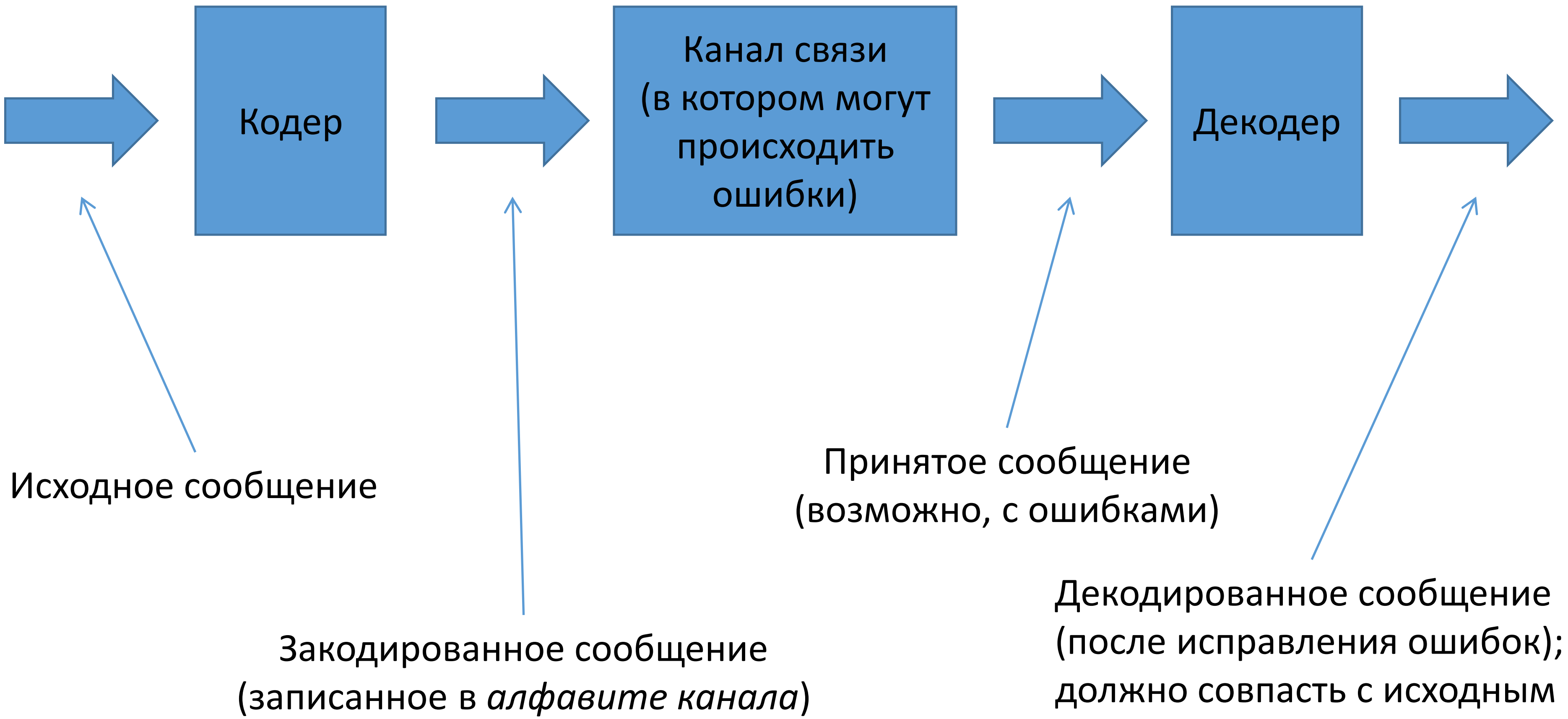
Коды, исправляющие ошибки

Основная модель канала связи:



Типы ошибок

- Ошибки замещения: муха → мука
  - Симметричные
  - Несимметричные
- Ошибки стирания: муха → му?а
- Ошибки выпадения: муха → уха
- Ошибки вставки: мука → мурка
- Комбинации перечисленных типов

Граница Плоткина

Теорема. (M. Plotkin)

Пусть $n < 2d$. Тогда для любого (n, M, d)-кода

$$M \leq \left\lfloor \frac{2d}{2d - n} \right\rfloor$$

Граница Плоткина

Доказательство:

Рассмотрим матрицу, в которой по строкам выписаны все кодовые слова:

$$\begin{pmatrix} \boldsymbol{a}_1 \\ \vdots \\ \boldsymbol{a}_M \end{pmatrix}$$

Элементы этой матрицы будем обозначать a_{ij}.
Оценим снизу и сверху следующую сумму:

$$T \coloneqq \sum_{\substack{1 \le k \le n \\ 1 \le i < j \le M}} \mathbb{1}_{a_{ik} \ne a_{jk}}$$

Граница Плоткина

Имеем

$$T = \sum_{1 \le i < j \le M} \sum_{1 \le k \le n} \mathbb{1}_{a_{ik} \ne a_{jk}} = \sum_{1 \le i < j \le M} d(\boldsymbol{a}_i, \boldsymbol{a}_j)$$

Отсюда

$$T \ge \frac{M \cdot (M-1)}{2} \cdot d$$

Граница Плоткина

С другой стороны

$$T = \sum_{1 \le k \le n} \sum_{1 \le i < j \le M} \mathbb{1}_{a_{ik} \ne a_{jk}}$$

Зафиксируем произвольное k.
Пусть среди кодовых слов ровно x_s слов имеют k-ю координату, равную s. Тогда

$$\sum_{1 \le i < j \le M} \mathbb{1}_{a_{ik} \ne a_{jk}} = x_0 \cdot x_1 \le \frac{M^2}{4}$$

Граница Плоткина

При любом k мы получаем

$$\sum_{1 \le i < j \le M} \mathbb{1}_{a_{ik} \ne a_{jk}} \le \frac{M^2}{4}$$

Значит

$$T = \sum_{1 \le k \le n} \sum_{1 \le i < j \le M} \mathbb{1}_{a_{ik} \ne a_{jk}} \le \frac{nM^2}{4}$$

Граница Плоткина

Сопоставим верхнюю и нижнюю оценки для T:

$$\frac{M \cdot (M-1)}{2} \cdot d \le T \le \frac{nM^2}{4}$$

Отсюда

$$M \le \left\lfloor \frac{2d}{2d-n} \right\rfloor$$

Матрицы Адамара (J. Hadamard)

Матрица Адамара — это квадратная матрица из $\{-1, 1\}^{n \times n}$, в которой любые две строки ортогональны.

Примеры:

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{pmatrix}$$

Теорема Адамара

Матрицы Адамара берут начало от следующей теоремы:

Теорема. (J. Hadamard)
Если $A = \left(a_{ij}\right)_{1 \le i,j \le n} \in \mathbb{R}^{n \times n}$ и $\left|a_{ij}\right| \le 1$ для любых i, j, то тогда

$$|\det A| \le n^{n/2}$$

Доказательство:

- $|\det A|$ — это объём параллелепипеда, построенного на векторах-строках матрицы A
- Объём максимален, когда длины сторон максимальны (максимум равен $\sqrt{n}$ при $\left|a_{ij}\right| = 1$) и углы между сторонами прямые (т.е. векторы ортогональны).

Матрицы Адамара

Если H — матрица Адамара, то

- Матрица, полученная из H перестановками строк/столбцов, тоже является матрицей Адамара.
- Матрица, полученная из H умножением строк/столбцов на -1, тоже является матрицей Адамара.

Матрицы Адамара, получаемые друг из друга такими преобразованиями, *эквивалентны*.

Матрицы Адамара

Любую матрицу Адамара умножением строк/столбцов на -1 можно привести к виду

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & & & \\ \vdots & & & \\ 1 & & & \end{pmatrix}$$

Такая матрица Адамара называется *нормализованной*.

Порядок матриц Адамара

Утверждение.

Если $H \in \{-1, 1\}^{n\times n}$ — матрица Адамара, и $n > 2$, то $4|n$.

Доказательство:

От матрицы H перейдём к эквивалентной матрице, в которой первые три строки такие:

$$\begin{matrix} \underbrace{\begin{matrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \end{matrix}}_{i} & \underbrace{\begin{matrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ -1 & -1 & \cdots & -1 \end{matrix}}_{j} & \underbrace{\begin{matrix} 1 & 1 & \cdots & 1 \\ -1 & -1 & \cdots & -1 \\ 1 & 1 & \cdots & 1 \end{matrix}}_{k} & \underbrace{\begin{matrix} 1 & 1 & \cdots & 1 \\ -1 & -1 & \cdots & -1 \\ -1 & -1 & \cdots & -1 \end{matrix}}_{l} \end{matrix}$$

Отсюда

$$\begin{cases} i + j + k + l = n \\ i + j - k - l = 0 \\ i - j - k + l = 0 \\ i - j + k - l = 0 \end{cases}$$

Решение этой системы: $i = j = k = l = n/4$.

Порядок матриц Адамара

Гипотеза Адамара (не доказана).
Матрицы Адамара порядка n существуют(?) для *всех* натуральных n, кратных четырём.

Наименьший порядок, для которого пока не доказано существование матрицы Адамара, равен 668.

Матрицы Адамара

Теорема. (R. E. A. C. Paley ’1933)
Если p простое и $4|(p+1)$, то существует матрица Адамара порядка $(p+1)$.

(Конструкция Пэли на основе квадратичных вычетов.)

Квадратичные вычеты

Элемент $a \in \mathbb{Z}_p \setminus \{0\}$ называется *квадратичным вычетом*, если $a = x^2$ для некоторого $x \in \mathbb{Z}_p$.

Остальные элементы из $\mathbb{Z}_p \setminus \{0\}$ называются *квадратичными невычетами*.

Например, в $\mathbb{Z}_7$ элементы $1,2,4$ — к.в.,
а $3,5,6$ — к.н.

Квадратичные вычеты

Утверждение.

- Если $p > 2$ простое, то ровно половина элементов из $\mathbb{Z}_p \setminus \{0\}$ являются к.в., а половина — к.н.

<u>Везде далее будем предполагать, что $p > 2$.</u>

Символ Лежандра

Символ Лежандра $\chi(a)$ определяется так:

$$\chi(a) = \begin{cases} 0, \text{если } a = 0 \\ 1, \text{если } a \text{ к. в.} \\ -1, \text{если } a \text{ к. н.} \end{cases}$$

Утверждение.

Для любых $a, b \in \mathbb{F}_q$ имеет место равенство

$$\chi(a) \cdot \chi(b) = \chi(ab)$$

Квадратичные вычеты

Утверждение.

Для любого $c \in \mathbb{Z}_p \setminus \{0\}$ имеет место равенство

$$\sum_{b \in \mathbb{Z}_p} \chi(b) \cdot \chi(b+c) = -1$$

Доказательство:

Т.к. *ровно* половина элементов $\mathbb{Z}_p \setminus \{0\}$ квадратичными вычетами, то $\sum_{a \in \mathbb{Z}_p} \chi(a) = 0.$

Также заметим, что

$$\sum_{b \in \mathbb{Z}_p} \chi(b) \cdot \chi(b+c) = \sum_{b \in \mathbb{Z}_p \setminus \{0\}} \chi(b) \cdot \chi(b+c)$$

Квадратичные вычеты

С учётом замеченного, получаем

$$\sum_{b\in\mathbb{Z}_p\setminus\{0\}} \chi(b)\cdot\chi(b+c) = \sum_{b\in\mathbb{Z}_p\setminus\{0\}} \chi(b)\cdot\chi\big(b\cdot b^{-1}(b+c)\big) =$$

$$= \sum_{b\in\mathbb{Z}_p\setminus\{0\}} \big(\chi(b)\big)^2\cdot\chi\big(b^{-1}(b+c)\big) = \sum_{b\in\mathbb{Z}_p\setminus\{0\}} \chi\big(b^{-1}(b+c)\big) =$$

$$= \sum_{b\in\mathbb{Z}_p\setminus\{0\}} \chi(1+b^{-1}c) = \sum_{a\in\mathbb{Z}_p\setminus\{1\}} \chi(a) = \sum_{a\in\mathbb{Z}_p} \chi(a) - \chi(1) = -1$$

Матрица Якобшталя (E. Jacobsthal)

Рассмотрим матрицу $\left(t_{a,b}\right)_{a,b\in\mathbb{Z}_p} \in \{-1,0,1\}^{p\times p}$, в которой $t_{a,b} \coloneqq \chi(a-b)$.

Скалярное произведение любых двух различных строк $\left(t_{a',b}\right)_{b\in\mathbb{Z}_p}$ и $\left(t_{a'',b}\right)_{b\in\mathbb{Z}_p}$ равно

$$\sum_{b\in\mathbb{Z}_p} \chi(a'-b)\cdot\chi(a''-b) = \sum_{b\in\mathbb{Z}_p} \chi(b)\cdot\chi\big(b+(a''-a')\big) = -1$$

«Подправленная» матрица Якобшталя

Рассмотрим матрицу $\left(t'_{a,b}\right)_{a,b\in\mathbb{Z}_p} \in \{-1,1\}^{p\times p}$, в которой $t'_{a,b} = \chi(a-b)$, если $a \neq b$ и $t'_{a,b} = -1$ иначе.

Скалярное произведение различных строк $\left(t'_{a',b}\right)_{b\in\mathbb{Z}_p}$ и $\left(t'_{a'',b}\right)_{b\in\mathbb{Z}_p}$ равно

$$\left(\sum_{b\in\mathbb{Z}_p} \chi(a'-b)\cdot\chi(a''-b)\right) - \chi(a'-a'') - \chi(a''-a') =$$
$$= -1 - \chi(a'-a'') - \chi(a''-a')$$

Если (-1) является квадратичным *не*вычетом в $\mathbb{Z}_p$, то

$$\chi(a''-a') = \chi(-1)\cdot\chi(a'-a'') = -\chi(a'-a''),$$

и скалярное произведение получается равным -1.

«Подправленная» и «дополненная» матрица Якобшталя

$T' \coloneqq \left(t'_{a,b}\right)_{a,b \in \mathbb{Z}_p} \in \{-1, 1\}^{p \times p}$,

где $t'_{a,b} = \chi(a - b)$, если $a \neq b$, и $t'_{a,b} = -1$ иначе.

Если (-1) является квадратичным невычетом в $\mathbb{Z}_p$, то скалярное произведение любых двух строк матрицы T' равно -1. Тогда матрица

$$\begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & & & \\ \vdots & & T' & \\ 1 & & & \end{pmatrix}$$

является нормализованной матрицей Адамара.

Матрицы Адамара

Утверждение. (Без доказательства)
При $4|(q+1)$ элемент (-1) является квадратичным невычетом в $\mathbb{Z}_p$.

Следствие.
Если p простое и $4|(p+1)$, то существует матрица Адамара порядка $(p+1)$.

Коды Адамара

Введены R. C. Bose, S. S. Shrikhande ’1959.

Идея:
В матрице Адамара любые две строки $\boldsymbol{a}, \boldsymbol{b}$ ортогональны. Т.к. $\boldsymbol{a}, \boldsymbol{b} \in \{-1, 1\}^n$, это значит, что ровно половина координат у них совпадает, а половина противоположны.

Заменяем координаты $-1 \to 0$ и получаем из строк матрицы двоичный код с большим кодовым расстоянием.

Коды Адамара

Пусть $A \in \{0,1\}^{n\times n}$ — матрица, полученная из нормализованной матрицы Адамара заменой элементов -1 на 0.

- Множество строк матрицы A с отброшенной первой координатой образует двоичный $\left(n-1, n, \frac{n}{2}\right)$-код
- Множество строк матрицы A и их дополнений образует $\left(n, 2n, \frac{n}{2}\right)$-код

Оптимальность кодов Адамара

Граница Плоткина.

Если $N < 2d$, то для любого (N, M, d)-кода

$$M \leq \frac{2d}{2d - N}$$

Коды Адамара с параметрами $\left(n-1, n, \frac{n}{2}\right)$ достигают границы Плоткина, имея максимально число слов при заданных длине и кодовом расстоянии.