ИНСТИТУТ
МАТЕМАТИКИ
МЕХАНИКИ
КОМПЬЮТЕРНЫХ
НАУК
имени И.И. Воровича —

Архитектура компьютера и операционные системы

Лекция 6. Помехоустойчивое кодирование

Андреева Евгения Михайловна
доцент кафедры информатики и вычислительного эксперимента

План лекции

- Помехоустойчивое кодирование (ECC)
  - бит четности
  - код троекратного повторения
  - (7, 4) - код Хэмминга
- Домашнее задание

Подсистема ввода-вывода

ИНСТИТУТ
МАТЕМАТИКИ
МЕХАНИКИ
КОМПЬЮТЕРНЫХ
НАУК
имени И.И. Воровича

Откуда берутся ошибки?

- Всплеск напряжения (при записи-чтении, память, диск)
- Физические повреждения (жесткие и оптические диски, шины)
- Ошибки приема/передачи данных (сети)
- …

Что делать с ошибками?

- обнаружение ошибок и автоматический запрос повторной передачи (на канальном и транспортном уровнях);
- обнаружение ошибок и отбрасывание повреждённых блоков (в системах потокового мультимедиа);
- обнаружение и исправление ошибок (применяется на физическом уровне).

Код коррекции/контроля ошибок

ECC (error-correcting/controlling code, код коррекции/контроля ошибок):

- сообщения длины K и кодовые слова длины N, N>K, (N, K)-код
- метрика Хэмминга (1950)
- простейший код контроля ошибок
  - код проверки четности
- простейший код исправления ошибки
  - код троекратного повторения

Кодовое слово

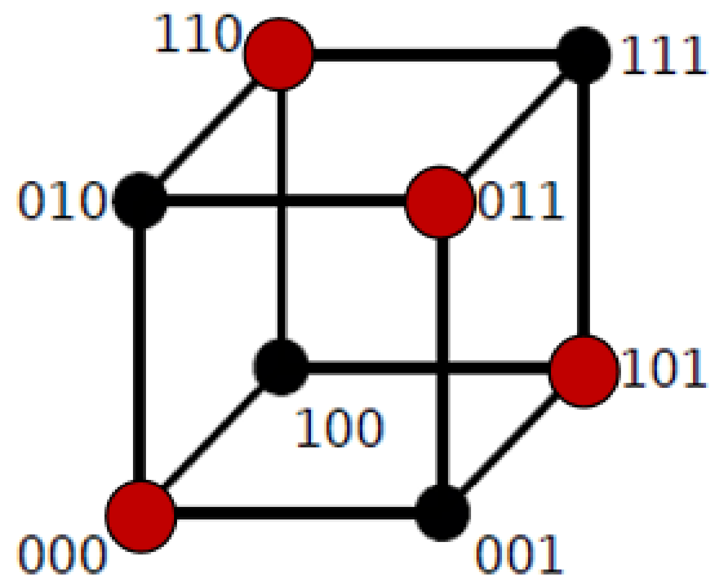
- Последовательность из N бит, содержащую K бит данных и R контрольных разрядов называют **кодовым словом**.
- Предположим, что слово состоит из K бит данных, к которым мы дополнительно прибавляем R бит (контрольных разрядов). Общая длина кодового слова составит N бит (N = K + R).
- Пример. Бит четности.
  - исходное слово – 01011101 (K=8)
  - контрольный разряд – 1 (R=1)
  - кодовое слово – 010111011 (N=9)

Кодовое пространство

- **Кодовое пространство** – совокупность всех кодовых слов (векторов) которые используются в данной системе кодирования. Максимальное количество кодовых векторов для двоичной системы в N-разрядном кодовом пространстве - 2^N.
- **Расстояние по Хэммингу** (метрика Хэмминга) - число позиций, в которых два кодовых вектора отличаются друг от друга (для 0110 и 0011 – расстояние по Хеммингу = 2).
- **Минимальное кодовое расстояние - d** -наименьшее кодовое расстояние, взятое по всем допустимым кодовым комбинациям в данной системе кодирования.

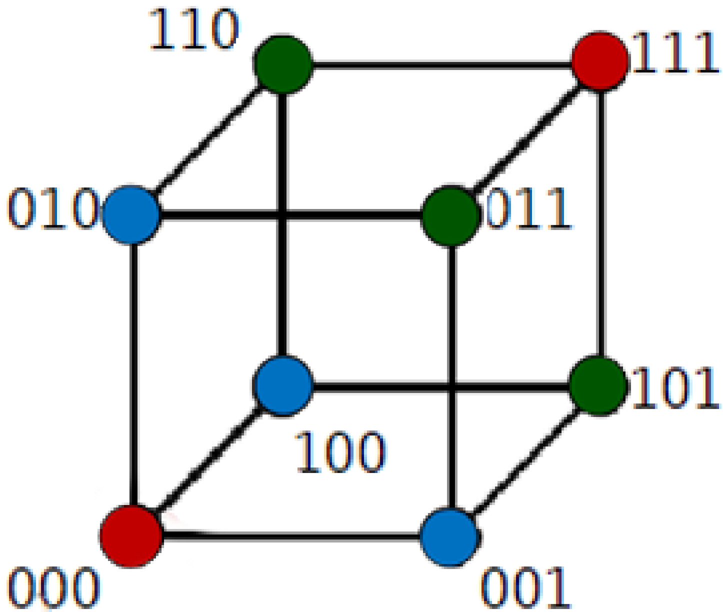
Пример. Бит четности

- N=3, K=2, R=1
- Допустимые кодовые слова
  - 00 0
  - 01 1
  - 10 1
  - 11 0
- d=2
- Обнаруживает, но не исправляет ошибки
- 2^K - количество разрешенных к передаче кодов (разрешенные кодовые комбинации)
- R/N – избыточность кода, здесь 1/3



Код троекратного повторения

- N=3, K=1, R=2
- Допустимые кодовые слова
  - 0 00
  - 1 11
- d=3
- Исправляет ошибку
- Избыточный, 2/3.



Обнаружение и исправление ошибок

- Пусть g — количество ошибок
- Для обнаружения ошибок:

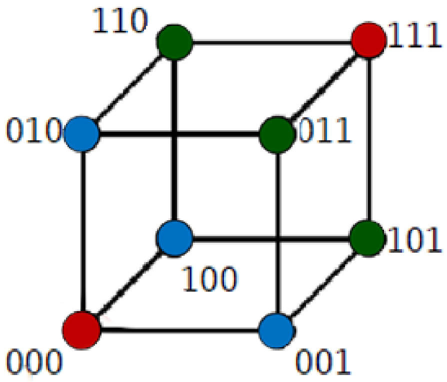
  $d \geq g+1$
- Для исправления ошибок:

  $d \geq 2*g+1$
- Пример:
  - допустимые кодовые слова – 0000000000, 0000011111, 1111100000, 1111111111
  - d=5
  - для слова 0000000111, при $g \leq 2$ можем исправить ошибки, при $g>2$ можем только обнаружить

Число контрольных разрядов

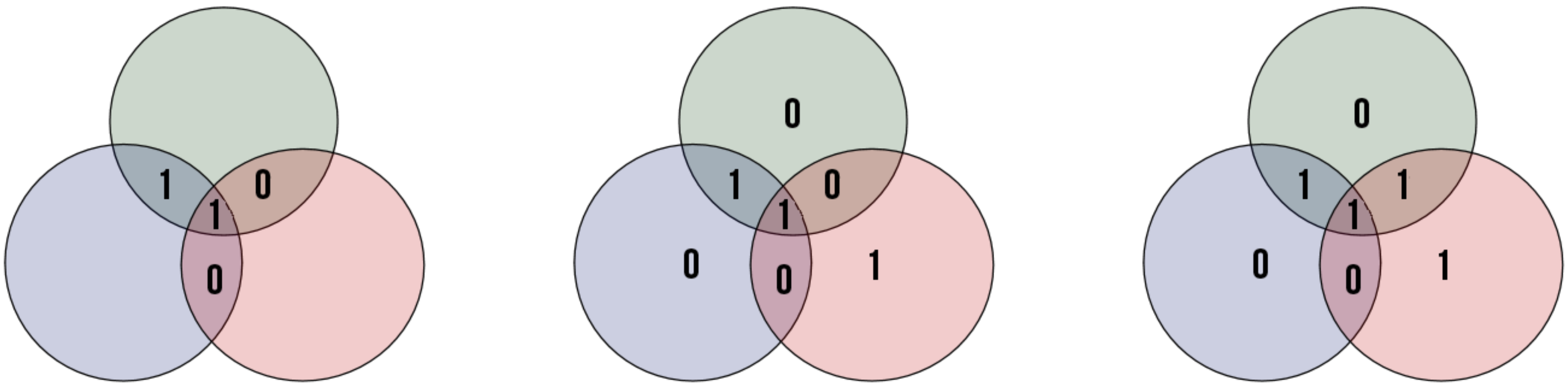
- Общая длина кодового слова составит N бит (N = K + R). K бит данных. R бит контрольных разрядов.
- Нижний предел числа контрольных разрядов, необходимых для исправления одиночных ошибок выражается формулой $K+R+1 \leq 2^R$

Размер исходного слова	Количество контрольных разрядов	Общий размер слова	Процент увеличения слова
8	4	12	50
16	5	21	31
32	6	38	19
64	7	71	11
128	8	136	6
256	9	265	4
512	10	522	2



$(N+1)\, 2^K \leq 2^N$

(7, 4) - код Хэмминга



- Принцип тот же - проверка на четность числа единичных битов: к последовательности добавляется такой элемент, чтобы число единичных символов в получившейся последовательности было четным.
- (7, 4) –код N=7, K=4, R=3

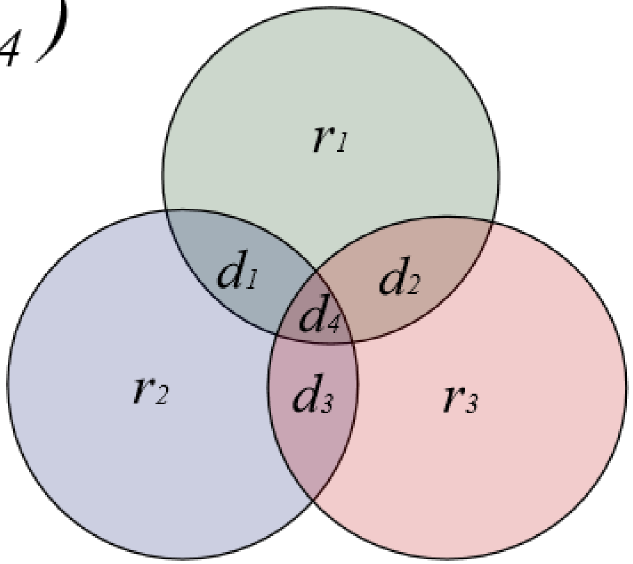
ИНСТИТУТ
МАТЕМАТИКИ
МЕХАНИКИ
КОМПЬЮТЕРНЫХ
НАУК
имени И.И. Воровича

Основные характеристики

- Если расстояние Хэмминга для двух слов равно d, значит, достаточно d одноразрядных ошибок, чтобы превратить одно слово в другое.
- Кол-во контролируемых ошибок d-1.
- Кол-во исправляемых ошибок (d-1)/2, округленное с недостатком

(7, 4) - код Хэмминга

(d_1, d_2, d_3, d_4)



$r_1 = d_1 + d_2 + d_4$
$r_2 = d_1 + d_3 + d_4$
$r_3 = d_2 + d_3 + d_4$

- Кодовое слово с=$(r_1, r_2, d_1, r_3, d_2, d_3, d_4)$.
- В позиции, соответствующие степеням 2 надо вписать контрольные разряды, позиции слева-направо.
- Семейство кодов Хэмминга – $(2^R\text{-}1, 2^R\text{-}R\text{-}1)$.

(7, 4) - код Хэмминга

- Пусть исходное слово (d_1, d_2, d_3, d_4). Контрольные разряды вычисляются по формулам:

$$r_1 = d_1 + d_2 + d_4$$
$$r_2 = d_1 + d_3 + d_4$$
$$r_3 = d_2 + d_3 + d_4$$

Кодовое слово $c=($	$r_1,$	$r_2,$	$d_1,$	$r_3,$	$d_2,$	$d_3,$	$d_4)$
номера позиций	1	2	3	4	5	6	7

- Тогда кодовое слово будет иметь вид $c=(d_1+d_2+d_4,\ d_1+d_3+d_4,\ d_1,\ d_2+d_3+d_4,\ d_2,\ d_3,\ d_4)$.
- Минимальное кодовое расстояние d=3

Проверочная матрица H

	r_1	r_2	d_1	r_3	d_2	d_3	d_4
r_1	1	0	1	0	1	0	1
r_2	0	1	1	0	0	1	1
r_3	0	0	0	1	1	1	1

- Пусть v – принятое по каналу слово, которое отличается от кодового слова с (7, 4) – кода Хэмминга не более, чем в одной позиции, тогда вектор s (синдром): s=Hv дает двоичную запись номера этой позиции. Если s=0, то v=c.

$$(d_1, d_2, d_3, d_4)$$

$$r_1 = d_1 + d_2 + d_4$$
$$r_2 = d_1 + d_3 + d_4$$
$$r_3 = d_2 + d_3 + d_4$$

ИНСТИТУТ МАТЕМАТИКИ МЕХАНИКИ КОМПЬЮТЕРНЫХ НАУК имени И.И. Воровича

Пример

	r_1	r_2	d_1	r_3	d_2	d_3	d_4
r_1	1	0	1	0	1	0	1
r_2	0	1	1	0	0	1	1
r_3	0	0	0	1	1	1	1

- v=(0111100)

$$s = Hv = \begin{matrix}1010101\\0110011\\0001111\end{matrix} \times 0111100^{T} = 000^{T}$$

- Пусть v_1=(0101100), тогда $s_1 = 110$
- Пусть v_2=(0110100), тогда $s_2 = 001$

ИНСТИТУТ
МАТЕМАТИКИ
МЕХАНИКИ
КОМПЬЮТЕРНЫХ
НАУК
имени И.И. Воровича

Домашнее задание

- Для закрепления лекций - читать [Таненбаум Э] стр. 98-168, 239-259
- Подготовка к тестированию по материалам лекций 5 и 6
- Для лабораторных занятий
  Приложение В из [Таненбаум Э] – стр. 784-789
- Подготовка к лабораторной 5.