение. **Покрывтием** $M(n, m, t)$ называется такой набор $F \subset 2^{N_n}$, где $N_n = \{1 \dots n\}$, $\forall f \in F \ |f| = m$ и любое t -элементное подмножество N_n содержится в одном из $f \in F$.

Для декодирования по информационной совокупности с исправлением не более t ошибок необходимо покрыть все исправимые конфигурации ошибок. Элементы такого покрытия — проверочные совокупности.

Пример. Не дописано

Построение минимального покрытия — NP-полная задача. Но есть итеративный приближенный алгоритм.

2.3 Дуальные коды

Не дописано

2.4 Граница Хэмминга

Теорема 4. Для любого q -ичного кода с минимальным расстоянием $d = 2t + 1$ число кодовых слов удовлетворяет

$$A_q(n, d) \leq \frac{q^n}{\sum_{i=0}^t C_n^i (q-1)^i}$$

Доказательство. Если код способен исправить t ошибок, то вокруг всех кодовых слов можно описать хэмминговы шары радиуса t , не пересекающиеся друг с другом. \square

2.4.1 Асимптотическая оценка

$$\begin{aligned} A_q(n, d) &\leq \frac{q^n}{\sum_{i=0}^t C_n^i (q-1)^i} \\ &\leq \frac{q^n}{C_n^{\frac{d-1}{2}} (q-1)^{\frac{d-1}{2}}} \\ &= \frac{q^n \left(\frac{d-1}{2}\right)! \left(n - \frac{d-1}{2}\right)!}{n! (q-1)^{\frac{d-1}{2}}} \\ &= ??? \end{aligned}$$

2.5 Граница Варшамова–Гилберта

Теорема 5. Существует q -ичный код длины n с минимальным расстоянием d , число слов которого удовлетворяет:

$$A_q(n, d) \geq \frac{q^n}{\sum_{i=0}^{d-1} C_n^i (q-1)^i}$$

Доказательство. Если код C имеет максимальную мощность, для любого вектора $x \notin C$ существует кодовое слово c такое, что $d(x, c) \leq d-1$. Не дописано \square

2.6 Граница Варшамова–Гилберта для линейных кодов

Теорема 6. Если

$$q^r > \sum_{i=0}^{d-2} C_{n-1}^i (q-1)^i$$

, то существует линейный код над $GF(q)$ длины n с минимальным расстоянием не менее d и не более чем $r = n - k$ проверочными символами.

Доказательство. Построим матрицу H размера $(n - k) \times n$ такую, что любые $d - 1$ её столбцов линейно независимы.

Первый столбец пусть будет произвольным ненулевым вектором. Если уже выбраны j столбцов, то $j + 1$ -й столбец не может быть никакой линейной комбинацией любых $d - 2$ выбранных столбцов. Таких столбцов $\sum_{i=0}^{d-2} C_j^i (q-1)^i$. Пока не запрещены все q^{n-k} столбцов, то можно выбрать ещё хотя бы один столбец. \square

Не дописано

2.7 Граница Грайсмера

Обозначение. $N(k, d)$ — минимальная длина двоичного линейного кода размерности k с минимальным расстоянием d .

Теорема 7. $N(k, d) \geq d + N(k - 1, \lceil d/2 \rceil)$

Доказательство. Пусть порождающая матрица (n, k, d) кода C наименьшей длины $n = N(k, d)$ имеет вид:

$$G = \begin{pmatrix} 0 & \dots & 0 & 1 & \dots & 1 \\ & & G' & & * & \end{pmatrix}$$

Не дописано

\square

Лекция 4

23 сентября

3 Универсальные методы декодирования линейных кодов

Мы изучали методы жесткого декодирования, но на практике чаще применяются методы мягкого декодирования, которые учитывают надежность символов.

Мы уже знаем, что декодирование кода по критерию максимального правдоподобия в канале с АБГШ эквивалентно декодированию по критерию минимального расстояния Евклида.

Несложными преобразованиями получим:

???

Пусть тогда $\hat{c}_i = \begin{cases} 0, & y_i > 0 \\ 1, & y_i \leq 0 \end{cases}$ — жесткие решения.

Тогда:

$$\begin{aligned} \operatorname{argmax}_{c \in C} \sum_{i=0}^{n-1} (-1)^{c_i} y_i &= \operatorname{argmax}_{c \in C} \sum_{i=0}^{n-1} (-1)^{c_i} y_i \\ &= ??? \end{aligned}$$

Не дописано

3.1 Метод порядковых статистик

Рассмотрим передачу кодовых слов $c_0 \dots c_{n-1}$ двоичного (n, k) кода с помощью символов 2-АМ, например $y_i = (-1)^{c_i} + \eta_i$ — с АБГШ.

Пусть тогда $L_i = \log \frac{P\{c_i=0|y_i\}}{P\{c_i=1|y_i\}}$ — логарифмические отношения правдоподобия. Вероятность ошибки в i -том жестком решении убывает с увеличением $|L_i|$. Тогда выберем информационную совокупность J для кода, соответствующего наибольшим значениям $|L_i|$. Приведем порождающую матрицу кода к виду G_J с единичной подматрицей в столбцах J .

С большой вероятностью число неверных решений \hat{c}_j , где $j \in J$, мало. Переберем все конфигурации ошибок e веса не более t на J и построим кодовые слова $c_e = (\hat{c}_J + e)G_J$. Из всех полученных кодовых слов выберем наиболее правдоподобное.

Сложность алгоритма $\mathcal{O}(k^2n + \sum_{i=0}^t \ln C_k^i)$.

Есть два¹ способа ускорить этот алгоритм:

1. Обменять память на скорость каким-то образом (придумать дома).
2. Реализовать раннюю остановку.

Рассмотрим двоичный (n, k, d) код C с порождающей матрицей G , пусть $D(x, y)$ — функция расстояния Хемминга.

Существует много кодов, содержащих одинаковые префиксы некоторой длины a , поэтому можно не пересчитывать $\sum_{i=0}^a D(c_i, y_i)$ заново.

Не дописано

3.2 Декодирование по решеткам

Определение. Решетка (англ. ???) — граф, обладающий следующими свойствами:

1. Вершины графа разбиты на непересекающиеся подмножества, называемые **уровнями** или **ярусами**.
2. Нулевой и последний ярусы содержат по одной вершине, называемой **терминальной**.
3. Граф направленный и допускается движение только от уровня с меньшим номером к уровню с большим номером.
4. Ребрам графа сопоставлены метки, соответствующие символам кодовых слов, а также метрики, называемые весами.

На таком графе можно запустить алгоритм Дейкстры или алгоритм Витерби

¹ Не взаимозаменяющих

Определение. Профиль сложности решетки это $\xi_0 \dots \xi_n$, где $\xi_i = |V_i|$.

Определение. Решетка называется **минимальной**, если профиль сложности решетки минимален среди всех решеток с заданным количеством ярусов.

Рассмотрим все кодовые слова $c_m = c_{m,0} \dots c_{m,n-1}$ кода.

Для любого i определим префикс длины i , называемый **прошлым** и обозначим его c_m^p , а оставшийся суффикс длины $n - i$ — **будущим** и обозначим его c_m^f .

Очевидно, что в произвольной решетке пути, входящие в фиксированную вершину, имеют общее будущее, а пути, исходящие из фиксированной вершины, имеют общее прошлое.

Не дописано

Докажем, что построенная решетка минимальна.

Доказательство. Рассмотрим произвольную решетку T' этого кода.

В T' два слова $c_1 = (c_1^p, c_1^f)$ и $c_2 = (c_2^p, c_2^f)$ могут иметь общую вершину на ярусе i только если $F_i(c_1^p) = F_i(c_2^p)$. По построению два пути, проходящие через общую вершину в T' , проходят также через общую вершину в T .

Таким образом, число вершин на ярусе i в T' не меньше числа вершин на ярусе i в T . \square

Теорема 8. Любой код имеет минимальную решетку, и все минимальные решетки совпадают с точностью до нумерации вершин яруса.

Не дописано

Теорема 9. Решетка, получаемая по порождающей матрице в минимальной спановой форме, минимальна.

Доказательство. Докажем, что для любого $l \in \mathbb{N}$ пути, определяющие слова с одинаковыми c^f длины $n - l$, не проходят через различные узлы на ярусе с номером l .

Узел, через который проходит путь на ярусе l , определяется значениями информационных символов, которые соответствуют активным на этом ярусе строкам. Т.к. эти строки линейно независимы и заканчиваются на ярусах с номерами $> l$, следовательно, нетривиальные линейные комбинации этих строк отличаются хотя бы на одной позиции $> l$.

Предположим, что есть два слова с одинаковым будущим, проходящие через разные узлы на ярусе l . Их сумма образует слово, активное на ярусе l и равное 0 на позициях правее l . Но из соображений выше таких слов быть не может, а следовательно слова с одинаковым будущим проходят через одни и те же узлы, а следовательно решетка минимальна. \square

Не дописано

Теорема 10. Решетка, построенная по проверочной матрице, минимальна.

Доказательство. Докажем, что пути с одинаковыми c^f не проходят через разные узлы.

Для кодового слова $c = (c^p, c^f)$ частичные синдромы, вычисленные по c^p и c^f , совпадают. Следовательно, все совпадающие c^f исходят из одного и того же узла, определенного частичным синдромом c^p . \square

4 Декодирование с мягким выходом

Длинные коды можно строить путём комбинирования более коротких кодов. Как — мы узнаем позже. Декодеры таких кодов могут быть построены из декодеров кодовых компонент. Такие декодеры могут взаимодействовать путем обмена апостериорными вероятностями:

$$p\{c_i = a \mid y_0^{n-1}\} = \sum_{c \in C_i(a)} p\{c \mid y_0^{n-1}\}$$

???

4.1 Алгоритм Бала–Коке–Елинека–Равива

Нужно вычислить

$$L_i = \ln \frac{P\{c_i = 0 \mid y_0^{n-1}\}}{P\{c_i = 1 \mid y_0^{n-1}\}} = \ln \frac{\sum_{(s', s) \in S_0} \frac{p(s_i = s', s_{i+1}, y_0^{n-1})}{p(y_0^{n-1})}}{\sum_{(s', s) \in S_1} \frac{p(s_i = s', s_{i+1}, y_0^{n-1})}{p(y_0^{n-1})}}$$

, где S_0 и S_1 — множества пар состояний $s' \in V_i, s \in V_{i+1}$, переход между которыми помечен 0 и 1 соответственно, $p(y_0^{n-1})$ — совместная плотность распределения принятых сигналов, $p(s_i = s', s_{i+1} = s, y_0^{n-1})$ — совместная плотность распределения принятых сигналов и состояний кодера на ярусах i и $i + 1$.

Поведение кодера при обработке i -го информационного бита определяется только его состоянием s' на предыдущем шаге и канал не имеет памяти:

$$p(s_i = s', s_{i+1} = s, y_0^{n-1}) = p(s_i = s', y_0^{n-1})p(s_{i+1} = s, y_i \mid s_i = s', y_0^{y-1})p(y_{i+1}^{n-1} \mid s_{i+1} = s, s_i = s', y_0^i)$$

$$= \underbrace{p(s_i = s', y_0^{n-1})}_{\alpha_i(s')} \underbrace{p(s_{i+1} = s, y_i \mid s_i = s')}_{\gamma_{i+1}(s', s)} \underbrace{p(y_{i+1}^{n-1} \mid s_{i+1} = s)}_{\beta_{i+1}(s)}$$

Примечание. Предыдущая строка — определения α, β, γ

Лекция 5

30 сентября

По формуле Байеса:

$$\alpha_i(s) = \sum_{\tilde{s} \in V_{i-1}} \alpha_{i-1}(\tilde{s}) \gamma_i(\tilde{s}, s), s \in V_i$$
$$\beta_i(\tilde{s}) = \sum_{s \in V_{i+1}} \gamma_{i+1}(\tilde{s}, s) \beta_{i+1}(s)$$

Не дописано

Начальные значения для рекуррентной формулы:

$$\beta'_n(s) = \beta_n(s) = \begin{cases} 1, & s = 0 \\ 0, & s \neq 0 \end{cases}$$

$$\gamma_{i+1}(s', s) = ???$$

Не дописано

4.2 Выводы

Метод порядковых статистик позволяет выполнить мягкое декодирование произвольного линейного блочного кода. Увеличение сложности позволяет приблизиться к декодированию по максимум правдоподобия.

Всякий линейный блочный код может быть представлен в виде решетки.

Алгоритм Витерби реализует декодирование линейного кода по максимум правдоподобия.

Алгоритм Бала-Коке-Елинека-Равива реализует посимвольное декодирование линейного кода по максимуму апостериорной вероятности.

5 Сверточные коды

Задача кодера — сделать передаваемые символы статистически зависимыми. Сверточный код отображает автоматом блоки данных в кадры кодового слова.

Пример. Простейший автомат — регистр сдвига.

Не дописано

Лекция 6

7 октября

6 Методы модификации и комбинирования кодов

Мы не хотим строить сложные коды целиком, проще комбинировать коды поменьше.

6.1 Конструкция Плоткина

Пусть дано два кода C_1 и $C_2 : (n, k_i, d_i)$. Построим тогда код $C = \{(c_1, c_1 + c_2 \mid c_i \in C_i, i \in \{1, 2\})\}$. Тогда этот код $(2n, k_1 + k_2, \min(2d_1, d_2))$.

Порождающая матрица этого кода, если G_1 — порождающая матрица для C_1 и G_2 — для C_2 :

$$G = \begin{pmatrix} G_1 & G_1 \\ 0 & G_2 \end{pmatrix}$$

Доказательство. C содержит кодовые слова вида $(c_1, c_1) \mid c_1 \in C_1$ и вида $(0, c_2) \mid c_2 \in C_2$, следовательно, $d \leq 2d_1$ и $d \leq d_2$

Пусть $c_1, c'_1 \in C_1 \setminus \{0\}$ и $c_2, c'_2 \in C_2 \setminus \{0\}$ — ненулевые слова компонентных кодов.

$$d((c_1, c_1 + c_2), (c'_1, c'_1 + c'_2)) = d(c_1, c'_1) + d(c_1 + c_2, c'_1 + c'_2)$$

???

□

Будем декодировать $(y', y'') = (c_1, c_1 + c_2) + (e', e'')$ в метрике Хэмминга.

$$y''' = y'' - y' = c_1 + c_2 + e'' - c_1 - e' = c_2 + e''$$

Продекодируем получившийся y''' в коде C_2 . Если $\text{wt}((e', e'')) \leq \lfloor (d-1)/2 \rfloor$, то $\text{wt}(e''') \leq \text{wt}(e') + \text{wt}(e'') \leq \lfloor (d-1)/2 \rfloor \leq \lfloor (d_2-1)/2 \rfloor$ и декодирование успешно, т.е. мы получим c_2 .

Предположим, что c_2 получено правильно. Продекодируем в C_1 вектора $y' = c_1 + e'$ и $y'' - c_2 = c_1 + e$. Если $\text{wt}((e', e'')) \leq ???$, то декодирование y' или y'' даст правильный результат.

6.1.1 Коды Рида-Маллера

Это семейство кодов определяется рекурсивным применением конструкции Плоткина. Код Рида-Маллера порядка r длины 2^m обозначается как $RM(r, m)$.

Код $RM(0, m)$ — код $(2^m, 1, 2^m)$. Код $RM(m, m)$ — $(2^m, 2^m, 1)$.

Код $RM(r+1, m+1)$ получается применением конструкции Плоткина к $RM(r+1, m)$ и $RM(r, m)$.

Недавно было доказано, что коды Рида-Маллера оптимальны в том смысле, что он достигает границы Шеннона для двоичного канала с АБГШ и некоторых других каналов. Проблема этих кодов в том, что их сложно эффективно декодировать.

6.2 Модификации кодов

6.2.1 Укорочение

Укороченный код получается путем выбора кодовых слов исходного кода, содержащий нули на заданных позициях. Эти нули удаляются.

Если дан (n, k, d) код с $G = (I \mid A)$, то мы удаляем из порождающей матрицы m столбцов единичной подматрицы и соответствующие им m строк. Если появятся нулевые столбцы, то мы их тоже удаляем. Таким образом получится порождающая матрица $(\leq n-m, k-m, \geq d)$ кода.

6.2.2 Выкалывание

Удалим из всех кодовых слов заданные символы (обычно проверочных ???).

Если дан (n, k, d) код с $H = (A \mid -I)$, то удалим из H m столбцов единичной подматрицы и соответствующие им m строк. Если появятся линейно зависимые строки, то мы их тоже удаляем. Тогда мы получим $(n-m, \leq k, \geq d-m)$ код.

Пример: таким образом можно построить оптимальный код $(10, 3, 5)$.

6.2.3 Расширение

Наиболее распространенный способ — проверки на четность, так мы получаем из (n, k, d) кода $(n+1, k, d')$ код.

Если d нечетно, то $d' = d + 1$, иначе¹ $d' = d$.

6.3 Каскадные коды

Переमेжитель переставляет символы таким образом, чтобы последствия ошибочного декодирования одного кода могут быть легко ликвидированы декодером другого кода.

6.4 Прямое произведение кодов

Для двух заданных кодов (n_1, k_1, d_1) и (n_2, k_2, d_2) построим прямое произведение этих двух кодов — один из кодов будет декодировать по строкам, а другой — по столбцам.

Порождающая матрица такого кода будет иметь вид

$$G' \oplus G'' = \begin{pmatrix} G'_{11}G'' & G'_{12}G'' & \dots & G'_{1n_1}G'' \\ G'_{21}G'' & G'_{22}G'' & \dots & G'_{2n_1}G'' \\ \vdots & \vdots & \ddots & \vdots \\ G'_{k_11}G'' & G'_{k_12}G'' & \dots & G'_{k_1n_1}G'' \end{pmatrix}$$

Пропускная способность такого кода будет равна

$$R = \frac{k_1 k_2}{n_1 n_2} < \frac{k_1}{n_1}, \frac{k_2}{n_2}$$

Код способен исправить многие (но не все) ошибки веса большего, чем $d_1 d_2 / 2$.

Алгоритм декодирования можно параллелизировать.

6.4.1 Пример: код Рао-Редди (48, 31, 8)

РЖД использует этот код для передачи данных машинисту.

Рассмотрим прямое произведение расширенного (16, 11, 4) кода Хемминга C_1 и (3, 2, 2) кода C_2 с проверкой на четность. Получится код (48, 22, 8).

Дополним этот код кодовыми словами кода Рида-Маллера (16, 5, 8) C_3 , к которым дописали 32 нуля в конец для размерности.

Тогда кодовые слова будут иметь вид $(c_1 + c_3, c_2, c_1 + c_2) \mid c_1, c_2 \in C_1, c_3 \in C_3$.

Вес кодового слова ≥ 8 , т.к.:

- ???
- ???

¹ Мне так кажется

Теперь получен код $(48, 27, 8)$.

Пусть C_4 — код с порождающей матрицей ???.

Итого код $(48, 31, 8)$ Рао-Редди состоит из кодовых слов вида:

$$(c_1 + c_3 + c_4, c_2 + c_4, c_1 + c_2 + c_4) \mid c_1, c_2 \in C_1, c_3 \in C_3, c_4 \in C_4$$

Вывод: создание кодов больше похоже на искусство, непонятно что лучше работает.

6.5 Лестничные коды

???

6.6 Граница Зяблова

Выберем внутренний (n, k, d) код на границе Варшавова–Гилберта с $r = k/n \geq 1 - h(d/n) = 1 - \delta$.

???

6.7 Обобщенные каскадные коды (Блох, Зяблов, Зиновьев)

??? внешние (N, K_i, D_i) коды \mathcal{A}_i над $GF(q^{m_i}) \mid 1 \leq i \leq s$.

??? вложенные внутренние (n, k_i, d_i) коды $\mathcal{B}_i : \mathcal{B}_1 \supset \mathcal{B}_2 \supset \dots \supset \mathcal{B}_s$ над $GF(q)$, при этом $k_i - k_{i+1} = m_i$ и код \mathcal{B}_i порождается последними k_i строками $k_1 \times m$ матрицы B .

Кодирование: ???

???

Декодирование: ???

6.8 Турбо-коды

Будем кодировать данные сразу двумя сверточными кодами.

???

Минимальное расстояние растет с увеличением длины крайне медленно и начиная с некоторого места перестает расти. Поэтому использовать длинные турбо-коды неразумно.

???

6.9 Построение перемежителей

Давайте найдем, как перемежитель должен переставлять символы.

Будем переставлять близкие позиции во входной последовательности в максимально удаленные позиции в выходной последовательности:

$$0 < |i - j| < d \Rightarrow |\pi(i) - \pi(j)| \geq S$$

- Можно генерировать случайные перестановки и отбрасывать неэффективные. Но это долго.
- Табличный перемежитель: запишем все символы в квадратную матрицу по строкам, а прочитаем ее по столбцам.
- Перестановочный полином: если ???

Перемежитель в целом нужен когда мы делаем композицию декодеров, т.к. канал с декодером на нем уже не является каналом с АБГШ.

???

Лекция 7

14 октября

7 Полярные коды

7.1 Некоторые определения

7.1.1 Функция переходных вероятностей канала

Определение. Рассмотрим канал без памяти с входным алфавитом $\mathcal{X} = \{0, 1\}$ и выходным алфавитом \mathcal{Y} .

Если \mathcal{Y} дискретен, то **функция переходных вероятностей канала** $W(y | c)$ — вероятность наблюдения на выходе $y \in \mathcal{Y}$ при условии подачи на его вход $c \in \mathcal{X}$.

Если \mathcal{Y} непрерывен, то **функция переходных вероятностей канала** $W(y | c)$ — плотность распределения выходного символа при подаче c на его вход

Пример (Двоичный симметричный канал).

$$\mathcal{Y} = \mathcal{X} \quad W(y | c) = \begin{cases} p, & y \neq x \\ 1 - p, & y = x \end{cases}$$

Пример (Двоичный стирающий канал).

$$\mathcal{Y} = \{0, 1, \varepsilon\} \quad W(y | c) = \begin{cases} p, & y = \varepsilon \\ 1 - p, & y = x \in \{0, 1\} \end{cases}$$

Пример (Двоичный симметричный канал со стираниями).

$$\mathcal{Y} = \{0, 1, \varepsilon\} \quad W(y | c) = \begin{cases} 1 - p - s, & y = x \\ s, & y = \varepsilon \\ p, & y \neq x, y \neq \varepsilon \end{cases}$$

Пример (Аддитивный гауссовский канал).

$$\mathcal{Y} = \mathbb{R}, y = (-1)^c + \eta, \eta \sim \mathcal{N}(0, \sigma^2) \quad W(y | c) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y - (-1)^c)^2}{2\sigma^2}\right)$$

7.1.2 Параметр Бхаттачарьи

Рассмотрим канал без памяти с двоичным кодом, приемник по максимуму правдоподобия.

Если передаваемые символы равновероятны, то вероятность ошибки:

$$\begin{aligned} P_e &= P\{c = 0\}P\{\text{err} | c = 0\} + P\{c = 1\}P\{\text{err} | c = 1\} \\ &= \frac{1}{2} \sum_{y: W(y|0) < W(y|1)} W(y | 0) + \frac{1}{2} \sum_{y: W(y|1) < W(y|0)} W(y | 1) \\ &= \frac{1}{2} \sum_{y: \frac{W(y|1)}{W(y|0)} > 1} W(y | 0) + \frac{1}{2} \sum_{y: \frac{W(y|0)}{W(y|1)} > 1} W(y | 1) \\ &= \frac{1}{2} \sum_{y \in \mathcal{Y}} \sum_{c \in \{0,1\}} \left(W(y | c) \chi\left(\frac{W(y | 1 - c)}{W(y | c)}\right) \right) \end{aligned}$$

, где χ — индикаторная функция:

$$\chi(z) = \begin{cases} 1, & z \geq 1 \\ 0, & z < 1 \end{cases}$$

Заметим, что $\chi(z) \leq \sqrt{z}$ для всех $z \in \mathbb{N}$. Таким образом, можно оценить вероятность ошибки:

$$\begin{aligned} P_e &\leq \frac{1}{2} \sum_{y \in \mathcal{Y}} \sum_{c \in \{0,1\}} W(y | c) \sqrt{\frac{W(y | 1 - c)}{W(y | c)}} \\ &= \frac{1}{2} \sum_{y \in \mathcal{Y}} \sum_{c \in \{0,1\}} \sqrt{W(y | 1 - c) W(y | c)} \\ &= \sum_{y \in \mathcal{Y}} \sqrt{W(y | 1) W(y | 0)} \\ &=: Z(W) \end{aligned}$$

Определение. $Z(W)$ — параметр Бхаттачарьи ФПВК W .

Пример (Двоичный стирающий канал).

$$Z(BEC(p)) = \sqrt{W(0 | 0)W(0 | 1)} + \sqrt{W(1 | 0)W(1 | 1)} + \sqrt{W(\varepsilon | 0)W(\varepsilon | 1)} = 0 + 0 + p = p$$

Пример (Аддитивный гауссовский канал).

$$\begin{aligned} Z(\mathcal{G}(\sigma)) &= \int_{-\infty}^{\infty} \sqrt{W(y|0)W(y|1)} dy \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \exp\left(-\frac{(y-1)^2 + (y+1)^2}{4\sigma^2}\right) dy \\ &= \exp\left(-\frac{1}{2\sigma^2}\right) \end{aligned}$$

Пропускную способность канала можно вычислить по формуле:

$$I(W) = \max_{\{p(x)\}} \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} W(y|x) P\{x\} \log \frac{W(y|x)}{W(y)}$$

Для многих каналов оптимальным распределением символов на входе $P\{x\}$ является равномерное, что мы и будем рассматривать дальше.

7.2 Поляризация канала

Рассмотрим следующее линейное преобразование:

$$\begin{pmatrix} c_0 & c_1 \end{pmatrix} = \begin{pmatrix} u_0 & u_1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

Пропустим полученные символы через двоичный стирающий канал с вероятностью стирания p , получим y_0, y_1 .

Если оба символа не стерты, то u_0 можно восстановить из y_0, y_1 как $y_0 \oplus y_1$. Вероятность того, что этого не произойдет — $1 - (1-p)^2 = 2p - p^2 \geq p$

u_1 восстанавливается либо как y_1 , либо как $y_0 \oplus u_0$ и тогда нам нужна подсказка свыше о u_0 . Вероятность того, что восстановить не получится — вероятность того, что оба символа стерты, т.е. $p^2 \leq p$.

Таким образом, мы с помощью **поляризации** получили два виртуальных канала — один чуть получше, другой чуть похуже, чем исходный канал.

Пусть $n = 2^m$, $A_m = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}^{\oplus m}$, где $\oplus m$ — m -кратное произведение Кронекера матрицы с собой. Введем $W_m(y_0^{n-1} | c_0^{n-1}) = \prod_{i=0}^{n-1} W(y_i | c_i)$.

Вход канала — u_i , выход — реальный выход канала y_0^{n-1} и подсказка свыше о предыдущих символах u_0^{i-1} .

Определим ФПВК для синтетических каналов:

$$W_m^{(i)}(y_0^{n-1}, u_0^{i-1} \mid u_i) = \frac{W_m^{(i)}(y_0^{n-1}, u_0^i)}{P\{u_i\}}$$

Не дописано