

Границы 16

Формула Шеннона для **пропускной способности гауссовского канала:**

$$C_t = W \log\left(1 + \frac{P_s}{P_n}\right).$$

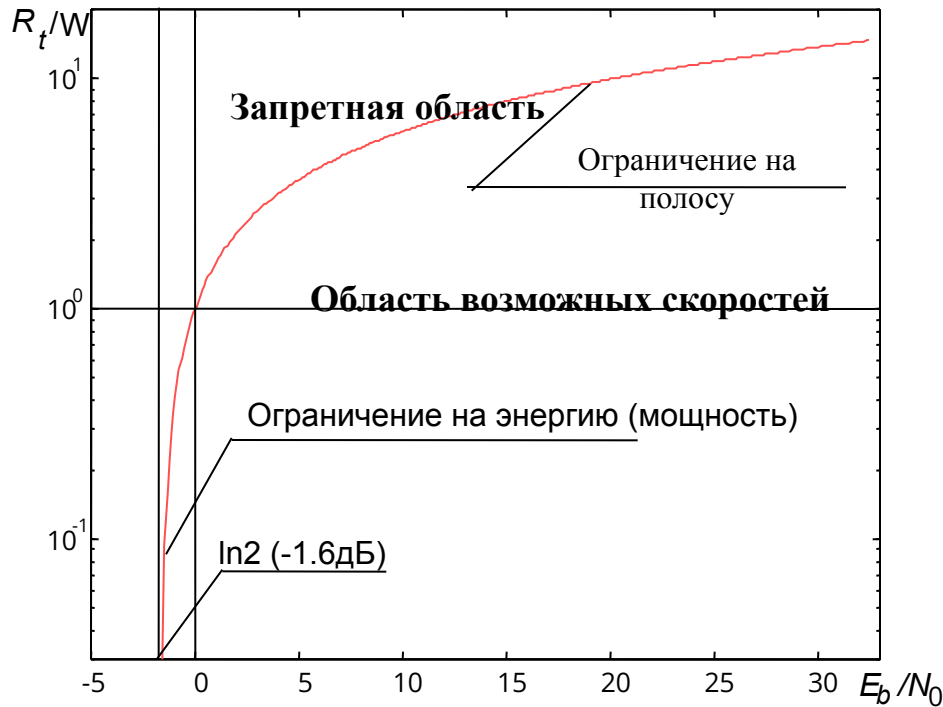
пропускной способностью гауссовского канала с ограниченной полосой W

В реальности спектральная плотность мощности шума N_0 может зачастую считаться постоянной в произвольной полосе W , так что $P_n = N_0 W$ и, когда полоса расширяется, пропускная способность растет, стремясь к пределу

$$C_{t\infty} = \frac{P_s}{N_0 \ln 2},$$

пропускная способностью гауссовского канала с неограниченной полосой

Граница Шеннона



$$\frac{E_b}{N_0} > \frac{2^{\frac{R_t}{W}} - 1}{\frac{R_t}{W}},$$

Зависимость **спектральной эффективности** R_t/W (скорости на 1 Гц) от **отношения сигнал-шум на бит**, где E_b – энергия сигнала на бит (но не кодовый символ!) полезной информации.

Если при проектировании системы высшим приоритетом является энергосбережение (величина E_b/N_0 не может быть значительной), единственным средством повышения надежности передачи служит расширение полосы ($W \gg R_t$).

Напротив, когда главная цель – высокая спектральная эффективность ($R_t \gg W$), проектировщик вынужден полагаться только на достаточную излучаемую энергию $E_b/N_0 \gg 1$.

Важнейшие границы теории кодирования. Граница Хэмминга

- **Теорема** Любой двоичный код, исправляющий вплоть до t ошибок, удовлетворяет неравенству

$$M \sum_{i=0}^t C_n^i \leq 2^n \iff \sum_{i=0}^t C_n^i \leq 2^{n(1-R)}.$$

- **Г.Х.** – утверждает, что если существует q -ичный линейный код длиной n со скоростью передачи информации R и минимальным расстоянием Хемминга d_{\min} или более, то

$$\sum_{i=0}^t C_n^i (q-1)^i \leq q^{n(1-R)}$$

Совершенные коды

- Коды, лежащие на границе Хэмминга (удовлетворяющие ей с равенством), называются **совершенными**.
- Совершенные коды исправляют любые ошибки кратности t и менее, но не исправляют и не обнаруживают никаких ошибок большей кратности
- Для совершенного кода, рассматриваемого как смежный класс, в качестве лидеров остальных смежных классов удаётся взять все векторы весом t и только их

Пример

- Пусть $n = 12, d = 5, t = 2, q = 2$
- Граница Хемминга дает следующий результат:

$$\sum_{i=0}^2 C_{12}^i \leq 2^{12-k} \Rightarrow 2^k \leq \frac{2^{12}}{\sum_{i=0}^2 C_{12}^i} \Rightarrow k = 5$$

- Таким образом, возможно построение кода с параметрами $(n, k) = (12, 5)$.

Совершенные коды

Геометрически такие коды реализуют так называемую «плотную упаковку», при которой все 2^n двоичных векторов распределены по M сферам радиуса t , не имеющих пересечений и промежутков.

Они названы совершенными, так как обеспечивают максимальную скорость R , допускаемую фиксированными $d = 2t + 1$ и n .

Среди двоичных существуют лишь три класса совершенных кодов – тривиальный код с повторением нечетной длины, коды Хэмминга, исправляющие однократные ошибки, и код Голея длины 23 с расстоянием 7 (исправляющий ошибки вплоть до трехкратных).

Код Голея

- Сфера – принадлежит ровно $1 + n + C_n^2 + \dots + C_n^t = \sum_{l=0}^t C_n^l$
- n -мерных векторов
- Заметим, что $\longrightarrow (C_{23}^0 + C_{23}^1 + C_{23}^2 + C_{23}^3)2^{12} = 2^{23}$
- Гипотеза: 23-мерное двоичное пространство можно плотно упаковать сферами радиуса 3.
- \swarrow Существует совершенный код (23, 12), исправляющий все тройные ошибки

Код Голея

- В основе конструкции кода лежит разложение
- $x^{23} - 1 = (1 + x) g_1(x) g_2(x)$
- $g_1(x) = 1 + x^2 + x^4 + x^5 + x^6 + x^{10} + x^{11}$
- $g_2(x) = 1 + x + x^5 + x^6 + x^7 + x^9 + x^{11}$

Порождающая матрица кода Голея (23, 12)

$$G := \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \end{bmatrix}$$

Граница Варшамова–Гилберта

Если $k = \log M$ целое число, эта граница модифицируется в более точную **границу Варшамова–Гилберта**, устанавливающую существование кода, удовлетворяющего неравенству:

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

$$\sum_{i=0}^{d-2} C_{n-1}^i < 2^{n(1-R)}.$$

Асимптотические версии нижних и верхних границ, выражают условия существования кодов в терминах скорости R и относительного кодового расстояния d/n :

Асимптотические версии

- Если $n \gg 1$, код не существует при нарушении любого из неравенств

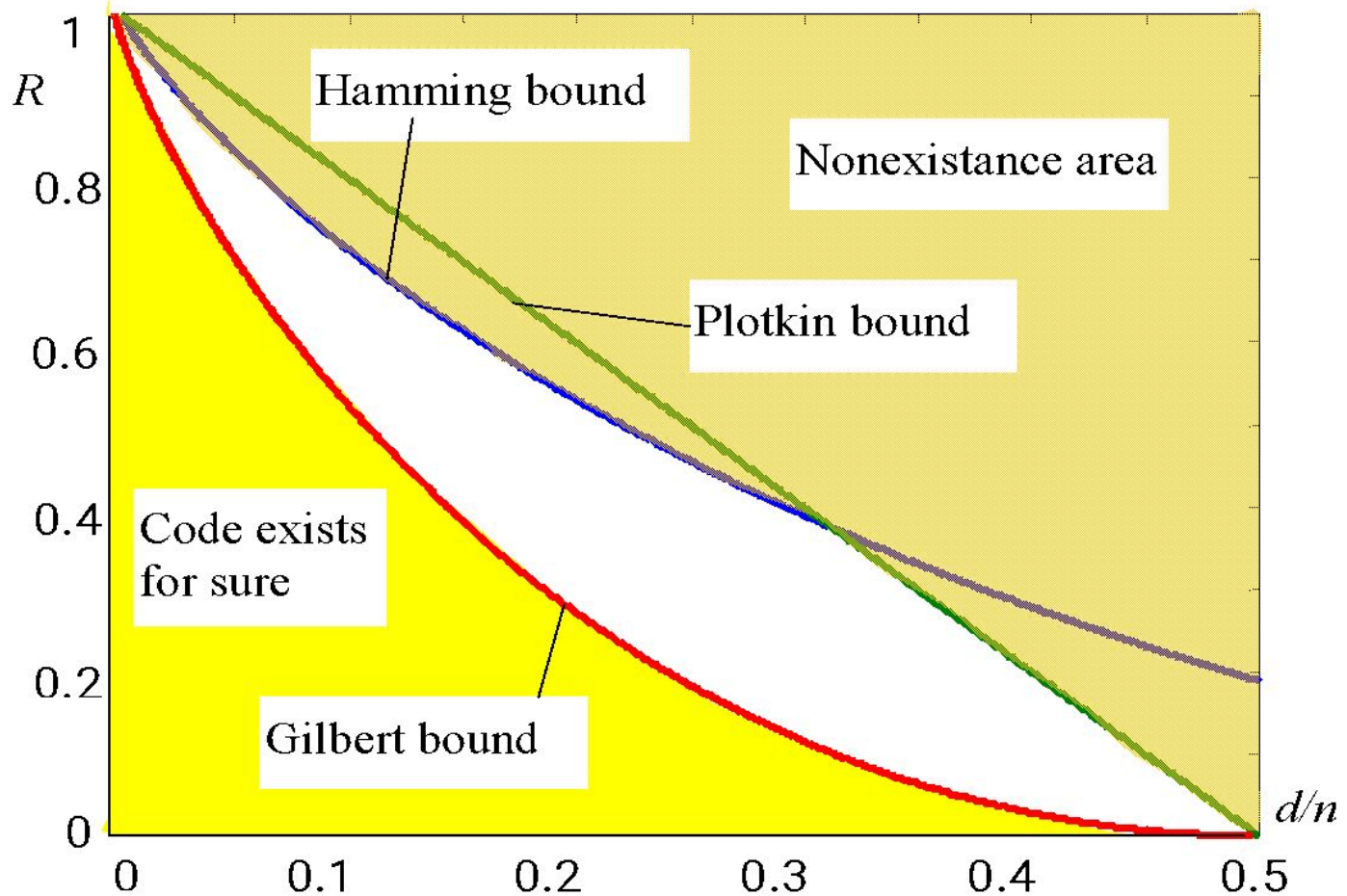
$$R < 1 - h\left(\frac{d}{2n}\right), \quad R < 1 - 2\frac{d}{n}$$

- **(асимптотические границы Хэмминга и Плоткина).**
- Но существует наверняка при условии **(асимптотическая граница Гильберта):**

$$R < 1 - h\left(\frac{d}{n}\right)$$

- где $h(\cdot)$ – энтропия двоичного ансамбля.

Границы



Комментарии

Коды с параметрами M , n , d , попадающими в область выше любой из границ Хэмминга или Плоткина, существовать не могут, тогда как в области ниже границы Гилберта существование кодов гарантировано.

Область между двумя упомянутыми является зоной неопределенности, для которой однозначный ответ о существовании кода нельзя получить с помощью рассмотренных границ (использование упоминавшихся более точных границ позволяет, разумеется, в той или иной мере сузить зону неопределенности) .

Пример

- Пусть исследуется соотношение $d = 5, (n, k) = (63, 51)$
- Находим границу Хемминга:

$$\sum_{i=0}^{t=2} C_{63}^i \leq 2^{n-k} \quad \rightarrow \quad 2017 \leq 2^{n-k}$$

$r = n - k = 11$

- Граница Варшамова – Гильберта равна

$$\sum_{i=0}^{d-2=3} C_{62}^i \leq 2^{n-k} \quad \rightarrow \quad 39774 > 2^{n-k}$$

$r = n - k = 16$

Т. о. $r = n - k = 12$ $11 < 12 < 16$

данный код близок к границе Хемминга

Hamming Bound

- **> HB := proc(n,d) local b,i,t,sum:**
- **t := floor((d-1)/2):**
- **sum := 1:**
- **for i from 1 to t do**
- **sum := sum + binomial(n,i):**
- **od:**
- **b := simplify(floor(n-log[2](sum))):**
- **printf("k is at most %d",b):**
- **end**

Gilbert-Varshamov Bound, version 1

- **> GV1 := proc(n,d) local b,i,sum:**
- **sum := 1:**
- **for i from 1 to d-1 do**
- **sum := sum + binomial(n,i):**
- **od:**
- **b := simplify(floor(n-log[2](sum))):**
- **printf("There is a code with dimension at least %d",b):**
- **end:**

Gilbert-Varshamov Bound, version 2

- **> GV2 := proc(n,d) local i,b,sum:**
- **sum := 1:**
- **for i from 1 to d-2 do**
- **sum := sum + binomial(n-1,i):**
- **od:**
- **b := simplify(floor(n-log[2](sum)))-1:**
- **printf("There is a code of dimension at least %d",b):**
- **end:**

Singleton Bound

- $SB := \text{proc}(n,d):$
- $> \text{printf}("k \text{ is at most } \%d", n-d+1):$
- $> \text{end}:$
- Пример
- $> \mathbf{SB(7,3)};$
- k is at most 5

Пример

- $> SB(7,3)$;
- k is at most 5
- $HB(7,3); k$ is at most 4
- $> GV1(7,3)$;
- There is a code with dimension at least 2
- $> GV2(7,3)$;
- There is a code of dimension at least 3

Источники сообщений, количество информации, энтропия

- **Идея определения количества информации**
- С теоретической точки зрения любая универсально применимая мера количества информации в сообщении, должна опираться только на степень предсказуемости последнего.
- Чем менее предсказуемо (вероятно) сообщение (событие), тем большее количество информации генерируется в результате его осуществления.

Математическая модель источника информации.

Дискретные источники

Дискретным называется источник, множество X возможных сообщений которого конечно или счетно $X = \{x_1, x_2, \dots\}$.

Подобный источник полностью описывается набором вероятностей сообщений: $p(x_i), i=1, 2, \dots$.

Условие нормировки:

$$\sum_{i=1}^M p(x_i) = 1 \quad \text{или} \quad \sum_{x \in X} p(x) = 1$$

Непрерывные источники

Рассматривая непрерывные источники, мы ограничимся только теми, несчетный ансамбль которых может быть ассоциирован с непрерывной случайной переменной, т. е. описан в терминах плотности вероятности $W(x)$.

Условие нормировки:

$$\int_{-\infty}^{\infty} W(x) dx = 1$$

Количество информации в сообщении

Аксиомы количества информации (требования к универсальной информационной мере):

1. Количество информации в сообщении x зависит только от его вероятности:

$$I(x) = f(p(x)), \forall x \in X.$$

2. Неотрицательность количества информации:

$$I(x) \geq 0, \forall x \in X,$$

причем $I(x)=0$ только для достоверного события ($p(x)=1$).

3. Аддитивность количества информации для независимых сообщений:

$$I(x, y) = I(x) + I(y).$$

Хартли

- Единственной функцией, удовлетворяющей этим трем аксиомам оказывается логарифм вероятности сообщения

$$I(x) = -\log p(x) = \log \frac{1}{p(x)}$$

Единица измерения количества информации зависит от выбора основания логарифма.

Традиционное основание два дает единицу измерения количества информации, называемую битом (*binary digit*).

Энтропия дискретного источника

Энтропия дискретного источника есть среднее количество информации в его сообщениях:

$$H(X) = \overline{I(X)} = - \sum_{x \in X} p(x) \log p(x) = \sum_{x \in X} p(x) \log \frac{1}{p(x)}.$$

Свойства энтропии:

1. Энтропия неотрицательна:

$$H(X) \geq 0,$$

где равенство нулю имеет место только для полностью детерминированного (неслучайного) источника.

Свойства энтропии:

2. Энтропия ограничена сверху соотношением

$$H(X) \leq \log M,$$

3. Энтропия ансамбля пар сообщений, генерируемых двумя независимыми источниками аддитивна

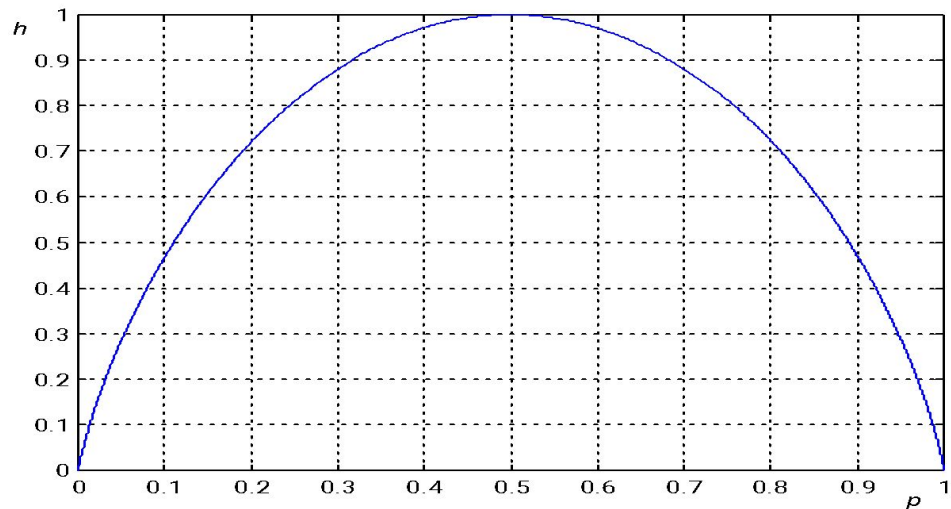
$$H(XY) = H(X) + H(Y).$$

Энтропия двоичного источника

Пусть p – вероятность одного из двух сообщений. Тогда

$$h(p) = -p \log p - (1 - p) \log(1 - p)$$

Энтропия двоичного источника в зависимости от p



Вопросы

