Лекция 1

Упрощенная модель цифровой системы связи

Общая схема передачи данных

Источник данных → Информационные символы → Кодер источника → Кодовые символы + шум → Канал → Кодовые символы + шум → Декодер источника → Декодированные данные → Получатель

Базовые элементы модели

1. Дискретные последовательности

- Работа в поле Галуа **GF(2) = {0, 1}**
- Пример последовательности: `01000100`

2. Двоичный симметричный канал (ДСК)

- На вход: `0` или `1`
- **С вероятностью** `p` происходит ошибка:
  - `0` передаётся как `1`
  - `1` передаётся как `0`
- **С вероятностью** `1-p` ошибки нет:
  - `0` → `0`

- `1` → `1`

  `p` — переходная вероятность (вероятность ошибки одного символа).

Методы борьбы с ошибками

Идея: минимизация ошибок путём дублирования данных

Пример:

- **Источник:** `0011`
- **Кодер:** `000 000 111 111` (трёхкратное повторение каждого бита)

Процесс кодирования

Информационные символы (K)	Кодовые символы (n)
0	000
1	111

$R = K/n$ — **скорость кода**

Теоретическая основа (Клод Шеннон, 1948)

Пропускная способность двоичного симметричного канала

Пропускная способность:

$$C = 1 - h(p)$$

Энтропия двоичного ансамбля:

$$h(x) = -xlog2x - (1 - x)log2(1 - x)$$

Теорема Шеннона:

- При скорости передачи $R < C$ может быть обеспечена сколь угодно малая вероятность ошибки декодирования за счёт увеличения длины кодов
- При $R > C$ — надёжная передача невозможна

Расстояние Хэмминга

Определение:
Если x — кодовое слово, то $w(x)$ — **вес Хэмминга**, определяется как число ненулевых элементов в x.
В двоичном случае — число единиц.

Расстояние Хэмминга между двумя кодовыми словами x и y:

$d(x, y)$ = количество элементов, которые отличаются друг от друга

Свойства:

- $d(x, y) = w(x + y)$ (побитовое сложение по модулю 2)
- $d(x, 0) = w(x)$

Пример:
x = 001101, w(x) = 3
y = 101001, w(y) = 3
d(x, y) = 2

Пример:

1 00 -> 00000

2 01 -> 10110

3 10 -> 01011

4 11 -> 11101

считаем $d(x, y)$

	1	2	3	4
1	0	3	3	4
2	3	0	4	3
3	3	4	0	3
4	4	3	3	0

приходит новое слово затем считаем расстояние до всех слов, к которому ближе то слово и прилетело

Минимальное расстояние кода

Определение:

$$dmin = mind(x, y), x \neq y$$

Связь с исправлением ошибок:

Если код исправляет ошибки кратности t, то:

$$t \leq \lfloor (dmin - 1)/2 \rfloor$$

Линейный код

Определение: Код, в котором сумма двух любых кодовых слов тоже является кодовым словом.

$$\forall x, y \in C : (x + y) \in C$$

Свойство расстояния:

$$d(x, y) = w(x + y) = w(z) = w(z + 0) = d(z, 0)$$

где $z = x + y \in C$

Минимальное расстояние:

$$dmin = mind(x, y) = minw(z), x \neq y, z \in C, z \neq 0$$

Линейный q-ичный (n, k) код C над $GF(q)$

Определение: Любое k-мерное подпространство пространства F_q^n всевозможных векторов длины n.

Пример: $q = 3, \quad k = 2, \quad n = 5$

- $GF(3) = 0, 1, 2$
- $F_3^2 = 00, 01, 02, 10, 11, 12, 20, 21, 22$
- F_3^5 — аналогично, векторы длины 5

Порождающая матрица

Определение: Порождающей матрицей (n, k)-кода называется матрица размера $k \times n$, строки которой — базисные векторы.

Обозначение: G — порождающая матрица

Формирование кодового слова:

$$c = m \cdot G$$

где $m = (m_1, m_2, \ldots, m_k)$ — информационное слово (2^k вариантов)

Пример базиса:

$$e1 = (10110), e2 = (01011)$$

Линейные комбинации:

- $0 \cdot e_1 + 0 \cdot e_2 = c_1 = 00000$
- $0 \cdot e_1 + 1 \cdot e_2 = c_3 = 01011$
- $1 \cdot e_1 + 0 \cdot e_2 = c_2 = 10110$
- $1 \cdot e_1 + 1 \cdot e_2 = c_4 = 11101$

Проверочная матрица

Определение: Для вектора $h = (h_1, h_2, \ldots, h_n)$ выполняется:

$$(c, h) = c_1 \cdot h_1 + c_2 \cdot h_2 + \ldots + c_n \cdot h_n = 0$$

для всех кодовых слов $c \in C$.

Проверочная матрица H:

- Размера $(n - k) \times n$
- Удовлетворяет условию: $G \cdot H^T = 0$
- Основное свойство: $c \cdot H^T = 0$ для всех кодовых слов

Пример: Для кода с $n = 5$, $k = 2$ матрица H имеет размер (3×5) и содержит $n - k = 3$ проверки.