# Помехоустойчивое кодирование

## **Линейные коды**

# Некоторые предположения

- Блоковый код- код, в котором все слова имеют одинаковую длину.
- Кодовое слово – слово из некоторого кода С.

Исходные предположения относительно канала

1. Сохранение длины. Слово на выходе канала имеет такую же длину, как кодовое слово на входе канала.
2. Независимость ошибок. Вероятность ошибки любого символа сообщения одна и та же.

# Исходная стратегия декодирования

- При декодировании мы используем принцип максимального правдоподобия, или стратегию ближайшего соседа, согласно которым получатель должен декодировать полученное слово *w'* как кодовое слово *w, ближайшее к w'*.

# Расстояние Хэмминга

- Интуитивное понятие “близости" двух слов формализуется с помощью расстояния Хэмминга $d(x, y)$ слов $x$, $y$.
- Для двух слов $x$, $y$

  $d(x, y)$ = число символов, в которых они различаются.
- Примеры: $h(10101, 01100) = 3$,
  $h(fourth, eighth) = 4$

# Свойства расстояния Хэмминга (1)

- (1) $d(x, y) = 0$ Ű $x = y$
- (2) $d(x, y) = d(y, x)$
- (3) $d(x, z)$ Ł $d(x, y) + d(y, z)$ (неравенство треугольника)

- Важнейшей характеристикой кода $C$ является его минимальное расстояние
  - $d(C) = \min \{d(x, y) \mid x,y \in C, x$ ą $y\}$,

- $d(C)$ дает наименьшее число ошибок, необходимое для перевода одного кодового слова в другое.

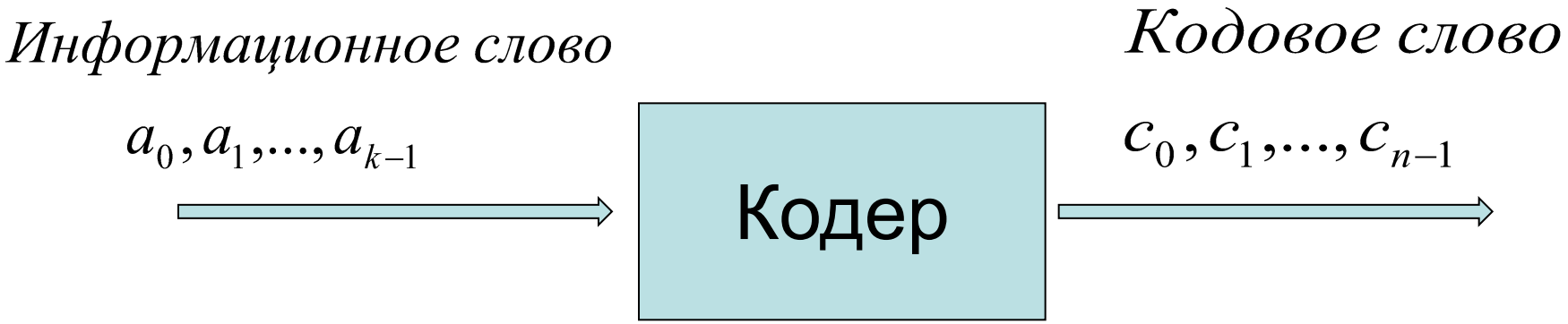
# Свойства расстояния Хэмминга (2)

- Теорема (Основная теорема  исправления ошибок)
- (1) Код $C$ может обнаруживать до s ошибок, если $d(C) \geq s + 1$.
- (2) Код $C$ может исправлять до $t$ ошибок, если $d(C) \geq 2t + 1$.

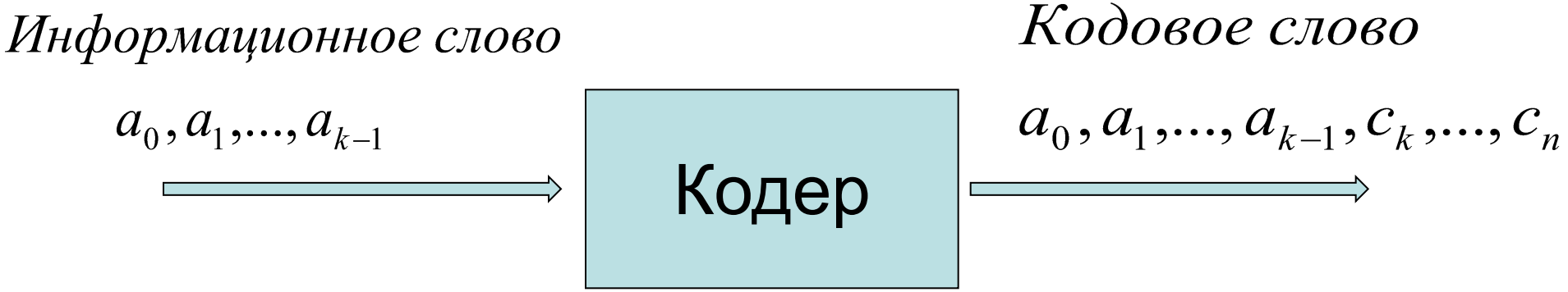
Доказательство (1) Очевидно.

- (2) Предположим $d(C) \geq 2t + 1$.
  Пусть передается кодовое слово $x$ и получено слово $y$  так что $d(x, y) \leq t$.
  Если $x' \neq x$ является кодовым словом, тогда $d(x' , y) \geq t + 1$ поскольку в противном случае $d(x', y) < t + 1$ и следовательно $d(x, x') \leq d(x, y) + d(y, x') < 2t + 1$ что противоречит предположению $d(C) \geq 2t + 1$.

# Кодирование – введение избыточности –алгебраический подход



# Систематическое кодирование



# Кодирование – введение избыточности (систематическое кодирование)

$$\left.\begin{aligned} c_0 &= a_0, \\ c_1 &= a_1, \\ &\dots \\ c_{k-1} &= a_{k-1}, \end{aligned}\right\} - \textit{информационные биты}$$

$$\left.\begin{aligned} c_k &= f_k(c_0,\dots,c_{k-1}), \\ &\dots \\ c_{n-1} &= f_{n-1}(c_0,\dots,c_{k-1}) \end{aligned}\right\} - \textit{проверочные биты}$$

# Линейное систематические кодирование – линейные функции

$$\left.\begin{array}{l} c_0 = a_0, \\ c_1 = a_1, \\ \ldots \\ c_{k-1} = a_{k-1}, \end{array}\right\} - \textit{информационные биты}$$

$$\left.\begin{array}{l} c_k = f_k(c_0, \ldots, c_{k-1}), \\ \ldots \\ c_{n-1} = f_{n-1}(c_0, \ldots, c_{k-1}) \end{array}\right\} - \textit{проверочные биты},$$

$$f_k, \ldots, f_{n-1} - \textit{линейные булевы функции}$$

# Пример линейного систематического кодирования - добавление проверки на четность(1)

**Пример.**

| Информационное слово | Кодовое слово |
|---|---|
| 000 | 0000 |
| 001 | 0011 |
| 010 | 0101 |
| 011 | 0110 |
| 100 | 1001 |
| 101 | 1010 |
| 110 | 1100 |
| 111 | 1111 |

$$c_0 = a_0,$$

$$c_1 = a_1,$$

$$c_2 = a_2,$$

$$c_3 = c_0 \oplus c_1 \oplus c_2.$$

# Линейный код (некоторые параметры) - *(n,k,d)*-код

- *n* – длина кодовых слов (длина кода)
- *k* – число информационных разрядов
- *d* – минимальное кодовое расстояние
- $R = \frac{k}{n}$ - скорость передачи
- Комментарий: Хороший (*n*,*k*,*d*)-код имеет маленькое *n* и большие *k* и *d*.

# Примеры

- $C_1$ = {00, 01, 10, 11} есть (2,2,1)-код.
- $C_2$ = {000, 011, 101, 110} есть (3,2,2)-код.
- $C_3$ = {00000, 01101, 10110, 11011} есть (5,2,3)-код.

# ISBN-код – недвоичный код

- Каждая книга имеет International Standard Book Number, которое представляет собой 10-разрядное кодовое слово создаваемое издателем и имеющее следующую структуру:

- $l \quad p \quad m \quad w \quad = \quad x_1 \ldots x_{10}$
- язык издатель номер взвешенная контрольная сумма
- 0 07 709503 0
- так что

$$\sum_{i=1}^{10} i x_i \equiv 0 \pmod{11}$$

Издатель добавляет $X$ в 10-ю позицию, если $x_{10} = 10$.

# ISBN-код – недвоичный код

- Обнаружение одиночной ошибки

- Пусть $X = x_1 \ldots x_{10}$ - правильный код и пусть
  - $Y = x_1 \ldots x_{J-1}\, y_J\, x_{J+1} \ldots x_{10}$, причем $y_J = x_J + a$, $a \neq 0$
- В таком случае:

$$\sum_{i=1}^{10} iy_i = \sum_{i=1}^{10} ix_i + ja \neq 0 \pmod{11}$$

# ISBN-код – недвоичный код

- Обнаружение ошибки перестановки

- Пусть $x_J$ и $x_k$ поменялись местами.

$$\sum_{i=1}^{10} iy_i = \sum_{i=1}^{10} ix_i + (k-j)x_j + (j-k)x_k$$

$$= (k-j)(x_j - x_k) \neq 0 \pmod{11}$$

при $k \neq j$ и $x_j \neq x_k$.

# Пример линейного систематического кодирования - добавление проверки на четность(2)

**Пример.**

| Информационное слово | Кодовое слово |
|---|---|
| 000 | 0000 |
| 001 | 0011 |
| 010 | 0101 |
| 011 | 0110 |
| 100 | 1001 |
| 101 | 1010 |
| 110 | 1100 |
| 111 | 1111 |

$$c_0 = a_0,$$

$$c_1 = a_1,$$

$$c_2 = a_2,$$

$$c_3 = c_0 \oplus c_1 \oplus c_2.$$

$$\begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix}$$

# Порождающая матрица

Пусть $\gamma$ - кодовое слово длины *n*

$\alpha$ - информационное слово длины *k*

$$\gamma = G \cdot \alpha$$

G – *nxk порождающая матрица* кода

# Систематический код

- **Первые разрядов кодового слова совпадают с информационными битами**

$$G = \begin{pmatrix} I_k \\ G_1 \end{pmatrix}$$

# Порождающая матрица

- *Пример.* $\gamma = G \cdot \alpha$

$$\begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \\ c_6 \end{bmatrix} = \left(\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \hline 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{array}\right) \cdot \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix}$$

- *Длина слов n=7, число иформационных разрядов =4, число проверочных разрядов n-k=3*

# Проверки

- *Пример.* Получаем *проверки*

$$c_4 = c_0 \oplus c_2 \oplus c_3,$$
$$c_5 = c_0 \oplus c_1 \oplus c_2,$$
$$c_6 = c_1 \oplus c_2 \oplus c_3,$$

$$\begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \\ c_6 \end{bmatrix} = \left(\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \hline 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{array}\right) \cdot \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix}$$

# Проверочная матрица

- *Пример.*

$$\begin{pmatrix} 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \end{pmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \\ c_6 \end{bmatrix} = 0$$

$$c_0 \oplus c_2 \oplus c_3 \oplus c_4 = 0$$
$$c_0 \oplus c_1 \oplus c_2 \oplus c_5 = 0,$$
$$c_1 \oplus c_2 \oplus c_3 \oplus c_6 = 0,$$

- *H – (n-k)xn проверочная матрица:*

$$H\gamma = 0$$

# Связь порождающей и проверочной матрицы систематического кода

- *Пример.*
-

$$H = \left(\begin{array}{cccc|ccc} 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \end{array}\right) = (P\ I_3)$$

$$G = \left(\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \hline 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{array}\right) = \begin{pmatrix} I_4 \\ P \end{pmatrix},$$

$$G^T H = 0$$

# Связь порождающей и проверочной матрицы систематического кода

- $$\underset{n\times k}{G}=\begin{pmatrix} I_k \\ P_{(n-k)\times k} \end{pmatrix}, \qquad \underset{(n-k)\times n}{H}=(P_{(n-k)\times k}\ \ I_{n-k})$$

$$G^T H = 0$$

# Сводка результатов по линейным кодам

- Линейный код задается порождающей ( $G$ ) или проверочной ( $H$ ) матрицами.
- Код (множество кодовых слов) – линейное подпространство, порожденное столбцами $G$
- С другой стороны – линейный код – дуальное подпространство столбцов матрицы $H^T$ - дуальный код