ПРОБЛЕМЫ ПЕРЕДАЧИ ИНФОРМАЦИИ

Том 56 **2020** **Вып. 2**

УДК 621.391.1 : 519.2

© 2020 г. **М.В. Бурнашев**

НОВЫЕ ГРАНИЦЫ В ЗАДАЧЕ ПРОВЕРКИ ГИПОТЕЗ С ИНФОРМАЦИОННЫМИ ОГРАНИЧЕНИЯМИ[1]

Рассматривается задача проверки гипотез, в которой мы не можем наблюдать часть данных. Наш помощник наблюдает пропущенные данные и может передать нам некоторую ограниченную информацию о них. Какая ограниченная информация позволит нам сделать наилучшие статистические выводы? В частности, какая минимальная информация достаточна для получения тех же результатов, как если бы мы непосредственно наблюдали все данные? Получены оценки для величины этой минимальной информации и некоторые подобные результаты.

Ключевые слова: проверка гипотез, информационные ограничения, вероятности ошибки.

DOI: 10.31857/S0555292320020023

§ 1. Введение и основные результаты

1. Постановка задачи. Как и в [1, 2], на длине n рассматривается двоичный симметричный канал ДСК(p) с входным и выходным алфавитами $E = \{0, 1\}$ и неизвестной переходной вероятностью p. Для различения входного и выходного множеств блоков $E^n = \{0, 1\}^n$ канала будем обозначать их через $E^n_{\rm in}$ и $E^n_{\rm out}$ соответственно. Относительно величины p имеются две гипотезы (одна из которых верна): $H_0 : p = p_0$ и $H_1 : p = p_1$, где $0 < p_0, p_1 \leqslant 1/2$.

Обозначим через $\mathbf{P}(\boldsymbol{y} \,|\, \boldsymbol{x})$ и $\mathbf{Q}(\boldsymbol{y} \,|\, \boldsymbol{x})$ вероятности получить на выходе канала блок $\boldsymbol{y} = (y_1, \ldots, y_n)$ при условии, что входным был блок $\boldsymbol{x} = (x_1, \ldots, x_n)$ для гипотез H_0 и H_1 соответственно. Тогда

$$\mathbf{P}(\boldsymbol{y} \,|\, \boldsymbol{x}) = (1 - p_0)^{n - d(\boldsymbol{x}, \boldsymbol{y})} p_0^{d(\boldsymbol{x}, \boldsymbol{y})}, \quad \mathbf{Q}(\boldsymbol{y} \,|\, \boldsymbol{x}) = (1 - p_1)^{n - d(\boldsymbol{x}, \boldsymbol{y})} p_1^{d(\boldsymbol{x}, \boldsymbol{y})},$$

где $d(\boldsymbol{x}, \boldsymbol{y})$ – расстояние Хэмминга между блоками $\boldsymbol{x}$ и $\boldsymbol{y}$ (т.е. число несовпадающих компонент этих векторов на всей длине n).

Рассматривается следующая задача минимаксного различения гипотез H_0 и H_1. Мы (т.е. "статистик") наблюдаем только блок $\boldsymbol{y} \in E^n_{\rm out}$ на выходе канала, а наш помощник ("helper") наблюдает только блок $\boldsymbol{x} \in E^n_{\rm in}$ на входе канала. Предполагается, что у нас нет никакой априорной информации о входном блоке $\boldsymbol{x}$. Ясно, что основываясь только на выходном блоке $\boldsymbol{y}$, мы не можем сделать никаких содержательных заключений относительно неизвестной величины p.

Предположим далее, что для заданной величины $R > 0$ нашему помощнику разрешается заранее разбить все входное пространство $E^n_{\rm in} = \{0, 1\}^n$ на $N \leqslant 2^{Rn}$ произвольных частей $\{X_1, \ldots, X_N\}$ и сообщить нам (каким-то дополнительным образом)

[1] Работа выполнена при частичной финансовой поддержке Российского фонда фундаментальных исследований (номер проекта 19-01-00364).

только то, какой части X_i принадлежит входной блок $\boldsymbol{x}$. Ясно, что только случай $N < 2^n$, т.е. $R < 1$, является интересным (иначе помощник может просто сообщить нам блок $\boldsymbol{x}$).

Например, помощник может сообщить статистику точные значения первых Rn величин $x_1, \ldots, x_{Rn}$ (но тогда ничего не сообщить о последующих величинах x_i). Однако такой простой способ разбиения входного пространства E^n_{in} (на цилиндрические множества $\{X_i\}$) не является, вообще говоря, оптимальным. С точки зрения статистика входные данные $(x_1, \ldots, x_n)$ представляют собой очень сильный мешающий вектор.

Есть много практических ситуаций, где встречается подобная задача. Например, в некоторых приложениях входной блок $\boldsymbol{x} \in E^n_{\text{in}}$ представляет собой "мешающий шум", который "загрязнил" выходной блок $\boldsymbol{y} \in E^n_{\text{out}}$, и поэтому нам хотелось бы "уменьшить" (по возможности) это "загрязнение", для того чтобы улучшить статистические выводы. Конечно, при этом очень важно качество канала связи от помощника к статистику. Для упрощения задачи мы рассматриваем здесь только идеализированный случай бесшумного канала с ограниченной пропускной способностью.

Можно также сказать, что оптимальная ограниченная информация о блоке $\boldsymbol{x} \in E^n_{\text{in}}$ означает оптимальное "сжатие" полной информации о блоке $\boldsymbol{x}$. Конечно, это оптимальное "сжатие" зависит от имеющейся априорной информации о переходной вероятности p и используемого критерия качества.

Замечание 1. Ясно, что задача не изменится, если, наоборот, статистик наблюдает вход, а помощник – выход канала.

Основываясь на наблюдении $\boldsymbol{y} \in E^n_{\text{out}}$ и номере (индексе) i части X_i, статистик принимает решение в пользу одной из гипотез H_0 или H_1. Для того чтобы избежать излишних усложнений, рассмотрим только нерандомизованные методы принятия решения (при этом существо задачи и результаты сохраняются).

Нас интересуют разбиения $\{X_1, \ldots, X_N\}$ и методы принятия решения, которые являются асимптотически (при $n \to \infty$) оптимальными. Аналогичные, но значительно более общие постановки такой задачи рассматривались, например, в [3–8].

Замечание 2. Забегая вперед, отметим, что насколько нам известно, все результаты в этой области (см., например, [1–8]) имеют вид "можно получить следующие характеристики проверки гипотез: ...". Нашей целью являются противоположные результаты, т.е. показать, что "нельзя получить характеристики лучше, чем ...".

Всюду далее $\log x = \log_2 x$. Введем шары и сферы в E^n:

$$\begin{aligned} &\mathbf{B}_{\boldsymbol{x}}(p) = \{\boldsymbol{u} : \ d(\boldsymbol{x}, \boldsymbol{u}) \leqslant pn\}, \\ &\mathbf{S}_{\boldsymbol{x}}(p) = \{\boldsymbol{u} : \ d(\boldsymbol{x}, \boldsymbol{u}) = pn\}, \end{aligned} \qquad \boldsymbol{x}, \boldsymbol{u} \in E^n. \tag{1}$$

2. Экспоненты вероятностей ошибки и дуальная задача. Пусть выбрано разбиение $\{X_1, \ldots, X_N\}$ входного пространства $E^n_{\text{in}} = \{0,1\}^n$. Тогда общее правило принятия решения можно описать следующим образом. Для каждого элемента разбиения X_i выбирается некоторое множество $\mathcal{A}(X_i) \subset E^n_{\text{out}}$, и далее, основываясь на наблюдении $\boldsymbol{y}$ и известном X_i, принимается решение ($\mathcal{A}^c = E^n_{\text{out}} \setminus \mathcal{A}$):

$$\boldsymbol{y} \in \mathcal{A}(X_i) \implies H_0, \quad \boldsymbol{y} \in \mathcal{A}^c(X_i) \implies H_1.$$

Определим вероятности ошибки 1-го рода α_n и 2-го рода β_n:

$$\begin{aligned} \alpha_n &= \Pr(H_1 \,|\, H_0) = \max_{i=1,\ldots,N} \max_{\boldsymbol{x} \in X_i} \mathbf{P}\left(\mathcal{A}^c(X_i) \,|\, \boldsymbol{x}\right), \\ \beta_n &= \Pr(H_0 \,|\, H_1) = \max_{i=1,\ldots,N} \max_{\boldsymbol{x} \in X_i} \mathbf{Q}\left(\mathcal{A}(X_i) \,|\, \boldsymbol{x}\right). \end{aligned}$$

Пусть далее $\gamma \geqslant 0$ – заданная величина. Будем требовать, чтобы для вероятности ошибки 1-го рода α_n выполнялось условие

$$\alpha_n = \Pr(H_1 \,|\, H_0) \leqslant 2^{-\gamma n}. \tag{2}$$

Нас интересует минимально возможная (по всем разбиениям $\{X_i\}$ входного пространства $E_{\rm in}^n$ и всем решениям) вероятность ошибки 2-го рода $\min \beta_n$. Мы исследуем асимптотический случай, когда $n \to \infty$ и $N = 2^{Rn}$, где $0 < R < 1$ – заданная постоянная[2]. Тогда для наилучших разбиения $\{X_i\}$ и решения обозначим

$$e(\gamma, R) = \lim_{n\to\infty} \frac{1}{n} \log_2 \frac{1}{\min \beta_n} > 0, \tag{3}$$

где минимум берется по всем разбиениям $\{X_i\}$ и решениям, удовлетворяющим условию (2).

Основной нашей целью являются оценки сверху для функции $e(\gamma, R)$ (оценки снизу см. в [1]). В данной статье мы ограничиваемся случаем $\gamma \to 0$, исследуя функцию $e(0, R) = e(R)$ и связанную с ней функцию $r_{\rm crit}(p_0, p_1)$ (иногда этот случай называют задачей Неймана – Пирсона). В отдельной работе мы рассмотрим случай $\gamma > 0$.

Для нас будет удобно рассмотреть также эквивалентную дуальную задачу (без помощника). Пусть задана величина r, $0 < r < 1$, и нам разрешается заранее выбрать любое множество $\mathcal{X} \subset E_{\rm in}^n$, состоящее из $X = 2^{rn}$ входных блоков. Известно также, что входной блок $\boldsymbol{x}$ принадлежит выбранному множеству $\mathcal{X}$. Мы наблюдаем выход канала $\boldsymbol{y}$ и, зная множество $\mathcal{X}$, рассматриваем задачу проверки гипотез H_0 и H_1. Далее мы выбираем множество $\mathcal{A}$ и в зависимости от наблюдения $\boldsymbol{y}$ принимаем решение:

$$\boldsymbol{y} \in \mathcal{A} \implies H_0, \quad \boldsymbol{y} \in \mathcal{A}^c \implies H_1.$$

Вероятности ошибок 1-го рода α_n и 2-го рода β_n определяются как

$$\alpha_n = \max_{\boldsymbol{x}\in\mathcal{X}} \mathbf{P}\left(\mathcal{A}^c \,|\, \boldsymbol{x}\right), \quad \beta_n = \max_{\boldsymbol{x}\in\mathcal{X}} \mathbf{Q}\left(\mathcal{A} \,|\, \boldsymbol{x}\right).$$

Пусть для вероятности ошибки 1-го рода α_n выполняется условие (2), и мы хотим выбрать множество $\mathcal{X} \subset E_{\rm in}^n$ мощности $X = 2^{rn}$ и правило принятия решения таким образом, чтобы достичь минимально возможной вероятности ошибки 2-го рода $\min \beta_n$. Для этой дуальной задачи аналогично (3) определим функцию

$$e_{\rm d}(\gamma, r) = \lim_{n\to\infty} \frac{1}{n} \log_2 \frac{1}{\min \beta_n} > 0, \tag{4}$$

где минимум берется по всем множествам $\mathcal{X} \subset E_{\rm in}^n$ мощности $X = 2^{rn}$ и всем решениям.

Следующий результат устанавливает простую связь между функциями $e(\gamma, R)$ и $e_{\rm d}(\gamma, r)$.

Предложение 1 [1, предложение 1]. *Справедливо соотношение*

$$e(\gamma, 1 - R) = e_{\rm d}(\gamma, R), \quad 0 \leqslant R \leqslant 1, \quad \gamma \geqslant 0. \tag{5}$$

В силу предложения 1 и формулы (5) достаточно исследовать функцию $e_{\rm d}(\gamma, r)$. В данной статье мы ограничимся случаем $\gamma \to 0$, исследуя функцию $e_{\rm d}(0, r)$.

[2] Для упрощения формул здесь и далее не будем использовать знак целой части.

Замечание 3. По существу в статье рассматривается случай, когда распределения $P(x,y)$ и $Q(x,y)$ имеют вид $P(x,y) = p(x)P(y\,|\,x)$ и $Q(x,y) = p(x)Q(y\,|\,x)$, где распределение $p(x)$ одно и то же для $P(x,y)$ и $Q(x,y)$. В более общей постановке задачи это может быть не так.

3. Известный входной блок. Предположим, что мы знаем входной блок $\boldsymbol{x} \in E_{\mathrm{in}}^n$ (тогда можно считать, что $\boldsymbol{x} = \mathbf{0}$) и наблюдаем выходной блок $\boldsymbol{y} \in E_{\mathrm{out}}^n$. Если требуется только $\alpha_n \to 0$, $n \to \infty$ (т.е. $\gamma = 0$), и нас интересует только экспонента (по n) вероятности ошибки 2-го рода β_n, то при $n \to \infty$ в силу центральной предельной теоремы (или в силу леммы Неймана – Пирсона) оптимальным множеством решения в пользу H_0 (т.е. p_0) является сферический слой $\mathbf{B_0}(p_0 + \delta) \setminus \mathbf{B_0}(p_0 - \delta)$ в E_{out}^n (см. (1)), где $\delta > 0$ мало. Тогда для экспоненты (по n) вероятности ошибки 2-го рода β_n имеем

$$\frac{1}{n}\log\beta_n = \frac{1}{n}\log\left[\binom{n}{p_0 n}(1-p_1)^{(1-p_0)n}p_1^{p_0 n}\right] + o(1), \quad n \to \infty,$$

и поэтому при $n \to \infty$ получаем

$$\frac{1}{n}\log\frac{1}{\beta_n} = -(1-p_0)\log(1-p_1) - p_0\log p_1 - h(p_0) + o(1) = D(p_0\,\|\,p_1) + o(1), \quad (6)$$

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

$$D(a\,\|\,b) = a\log\frac{a}{b} + (1-a)\log\frac{1-a}{1-b}. \quad (7)$$

Замечание 4. Величина $D(a\,\|\,b)$ из (7) представляет собой расхождение (divergence) для двух бернуллиевских случайных величин с параметрами a и b соответственно. В русскоязычной литературе $D(a\,\|\,b)$ чаще называется расстоянием Кульбака – Лейблера. Величина $D(a\,\|\,b)$ дает наилучшую экспоненту для вероятности ошибки 2-го рода при заданной вероятности ошибки 1-го рода (т.е. когда ее экспонента равна нулю) при проверке простой гипотезы $H_0\colon p = a$ против простой альтернативы $H_1\colon p = b$.

При $\gamma = r = 0$ для величины $e_{\mathrm{d}}(0,0)$ (см. (4)) из (6) получаем

$$e_{\mathrm{d}}(0,0) = D(p_1\,\|\,p_0). \quad (8)$$

4. Неизвестный входной блок и критическая скорость. Если мы знаем входной блок $\boldsymbol{x} \in E_{\mathrm{in}}^n$ и $\alpha_n \to 0$, то наилучшая экспонента $e_{\mathrm{d}}(0,0)$ вероятности ошибки 2-го рода β_n дается формулой (8).

Если же мы знаем только, что входной блок $\boldsymbol{x}$ принадлежит множеству $\mathcal{X} \subseteq E_{\mathrm{in}}^n$ мощности $X \sim 2^{rn}$, то для наилучшего такого множества $\mathcal{X}$ экспонента $e_{\mathrm{d}}(0,r)$ вероятности ошибки 2-го рода β_n определяется формулой (4). Ясно, что

$$e_{\mathrm{d}}(\gamma,r) \leqslant e_{\mathrm{d}}(\gamma,0), \quad \gamma \geqslant 0, \quad 0 \leqslant r \leqslant 1. \quad (9)$$

Функция $e_{\mathrm{d}}(\gamma,r)$ не возрастает по r. Поэтому возникает естественный вопрос: существует ли $r(\gamma) > 0$, для которого в (9) выполняется равенство, и если да, то какова максимальная такая скорость $r_{\mathrm{crit}}(\gamma)$? Ограничиваясь случаем $\gamma = 0$, определим $r_{\mathrm{crit}}(p_0,p_1) = r_{\mathrm{crit}}(p_0,p_1,0)$ как (см. (8))

$$r_{\mathrm{crit}} = r_{\mathrm{crit}}(p_0,p_1) = \sup\{r:\ e_{\mathrm{d}}(0,r) = e_{\mathrm{d}}(0,0) = D(p_0\,\|\,p_1)\}. \quad (10)$$

Иными словами, какова наибольшая мощность 2^{rn} "наилучшего" множества $\mathcal{X}$, для которого можно достичь такой же асимптотической эффективности, как и при известном входном блоке $\boldsymbol{x}$ (хотя мы и не знаем входной блок $\boldsymbol{x}$)?

Аналогично введем критическую скорость R_{crit} для исходной задачи (см. (3))

$$R_{\mathrm{crit}}(p_0, p_1) = \inf\{R:\ e(0, R) = e(0, 1) = D(p_0 \,\|\, p_1)\}. \tag{11}$$

В силу предложения 1 и (11) имеем

$$R_{\mathrm{crit}}(p_0, p_1) = 1 - r_{\mathrm{crit}}(p_0, p_1). \tag{12}$$

Основной результат статьи составляет

Теорема 1. *Если $p_1 < p_0 \leqslant 1/2$, то существует $p_1^*(p_0) \leqslant p_0$, такое что для любого $p_1 \leqslant p_1^*(p_0)$ справедлива формула*

$$r_{\mathrm{crit}}(p_0, p_1) = 1 - R_{\mathrm{crit}}(p_0, p_1) = 1 - h(p_0), \quad 0 < p_1 \leqslant p_1^* < p_0 \leqslant 1/2. \tag{13}$$

Замечание 5. Хотя величина $r_{\mathrm{crit}}(p_0, p_1)$ в (13) совпадает с пропускной способностью канала ДСК(p_0), ее происхождение (10) связано с функцией $e_{\mathrm{d}}(0, r)$, аналогичной функции надежности $E(r, p)$ в теории информации [9, 10]. При этом точный вид функции $E(r, p)$ до сих пор известен только частично [11]. Поэтому, как и в [11–13], в доказательстве теоремы 1 используются достаточно недавние результаты о спектре двоичных кодов. Полное описание функции $e_{\mathrm{d}}(\gamma, r)$ выглядит трудной задачей.

В § 2 приводится граница снизу для r_{crit} (предложение 2). В § 3 выводится общая формула для вероятности ошибки 2-го рода β_n (лемма 1). В § 4, используя метод "двух гипотез", доказывается теорема 1. Но граница сверху (13) для r_{crit}, вообще говоря, слабее соответствующей границы снизу из § 2. В § 5 с помощью дополнительных комбинаторных соображений выводится еще одна граница сверху для r_{crit} (предложение 3). В § 6 показывается точность границы снизу для r_{crit} из предложения 2 при условии, что выполняется некоторое дополнительное условие. В Приложении приводятся некоторые необходимые аналитические результаты.

В статье $f \sim g$ означает, что $n^{-1} \ln f = n^{-1} \ln g + o(1)$, $n \to \infty$, а $f \lesssim g$ означает $n^{-1} \ln f \leqslant n^{-1} \ln g + o(1)$, $n \to \infty$.

§ 2. Граница снизу для r_{crit}

Следующий результат следует из [1, предложение 2].

Предложение 2. *Для $r_{\mathrm{crit}}(p_0, p_1)$ справедливы оценки снизу*

$$r_{\mathrm{crit}}(p_0, p_1) \geqslant 1 - h(p_0), \quad \textit{если} \quad 0 < p_1 < p_0 \leqslant 1/2, \tag{14}$$

и

$$r_{\mathrm{crit}}(p_0, p_1) \geqslant 1 - h(p_0) - D(p_0 \,\|\, p_1), \quad \textit{если} \quad 0 < p_0 < p_1 \leqslant 1/2. \tag{15}$$

Доказательство. Для заданного r, $0 < r < 1$, выберем случайно и равновероятно множество $\mathcal{X}$ из $X = 2^{rn}$ входных блоков $\boldsymbol{x}$. В [1, предложение 2] было показано, что если $p_0 < p_1 \leqslant 1/2$, то для любого τ, $p_0 \leqslant \tau \leqslant p_1$, существует множество $\mathcal{X}$ и метод принятия решения, для которого выполняются неравенства

$$\frac{1}{n} \log \frac{1}{\alpha_n} \geqslant D(\tau \,\|\, p_0), \quad \frac{1}{n} \log \frac{1}{\beta_n} \geqslant \min\{D(\tau \,\|\, p_1), 1 - h(\tau) - r\}. \tag{16}$$

Если достаточно иметь $\alpha_n \to 0$, $n \to \infty$, то полагая в (16) $\tau = p_0$, из (10) получаем (15).

Аналогично, если $p_1 < p_0 \leqslant 1/2$, то меняя в (16) местами p_0 с p_1 и α_n с β_n, для любого τ имеем

$$\frac{1}{n}\log\frac{1}{\alpha_n} \geqslant \min\{D(\tau\,\|\,p_0), 1-h(\tau)-r\}, \quad \frac{1}{n}\log\frac{1}{\beta_n} \geqslant D(\tau\,\|\,p_1). \tag{17}$$

Если $\alpha_n \to 0$, $n\to\infty$, то полагая в (17) $\tau = p_0$, из (10) получаем (14). ▲

§ 3. Общая формула для вероятности ошибки 2-го рода β_n

Пусть $\mathcal{C}_n(r) = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_M\}$ – множество (код) из $M = 2^{rn}$ различных входных (кодовых) блоков. Для кода $\mathcal{C}_n(r)$ и вероятности ошибки 1-го рода α_n обозначим через $\mathcal{D}_0 = \mathcal{D}_0(\mathcal{C}_n, \alpha_n) \subseteq E^n_{\text{out}}$ оптимальное множество решения в пользу H_0, минимизирующее вероятность ошибки 2-го рода β_n. Хотя множество $\mathcal{D}_0$ имеет довольно сложный вид, можно установить некоторые его свойства, достаточные для доказательства теоремы 1.

Выберем малое $\delta > 0$, и для каждого $\boldsymbol{x}_k$, $k = 1, \ldots, M$, введем сферический слой в E^n_{out}

$$SL_{\boldsymbol{x}_k}(p_0, \delta) = \mathbf{B}_{\boldsymbol{x}_k}(p_0+\delta) \setminus \mathbf{B}_{\boldsymbol{x}_k}(p_0-\delta) = \{\boldsymbol{u} :\ |d(\boldsymbol{x}_k, \boldsymbol{u}) - p_0 n| \leqslant \delta n\}, \tag{18}$$

где $\mathbf{B}_{\boldsymbol{x}}(p)$ определено в (1). Для каждого $\boldsymbol{x}_k$ введем также множество

$$D_{\boldsymbol{x}_k}(\delta) = \mathcal{D}_0 \cap SL_{\boldsymbol{x}_k}(p_0, \delta). \tag{19}$$

Так как необходимо иметь $\alpha_n \to 0$, $n\to\infty$, то оптимальное множество $\mathcal{D}_0$ содержит "существенную" часть каждого множества $SL_{\boldsymbol{x}_k}(p_0,\delta)$, $k = 1, \ldots, M$. Для того чтобы оценить это, заметим, что для любых $\boldsymbol{x}_k$ и $\boldsymbol{u}, \boldsymbol{z} \in SL_{\boldsymbol{x}_k}(p_0,\delta)$ имеем

$$\frac{\mathbf{P}(\boldsymbol{u}\,|\,p_0, \boldsymbol{x}_k)}{\mathbf{P}(\boldsymbol{z}\,|\,p_0, \boldsymbol{x}_k)} = \left(\frac{q_0}{p_0}\right)^{d(\boldsymbol{z},\boldsymbol{x}_k) - d(\boldsymbol{u},\boldsymbol{x}_k)} \leqslant \left(\frac{q_0}{p_0}\right)^{2\delta n}, \quad q_0 = 1-p_0. \tag{20}$$

По экспоненциальному неравенству Чебышева (граница Чернова) для любого $\boldsymbol{x}_k$ и малых $\delta > 0$ получаем

$$\log \mathbf{P}\{\boldsymbol{u} \notin SL_{\boldsymbol{x}_k}(p_0,\delta)\,|\,\boldsymbol{x}_k, p_0\} \leqslant -\frac{n\delta^2}{2p_0q_0}. \tag{21}$$

Тогда в силу (18), (19) и (21) для любого $\boldsymbol{x}_k$ имеем

$$\begin{aligned}&\mathbf{P}\{D_{\boldsymbol{x}_k}(\delta)\,|\,p_0, \boldsymbol{x}_k\} \geqslant 1 - \mathbf{P}\{\boldsymbol{u} \notin \mathcal{D}_0\,|\,p_0, \boldsymbol{x}_k\} - \mathbf{P}\{\boldsymbol{u} \notin SL_{\boldsymbol{x}_k}(p_0,\delta)\,|\,p_0, \boldsymbol{x}_k\} \geqslant \\ &\geqslant 1 - \alpha_n - e^{-n^2\delta^2/(2p_0q_0)},\end{aligned} \tag{22}$$

а в силу (20) также имеем

$$\begin{aligned}&\delta_1 |SL_{\boldsymbol{x}_k}(p_0,\delta)| \leqslant |D_{\boldsymbol{x}_k}(\delta)| \leqslant |SL_{\boldsymbol{x}_k}(p_0,\delta)|, \\ &\delta_1 = \left(1 - \beta_n - e^{-n^2\delta^2/(2p_0q_0)}\right)\left(\frac{p_0}{q_0}\right)^{2\delta n}.\end{aligned} \tag{23}$$

Так как $D_{\boldsymbol{x}_k}(\delta) \subseteq \mathcal{D}_0$ для любого $\boldsymbol{x}_k$, то в силу (19), (22) и (23) для вероятности $\mathbf{P}(e\,|\,p_1, \boldsymbol{x}_i)$ имеем

$$\mathbf{P}(e\,|\,p_1, \boldsymbol{x}_i) = \mathbf{P}\{\mathcal{D}_0\,|\,p_1, \boldsymbol{x}_i\} \sim \mathbf{P}\left\{\bigcup_{k=1}^{M} D_{\boldsymbol{x}_k}(\delta)\,|\,p_1, \boldsymbol{x}_k\right\} \sim$$

$$\sim \delta_1\,\mathbf{P}\left\{\bigcup_{k=1}^{M} SL_{\boldsymbol{x}_k}(p_0, \delta)\,|\,p_1, \boldsymbol{x}_i\right\}. \tag{24}$$

Для $t > 0$ и каждого $\boldsymbol{x}_i$ введем множество

$$D_{\boldsymbol{x}_i}(t, p) =$$
$$= \{\boldsymbol{u}:\ \text{существует } \boldsymbol{x}_k \neq \boldsymbol{x}_i,\ \text{такое что } d(\boldsymbol{x}_i, \boldsymbol{u}) = tn,\ d(\boldsymbol{x}_k, \boldsymbol{u}) = pn\}. \tag{25}$$

Лемма 1. *Для вероятности ошибки 2-го рода β_n кода $\mathcal{C}_n = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_M\}$ и оптимального решения $\mathcal{D}_0$ в пользу H_0 при $n \to \infty$ справедлива формула*

$$\frac{\log \beta_n}{n} \sim \max_{t>0}\left\{\frac{1}{n}\log\left[\frac{1}{M}\sum_{i=1}^{M}|D_{\boldsymbol{x}_i}(t, p_0)|\right] + t\log p_1 + (1-t)\log(1-p_1)\right\}. \tag{26}$$

Критическая скорость $r_{\mathrm{crit}}(p_0, p_1)$ определяется формулой ($M = 2^{rn}$)

$$r_{\mathrm{crit}}(p_0, p_1) = \sup\{r : F(p_0, p_1, r) \leqslant 0\} = \inf\{r : F(p_0, p_1, r) > 0\}, \tag{27}$$

где

$$F(p_0, p_1, r) = \lim_{n\to\infty} \min_{|\mathcal{C}_n| \leqslant M} \max_{t} F(p_0, p_1, r, \mathcal{C}_n, t),$$
$$F(p_0, p_1, r, \mathcal{C}_n, t) = \frac{1}{n}\log\left[\sum_{i=1}^{M}|D_{\boldsymbol{x}_i}(t, p_0)|\right] + (p_0 - t)\log\frac{1-p_1}{p_1} - r - h(p_0). \tag{28}$$

Доказательство. Используя (24) при $\delta = o(1)$ и $\delta_1 = e^{o(n)}$ при $n \to \infty$, имеем

$$\beta_n = \max_i \mathbf{P}(e\,|\,p_1, \boldsymbol{x}_i) \sim \frac{1}{M}\sum_{i=1}^{M}\mathbf{P}(e\,|\,p_1, \boldsymbol{x}_i) \sim$$

$$\sim \frac{\delta_1}{M}\sum_{i=1}^{M}\mathbf{P}\left\{\bigcup_{k=1}^{M} SL_{\boldsymbol{x}_k}(p_0, \delta)\,\big|\,p_1, \boldsymbol{x}_i\right\}. \tag{29}$$

Из (25) и (26) для каждого $\boldsymbol{x}_i$

$$\mathbf{P}\left\{\bigcup_{k=1}^{M} SL_{\boldsymbol{x}_k}(p_0, \delta)\,\big|\,p_1, \boldsymbol{x}_i\right\} \sim \mathbf{P}\left\{\bigcup_{t>0} D_{\boldsymbol{x}_i}(t, p_0)\,\big|\,p_1, \boldsymbol{x}_i\right\} \sim$$
$$\sim \max_{t>0}\left\{p_1^{tn}(1-p_1)^{(1-t)n}|D_{\boldsymbol{x}_i}(t, p_0|\right\}. \tag{30}$$

Поэтому из (29) и (30) следует формула (26).

Так как

$$\mathbf{P}\{SL_{\boldsymbol{x}_i}(p_0, \delta)\,|\,p_1, \boldsymbol{x}_i\} \sim \mathbf{P}\{d(\boldsymbol{x}_i, \boldsymbol{u}) \geqslant p_0 n\,|\,p_1, \boldsymbol{x}_i\} \sim 2^{-D(p_0\|p_1)n},$$

то правая часть (26) возрастает по r (т.е. по $M = 2^{rn}$), начиная с $-D(p_1\,\|\,p_0)$. Поэтому из (6) и (26) следует, что критическая скорость r_{crit} равна максимальной

скорости r, такой что

$$\min_{\{\boldsymbol{x}_i\}}\max_{t>0}\left\{\frac{1}{n}\log\left[\sum_{i=1}^{M}|D_{\boldsymbol{x}_i}(t,p_0)|\right]+t\log p_1+(1-t)\log(1-p_1)\right\}-r\leqslant$$
$$\leqslant -D(p_0\,\|\,p_1). \tag{31}$$

Заметим, что

$$D(p_0\,\|\,p_1)+t\log p_1+(1-t)\log(1-p_1)=-h(p_0)+(p_0-t)\log\frac{1-p_1}{p_1}. \tag{32}$$

Из (31) и (32) следуют формулы (27), (28). ▲

Отметим, в частности, что из (53) при $t=p_0$ имеем

$$F(p_0,p_1,r,\mathcal{C}_n,p_0)=o(1),\quad n\to\infty.$$

В анализе соотношений (27), (28) основную трудность составляет оценка мощностей $|D_{\boldsymbol{x}_i}(t,p_0)|$ в (28), которые зависят от геометрии кода $\mathcal{C}_n$. Аналогичная проблема возникала в [11–13] при исследовании функции надежности $E(R,p)$ канала ДСК(p). Прямая оценка этих мощностей ведет к весьма громоздким формулам.

§ 4. Граница сверху для r_{crit}: две гипотезы

Получим простую (но не очень точную) оценку сверху для $r_{\mathrm{crit}}(p_0,p_1)$, используя популярный в математической статистике (чаще в теории оценивания) метод "двух гипотез". Используя для этого формулу (26), выберем из кода $\mathcal{C}_n(r)=\{\boldsymbol{x}_1,\dots,\boldsymbol{x}_M\}$, $M=2^{rn}$, какие-либо два кодовых слова, скажем, $\boldsymbol{x}_1$ и $\boldsymbol{x}_2$ с $d(\boldsymbol{x}_1,\boldsymbol{x}_2)=\omega n$. Можно считать, что для скорости $r>0$ величина ω удовлетворяет ограничениям

$$0<\omega\leqslant\omega_{\min}(r),$$

где величина $\omega_{\min}(r)$ будет определена далее. Заменим код $\mathcal{C}_n(r)$ кодом $\mathcal{C}'$ из двух выбранных кодовых слов $\mathcal{C}'=\{\boldsymbol{x}_1,\boldsymbol{x}_2\}$. Тогда $\beta_n(\mathcal{C})\geqslant\beta_n(\mathcal{C}')$. Аналогично (29), (30) имеем

$$\beta_n(\mathcal{C}')\sim 2^{-D(p_0\|p_1)n}+\mathbf{P}\left\{SL_{\boldsymbol{x}_2}(p_0,\delta)\,\middle|\,p_1,\boldsymbol{x}_1\right\}.$$

Нас интересует, когда для $\boldsymbol{x}_1,\boldsymbol{x}_2$ справедливо неравенство

$$\frac{1}{n}\log\mathbf{P}\left\{SL_{\boldsymbol{x}_2}(p_0,\delta)\,\middle|\,p_1,\boldsymbol{x}_1\right\}>-D(p_0\,\|\,p_1). \tag{33}$$

Оценим вероятность в левой части (33). Для $d(\boldsymbol{x}_i,\boldsymbol{x}_k)=\omega n$ обозначим

$$S_{\boldsymbol{x}_i,\boldsymbol{x}_k}(t,p,\omega)=\{\boldsymbol{u}:\ d(\boldsymbol{x}_i,\boldsymbol{u})=tn,\ d(\boldsymbol{x}_k,\boldsymbol{u})=pn,\ d(\boldsymbol{x}_i,\boldsymbol{x}_k)=\omega n\}. \tag{34}$$

Тогда (см. Приложение)

$$\begin{aligned}&\frac{1}{n}\log|S_{\boldsymbol{x}_i,\boldsymbol{x}_k}(t,p,\omega)|=g(t,p,\omega)+o(1),\quad n\to\infty,\\&\frac{1}{n}\log\mathbf{P}\left\{S_{\boldsymbol{x}_i,\boldsymbol{x}_k}(t,p,\omega)\,\middle|\,p_1,\boldsymbol{x}_i\right\}=g(t,p,\omega)-t\log\frac{1-p_1}{p_1}+\log(1-p_1)+o(1),\end{aligned} \tag{35}$$

где $g(t,p,\omega)$ определено в (78). Поэтому при $n\to\infty$ (см. (76), (77))

$$\frac{1}{n}\log\mathbf{P}\left\{SL_{\boldsymbol{x}_2}(p_0,\delta)\,\middle|\,p_1,\boldsymbol{x}_1\right\}=$$
$$=\frac{1}{n}\max_t\log\mathbf{P}\left\{S_{\boldsymbol{x}_1,\boldsymbol{x}_2}(t,p_0,\omega)\,\middle|\,p_1,\boldsymbol{x}_1\right\}+o(1)=f(p_0,p_1,\omega)+o(1), \tag{36}$$

где

$$\begin{aligned}&f(p_0,p_1,\omega)=\max_t f(p_0,p_1,\omega,t),\\&f(p_0,p_1,\omega,t)=g(t,p_0,\omega)-t\log\frac{1-p_1}{p_1}+\log(1-p_1).\end{aligned} \tag{37}$$

Имеем

$$f'_t(p_0,p_1,\omega,t)=\log\frac{\omega-t}{t}-\log\frac{p_0+t-\omega}{1-p_0-t}-2\frac{1-p_1}{p_1},\quad f''_{tt}(p_0,p_1,\omega,t)<0. \tag{38}$$

В силу (32) и (35)–(37) неравенство (33) принимает вид

$$\max_t F(p_0,p_1,\omega,t)>0, \tag{39}$$

где

$$F(p_0,p_1,\omega,t)=f(p_0,p_1,\omega,t)+D(p_0\,\|\,p_1)=$$
$$=g(t,p_0,\omega)+(p_0-t)\log\frac{1-p_1}{p_1}-h(p_0). \tag{40}$$

Если для каких-либо p_0,p_1 и ω выполняется неравенство (39), то справедлива соответствующая граница сверху (14), (15). Обозначим через $t_1^0=t_1^0(p_0,p_1,\omega)$ максимизирующую величину t в (37) (она же остается максимизирующей в (39)). Тогда

$$f(p_0,p_1,\omega)=f(p_0,p_1,\omega,t_1^0(p_0,p_1,\omega)). \tag{41}$$

Из уравнения $f'_t(p_0,p_1,\omega,t)=0$ для t_1^0 из (38) получаем

$$\begin{aligned}&t_1^0=t_1^0(p_0,p_1,\omega)=\frac{\sqrt{1+(v_0-1)[(\omega-p_0)^2v_0-(1-\omega-p_0)^2+1]}-1}{v_0-1},\\&v_0(p_1)=\left(\frac{1-p_1}{p_1}\right)^2\geqslant 1.\end{aligned} \tag{42}$$

Тогда из (40) и (42) имеем

$$F(p_0,p_1,\omega,t_1^0)=g(t_1^0,p_0,\omega)+(p_0-t_1^0)\log\frac{1-p_1}{p_1}-h(p_0). \tag{43}$$

Можно проверить, что для функции $F(p_0,p_1,\omega,t_1^0)$ из (43) вытекают свойства $F(p_0,p_1,0,t_1^0)=0$ и $F''_{\omega\omega}<0$, $\omega>0$. Поэтому достаточно проверить неравенство (39) с $t=t_1^0$ только для минимальной для кода $\mathcal{C}_n(r)$ величины ω (т.е. для его кодового расстояния $d(\mathcal{C})$).

Пусть $\omega_{\min}(r)n$ – максимально возможное минимальное расстояние кода $\mathcal{C}_n(r)$. Для величины $\omega_{\min}(r)$ известна граница [14, формула (1.5)]

$$r\leqslant h\left[\frac{1}{2}-\sqrt{\omega_{\min}(1-\omega_{\min})}\right],\quad \omega_{\min}=\omega_{\min}(r). \tag{44}$$

Рассмотрим два возможных случая: 1) $p_1<p_0\leqslant 1/2$ и 2) $p_0<p_1\leqslant 1/2$.

1) Случай $p_1 < p_0 \leqslant 1/2$. Полагая $r = 1 - h(p_0)$, обозначим через $\omega_0 = \omega_0(p_0)$ корень уравнения (см. (44))

$$1 - h(p_0) = h\left[\frac{1}{2} - \sqrt{\omega(1-\omega)}\right].$$

Тогда неравенство (39) принимает вид ($\omega_0 = \omega_0(p_0)$)

$$F(p_0, p_1, \omega_0, t_1^0) = g(t_1^0, p_0, \omega_0) + (p_0 - t_1^0)\log\frac{1-p_1}{p_1} - h(p_0) > 0. \tag{45}$$

Можно проверить (с помощью Maple), что (45) выполняется, если $p_1 \leqslant p_1^*(p_0)$, где

p_0	0,1	0,12	0,15	0,2	0,3	0,4	0,45	0,49
$p_1^*(p_0)$	0,0003	0,003	0,016	0,056	0,17	0,317	0,4	0,48

.

Если $p_0 \leqslant 0{,}20707$ (т.е. $\omega < 0{,}273$), то в [14, формула (1.4)] имеется оценка чуть более точная (но более громоздкая), чем (44).

2) Случай $p_0 < p_1 \leqslant 1/2$. Можно проверить, что неравенство (39) не выполняется ни при каких $p_0 < p_1$!

§ 5. Граница сверху для r_{crit}: комбинаторика

Приведем еще одну границу сверху для r_{crit}, по-прежнему основанную на формуле (26), но использующую дополнительные комбинаторные соображения.

1. Комбинаторная лемма. В коде $\mathcal{C}_n = \{\boldsymbol{x}_i\}$ будем называть $(\boldsymbol{x}_i, \boldsymbol{x}_j)$ ω-парой, если $d(\boldsymbol{x}_i, \boldsymbol{x}_j) = \omega n$. Будем говорить, что точка $\boldsymbol{y} \in E^n$ является (ω, p, t)-покрытой, если существует ω-пара $(\boldsymbol{x}_i, \boldsymbol{x}_j)$, такая что $d(\boldsymbol{x}_i, \boldsymbol{y}) = pn$, $d(\boldsymbol{x}_j, \boldsymbol{y}) = tn$. Обозначим через $K(\boldsymbol{y}, \omega, p, t)$ число (ω, p, t)-покрытий точки $\boldsymbol{y}$ (учитывая кратность покрытий), т.е.

$$K(\boldsymbol{y}, \omega, p, t) = \\ = |\{(\boldsymbol{x}_i, \boldsymbol{x}_j) : d(\boldsymbol{x}_i, \boldsymbol{x}_j) = \omega n,\ d(\boldsymbol{x}_i, \boldsymbol{y}) = pn,\ d(\boldsymbol{x}_j, \boldsymbol{y}) = tn\}|, \quad \omega > 0. \tag{46}$$

Введем множества (ср. (25))

$$D_{\boldsymbol{x}_i}(t, p, \omega) = \bigcup_{\boldsymbol{x}_k} S_{\boldsymbol{x}_i, \boldsymbol{x}_k}(t, p, \omega) = \{\boldsymbol{u} : \text{ существует } \boldsymbol{x}_k, \\ \text{такое что } d(\boldsymbol{x}_i, \boldsymbol{x}_k) = \omega n,\ d(\boldsymbol{x}_i, \boldsymbol{u}) = tn,\ d(\boldsymbol{x}_k, \boldsymbol{u}) = pn\}. \tag{47}$$

Тогда

$$D_{\boldsymbol{x}_i}(t, p) = \bigcup_{\omega > 0} D_{\boldsymbol{x}_i}(t, p, \omega).$$

Для $t > 0$ введем величину

$$m_t(\boldsymbol{y}) = |\{\boldsymbol{x}_i : \boldsymbol{x}_i \in \mathbf{S}_{\boldsymbol{y}}(t)\}|. \tag{48}$$

Тогда для любых $\boldsymbol{y}, p, t > 0$

$$K(\boldsymbol{y}, t, p) = m_t(\boldsymbol{y}) m_p(\boldsymbol{y}). \tag{49}$$

Лемма 2. *Для кода* $\{\boldsymbol{x}_i\}$ *и* $\omega, p, t > 0$ *справедлива формула* (*см.* (46) *и* (47))

$$\sum_{i=1}^{M} |D_{\boldsymbol{x}_i}(t,p,\omega)| \leqslant \sum_{\boldsymbol{y}\in E^n} K(\boldsymbol{y},\omega,t,p). \tag{50}$$

Также, если (*см.* (48))

$$\max_{\boldsymbol{y}} m_p(\boldsymbol{y}) = 2^{o(n)}, \quad n\to\infty, \tag{51}$$

то для любых $\omega, t > 0$

$$\sum_{i=1}^{M} |D_{\boldsymbol{x}_i}(t,p,\omega)| = 2^{o(n)} \sum_{\boldsymbol{y}\in E^n} K(\boldsymbol{y},\omega,t,p), \quad n\to\infty. \tag{52}$$

Доказательство. Пусть $\boldsymbol{y}\in E^n$ и имеется m упорядоченных пар $(\boldsymbol{x}_i,\boldsymbol{x}_j)$ с $d(\boldsymbol{x}_i,\boldsymbol{x}_j)=\omega n$ и $d(\boldsymbol{x}_i,\boldsymbol{y})=tn$, $d(\boldsymbol{x}_j,\boldsymbol{y})=pn$. Эти m пар $(\boldsymbol{x}_i,\boldsymbol{x}_j)$ имеют $m_1\leqslant m$ различных первых аргументов $\{\boldsymbol{x}_i\}$. Тогда $\boldsymbol{y}$ присутствует m раз в правой части (50) и m_1 раз в левой части, что доказывает формулу (50). Если выполнено условие (51), то $m_1 = me^{o(n)}$, откуда следует равенство (52). Отметим также, что в силу (49) имеем

$$\sum_{i=1}^{M} |D_{\boldsymbol{x}_i}(t,p)| = \sum_{\boldsymbol{y}:\, m_p(\boldsymbol{y})\geqslant 1} \frac{K(\boldsymbol{y},t,p)}{m_p(\boldsymbol{y})} = \sum_{\boldsymbol{y}:\, m_p(\boldsymbol{y})\geqslant 1} m_t(\boldsymbol{y}) \sim$$
$$\sim M2^{h(t)n} - \sum_{\boldsymbol{y}:\, m_p(\boldsymbol{y})=0} m_t(\boldsymbol{y}). \tag{53}$$

Из первого равенства в (53) также следуют формулы (50) и (52). ▲

Формула (53) выглядит простой и привлекательной, однако ее правая часть имеет вид "большое минус большое", что неудобно. Отметим, что в (53) нельзя пренебрегать последней суммой, так как тогда получим только $r_{\mathrm{crit}}\leqslant 1$, что неинтересно.

2. Еще одна граница сверху для r_{crit}. Оценим сверху последнюю сумму в (53) следующим образом. Имеем

$$\sum_{\boldsymbol{y}:\, m_{p_0}(\boldsymbol{y})=0} m_t(\boldsymbol{y}) \leqslant 2^{h(t)n} |\{\boldsymbol{y}:\ m_{p_0}(\boldsymbol{y})=0\}|. \tag{54}$$

Максимум мощности $|\{\boldsymbol{y}:\ m_{p_0}(\boldsymbol{y})=0\}|$ достигается, когда код $\mathcal{C}$ является шаром $\mathbf{B_0}(\tau)$ радиуса τn, где $r=h(\tau)$. Поэтому

$$\begin{aligned} &\max_{\mathcal{C}} |\{\boldsymbol{y}:\ m_{p_0}(\boldsymbol{y})=0\}| = 2^n - |\mathbf{B_0}(\tau+p_0)| \sim 2^{h(\tau+p_0)n}, \quad \tau+p_0\geqslant 1/2,\\ &\max_{\mathcal{C}} |\{\boldsymbol{y}:\ m_{p_0}(\boldsymbol{y})=0\}| \sim 2^n, \quad \tau+p_0\leqslant 1/2. \end{aligned} \tag{55}$$

Если $\tau+p_0\geqslant 1/2$, т.е. если $r\geqslant h(1/2-p_0)$, то из (53)–(55) получаем

$$\sum_{i=1}^{M} |D_{\boldsymbol{x}_i}(t,p_0)| \geqslant 2^{h(t)n}\big[M - 2^{h(\tau+p_0)n}\big] = 2^{h(t)n}\big[2^{h(\tau)n} - 2^{h(1-\tau-p_0)n}\big] \sim M2^{h(t)n},$$

если $\tau > 1-\tau-p_0$, т.е. $\tau > (1-p_0)/2$, или, эквивалентно, если $r > h[(1-p_0)/2]$.

Поэтому если $r \geqslant \max\{h(1/2-p_0), h[(1-p_0)/2]\} = h[(1-p_0)/2]$, то при любом $p_0 \neq p_1$ равенство (28) принимает вид

$$F(p_0,p_1,r) = \max_{t>0}\left\{h(t) + (p_0-t)\log\frac{1-p_1}{p_1}\right\} - h(p_0) =$$
$$= h(p_1) + (p_0-p_1)\log\frac{1-p_1}{p_1} - h(p_0) > 0, \quad p_0 \neq p_1,$$

так как максимум по t достигается при $t = p_1$. Поэтому это дает следующую границу сверху для r_{crit} (более слабую, чем (13)):

$$r_{\mathrm{crit}}(p_0,p_1) \leqslant h[(1-p_0)/2], \quad p_0 \neq p_1. \tag{56}$$

Замечание 6. Отметим, что $1 - h(p_0) < h(1/2-p_0) < h[(1-p_0)/2]$, $0 < p_0 < 1/2$.

Усилим оценку (56). Наряду с (54) также имеем

$$\sum_{\boldsymbol{y}:\, m_{p_0}(\boldsymbol{y})=0} m_t(\boldsymbol{y}) \leqslant M|\{\boldsymbol{y} : m_{p_0}(\boldsymbol{y}) = 0\}|.$$

Поэтому если $\tau + p_0 \geqslant 1/2$ и $t \geqslant 1-\tau-p_0$, то

$$\sum_{i=1}^{M} |D_{\boldsymbol{x}_i}(t,p_0)| \geqslant M\left[2^{h(t)n} - 2^{h(1-\tau-p_0)n}\right] \sim M2^{h(t)n}.$$

В силу (39), (40) необходимо иметь

$$\begin{aligned} &\max_{t\geqslant 1-\tau-p_0} f(t,p_0,p_1) > 0,\\ &f(t,p_0,p_1) = h(t) + (p_0-t)\log\frac{1-p_1}{p_1} - h(p_0). \end{aligned} \tag{57}$$

Максимум по $t \geqslant 1-\tau-p_0$ функции $f(t,p_0,p_1)$ достигается при $t = \max\{p_1, 1-\tau-p_0\}$, так как

$$\begin{aligned} &\max_t f(t,p_0,p_1) = f(p_1,p_0,p_1) > 0, \quad p_0 \neq p_1, \quad f(p_0,p_0,p_1) = 0,\\ &f'_t(t,p_0,p_1) = \log\frac{1-t}{t} - \log\frac{1-p_1}{p_1}, \quad f''_{tt}(t,p_0,p_1) < 0,\\ &\operatorname{sign} f'_t(t,p_0,p_1) = \operatorname{sign}(p_1-t). \end{aligned} \tag{58}$$

Поэтому если $p_1 \geqslant 1-\tau-p_0$, то из (57), (58) для $p_0 \neq p_1$ получаем

$$\max_{t\geqslant 1-\tau-p_0} f(t,p_0,p_1) = h(p_1) + (p_0-p_1)\log\frac{1-p_1}{p_1} - h(p_0) > 0. \tag{59}$$

Тогда если $\tau \geqslant \max\{1/2-p_0, 1-p_0-p_1\} = 1-p_0-p_1$, то для $p_0 \neq p_1$ выполняется неравенство (59), откуда следует оценка

$$\tau_{\mathrm{crit}} \leqslant 1-p_0-p_1, \quad r_{\mathrm{crit}} = h(\tau_{\mathrm{crit}}). \tag{60}$$

Если же $p_1 < 1-\tau-p_0$, то максимум в (57) достигается при $t = 1-\tau-p_0$, и тогда

$$\max_{t\geqslant 1-\tau-p_0} f(t,p_0,p_1) = f(1-\tau-p_0,p_0,p_1).$$

Заметим, что

$$f(p_0,p_0,p_1)=0,\quad f'_{t=p_0}(t,p_0,p_1)\neq 0,\quad p_0\neq p_1,$$
$$f''_{tt}(t,p_0,p_1)<0,\quad \operatorname{sign} f'_t(t,p_0,p_1)=\operatorname{sign}(p_1-t).$$

Пусть также $p_0>1-\tau-p_0$ (т.е. $\tau>1-2p_0$). Тогда $\max\limits_{t\geqslant 1-\tau-p_0} f(t,p_0,p_1)>0$ (достаточно выбрать t близким к p_0). Тогда

$$\tau_{\mathrm{crit}}\leqslant 1-2p_0,\quad r_{\mathrm{crit}}=h(\tau_{\mathrm{crit}}). \tag{61}$$

В результате из (60) и (61) получаем

Предложение 3. *При любых $p_0,p_1\in[0,1/2]$ для r_{crit} справедлива оценка сверху*

$$\tau_{\mathrm{crit}}(p_0,p_1)\leqslant \min\{1-p_0-p_1,1-2p_0\},\quad r_{\mathrm{crit}}=h(\tau_{\mathrm{crit}}). \tag{62}$$

Следствие. *Если $p_0=1/2$, то из (62) следует $\tau_{\mathrm{crit}}(1/2,p_1)=r_{\mathrm{crit}}(1/2,p_1)=0$.*

Ранее этот частный результат был получен другим способом в [1, предложение 3]. Там же найдена наилучшая экспонента $e_{\mathrm{d}}(\gamma,r)$ из (4) для $\gamma\geqslant 0$, $0\leqslant r\leqslant 1$.

§ 6. "Потенциальная" аддитивная граница сверху для r_{crit}

Теорема 1 была доказана, заменяя в формуле (26) экспоненциальное число M кодовых слов $\{\boldsymbol{x}_i\}$ двумя ближайшими кодовыми словами $(\boldsymbol{x}_i,\boldsymbol{x}_j)$. Такой способ исследования дает оптимальный результат, только если можно выбрать пару $(\boldsymbol{x}_i,\boldsymbol{x}_j)$ с $d(\boldsymbol{x}_i,\boldsymbol{x}_j)=\omega n$ и малым $\omega>0$. В рассматриваемой постановке задачи этого сделать нельзя.

Для того чтобы усилить теорему 1, необходимо рассмотреть в (26) экспоненциальное число M кодовых слов $\{\boldsymbol{x}_i\}$, что значительно труднее (см. [11–13]). Усилим теорему 1 при условии, что в формуле (26) можно применить аддитивную аппроксимацию.

Предположим, что при $n\to\infty$ для всех $\{\boldsymbol{x}_i\}$ в формуле (26) справедливо аддитивное приближение

$$\mathbf{P}\left\{\bigcup_{k\neq i} SL_{\boldsymbol{x}_k}(p_0,\delta)\,\middle|\,p_1,\boldsymbol{x}_i\right\}=2^{o(n)}\sum_{k\neq i}\mathbf{P}\left\{SL_{\boldsymbol{x}_k}(p_0,\delta)\,\middle|\,p_1,\boldsymbol{x}_i\right\}. \tag{63}$$

Тогда (см. (36)) при $d(\boldsymbol{x}_i,\boldsymbol{x}_k)=\omega_{ik}n$

$$\mathbf{P}\left\{\bigcup_{k\neq i} SL_{\boldsymbol{x}_k}(p_0,\delta)\,\middle|\,p_1,\boldsymbol{x}_i\right\}=2^{o(n)}\sum_{k\neq i}2^{f(p_0,p_1,\omega_{ik})n}$$

и

$$\sum_{i=1}^{M}\mathbf{P}\left\{\bigcup_{k\neq i} SL_{\boldsymbol{x}_k}(p_0,\delta)\,\middle|\,p_1,\boldsymbol{x}_i\right\}=2^{o(n)}\sum_{i=1}^{M}\sum_{k\neq i}2^{f(p_0,p_1,\omega_{ik})n}. \tag{64}$$

Для того чтобы далее развить соотношение (64), введем некоторые дополнительные понятия. *Спектром (распределением расстояний)* $B(\mathcal{C})=(B_0,B_1,\dots,B_n)$ кода $\mathcal{C}$ длины n называется $(n+1)$-вектор с компонентами

$$B_i=|\mathcal{C}|^{-1}\,|\{(\boldsymbol{x},\boldsymbol{y}):\,\boldsymbol{x},\boldsymbol{y}\in\mathcal{C},\,d(\boldsymbol{x},\boldsymbol{y})=i\}|,\quad i=0,1,\dots,n. \tag{65}$$

Иными словами, B_i равно среднему числу кодовых слов $\boldsymbol{y}$ на расстоянии i от кодового слова $\boldsymbol{x}$. Общее число упорядоченных кодовых пар $(\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{C}$ с $d(\boldsymbol{x}, \boldsymbol{y}) = i$ равно $|\mathcal{C}|B_i$. Обозначим также $B_{\omega n} = 2^{b(\omega,r)n}$.

Тогда формулу (64) можно продолжить следующим образом:

$$\sum_{i=1}^{M} \mathbf{P}\left\{ \bigcup_{k \neq i} SL_{\boldsymbol{x}_k}(p_0, \delta) \,\middle|\, p_1, \boldsymbol{x}_i \right\} = 2^{o(n)} M \sum_{\omega > 0} 2^{[b(\omega,r) + f(p_0,p_1,\omega)]n}.$$

Поэтому (см. (36), (37))

$$\frac{1}{n} \log \left[\sum_{i=1}^{M} \mathbf{P}\left\{ \bigcup_{k \neq i} SL_{\boldsymbol{x}_k}(p_0, \delta) \,\middle|\, p_1, \boldsymbol{x}_i \right\} \right] =$$
$$= r + \max_{\omega, t} \{ b(\omega, r) + f(p_0, p_1, \omega, t) \} + o(1), \tag{66}$$

где $f(p_0, p_1, \omega, t)$ определено в (37). Тогда для функции $F(p_0, p_1, r)$ из (28) и (66) имеем

$$F(p_0, p_1, r) = \max_{\omega, t} \left\{ b(\omega, r) + g(p_0, t, \omega) + (p_0 - t) \log \frac{1 - p_1}{p_1} - h(p_0) \right\}. \tag{67}$$

В качестве оценки для $b(\omega, r)$ в (67) используем какую-либо функцию $b_{\mathrm{low}}(\omega, r)$ со следующим свойством: существует величина $\omega_{\max} = \omega_{\max}(r) > 0$, такая что

$$\max_{0 < \omega \leqslant \omega_{\max}} [b(\omega, r) - b_{\mathrm{low}}(\omega, r)] \geqslant 0, \quad r > 0. \tag{68}$$

Тогда для того чтобы выполнялось неравенство $F(p_0, p_1, r) > 0$ (см. (27)), достаточно, чтобы было справедливо условие (см. (37) и (67))

$$\min_{0 < \omega \leqslant \omega_{\max}} \max_{t > 0} \left\{ b_{\mathrm{low}}(\omega, r) + g(p_0, t, \omega) + (p_0 - t) \log \frac{1 - p_1}{p_1} - h(p_0) \right\} > 0. \tag{69}$$

Используем в (69) в качестве $b_{\mathrm{low}}(\omega, r)$ наилучшую из известных таких функций $\mu(r, \alpha, \omega)$, $h_2(\tau) = h_2(\alpha) - 1 + r$, с произвольным $\alpha \in [\delta_{GV}(r), 1/2]$ (см. (81), (82) и теорему 2 в Приложении). Для функции $\mu(r, \alpha, \omega)$ выполняется условие (68), она монотонно возрастает по r, и $\omega_{\max} = G(\alpha, \tau)$, где $G(\alpha, \tau)$ определено в (79). Тогда для того чтобы выполнялось неравенство (69), достаточно, чтобы было справедливо условие

$$\min_{0 < \omega \leqslant \omega_{\max}} \max_{t > 0} K(p_0, p_1, r, \omega, t) > 0, \tag{70}$$

где

$$K(p_0, p_1, r, \omega, t) = \mu(r, p_0, \omega) + g(p_0, t, \omega) + (p_0 - t) \log \frac{1 - p_1}{p_1} - h(p_0). \tag{71}$$

Заметим, что $K(p_0, p_1, r, 0, p_0) = 0$. Чтобы избежать громоздких вычислений, положим $t = p_0$. Функция $K(p_0, p_1, r, \omega, p_0) = 0$ вогнута по ω, т.е. $K''(p_0, p_1, r, \omega, p_0)_{\omega\omega} < 0$ (проще всего это проверить с помощью Maple). Поэтому минимум по ω достигается при $\omega = \omega_{\max} = G(\alpha, \tau)$, и условие (70) достаточно проверить для $\omega = G(\alpha, \tau)$. Известна полезная формула [11, лемма 4]

$$\mu(r, \alpha, G(\alpha, \tau)) = h_2(G(\alpha, \tau)) + r - 1, \quad h_2(\alpha) - h_2(\tau) = 1 - r. \tag{72}$$

Далее рассмотрим только более простой

Случай $p_1 < p_0 \leqslant 1/2$. Положим $r = r_0 = 1 - h(p_0)$ и $\alpha = p_0$ (заметим, что тогда $\delta_{GV}(r_0) = p_0$, $\tau = 0$). Тогда $G(\alpha, \tau) = 2p_0(1-p_0)$, и условие (70) достаточно проверить для $\omega = 2p_0(1-p_0)$. Из (71), (72) при $\alpha = p_0$, $\tau = 0$, $r = r_0 = 1 - h(p_0)$, $t = p_0$ и $\omega_{\max} = G(\alpha, \tau) = 2p_0(1-p_0)$ имеем

$$K(p_0, p_1, 1 - h(p_0), \omega_{\max}, p_0) = h_2(\omega_{\max}) + g(p_0, p_0, \omega_{\max}) - 2h(p_0),$$

где

$$g(p, p, 2p(1-p)) = 2p(1-p) + [1 - 2p(1-p)]h\left[\frac{p^2}{1-2p(1-p)}\right].$$

Можно проверить, что при $\omega_0 = 2p_0(1-p_0)$ справедливо равенство

$$K(p_0, p_1, 1 - h(p_0), \omega_0, p_0) = h_2(\omega_0) + \omega_0 + (1-\omega_0)h\left(\frac{p_0^2}{1-\omega_0}\right) - 2h(p_0) = 0. \quad (73)$$

Также имеем

$$[K(p_0, p_1, 1 - h(p_0), \omega_0, t)]'_t = \frac{1}{2}\log\frac{(1-t)^2 - (1-\omega_0-p_0)^2}{t^2 - (\omega_0 - p_0)^2} - \log\frac{1-p_1}{p_1}, \quad (74)$$

$$[K(p_0, p_1, 1 - h(p_0), \omega_0, t)]''_{tt} < 0.$$

Поэтому при $t = p_0$ имеем

$$[K(p_0, p_1, 1 - h(p_0), \omega_0, t)]'_{t=p_0} = \log\frac{1-p_0}{p_0} - \log\frac{1-p_1}{p_1} < 0, \quad p_1 < p_0, \quad (75)$$

Из (73)–(75) следует, что

$$K(p_0, p_1, 1 - h(p_0), \omega_0, t) > 0, \quad t < p_0.$$

Поэтому неравенство (70) выполняется для любых $r > r_0 = 1 - h(p_0)$ и $p_1 < p_0 \leqslant 1/2$.

В результате получаем следующий условный результат.

Предложение 4. *Если справедливо аддитивное приближение* (63), *то тогда* $r_{\text{crit}}(p_0, p_1) = 1 - h(p_0)$, $0 < p_1 < p_0 \leqslant 1/2$.

Замечание 7. Можно показать, что теорема 1 и формула (13) справедливы при любых $p_1 < p_0 \leqslant 1/2$. Для этого можно действовать аналогично [11], используя лемму 2 и рассматривая по отдельности случаи равенства в формуле (50) (по существу, это эквивалентно рассмотренному в § 6 случаю) и неравенства в ней. Доказательство во втором случае оказывается неоправданно громоздким (и ориентированным только на двоичный канал ДСК(p)). По этой причине мы его не приводим. Определенно, есть более простой способ доказательства.

ПРИЛОЖЕНИЕ

1. Функция $g(t, p, \omega)$ и формула (35). Рассмотрим кодовые слова $\boldsymbol{x} = \mathbf{0}$ и $\boldsymbol{x}_1$ с $d(\boldsymbol{x}, \boldsymbol{x}_1) = w(\boldsymbol{x}_1) = \omega n$, а также множество $S_{\boldsymbol{x}, \boldsymbol{x}_1}(t, p, \omega)$ из (34). Можно считать, что $\boldsymbol{x}_1 = (1, \ldots, 1, 0, \ldots, 0)$, причем $\boldsymbol{x}_1$ имеет сначала ωn "единиц", а затем $(1-\omega)n$ "нулей". Пусть также $\boldsymbol{u} \in S_{\boldsymbol{x}, \boldsymbol{x}_1}(t, p, \omega)$ имеет $u_1 n$ "единиц" на первых ωn позициях и $u_2 n$ "единиц" на следующих $(1-\omega)n$ позициях. Так как $u_1 + u_2 = t$, $\omega - u_1 + u_2 = p$, то

$$u_1 = \frac{t - p + \omega}{2}, \quad u_2 = \frac{t + p - \omega}{2}, \quad (76)$$

и при $n \to \infty$ получаем

$$\frac{1}{n}\log|S_{\boldsymbol{x},\boldsymbol{x}_1}(t,p,\omega)| = \frac{1}{n}\log\left[\binom{\omega n}{u_1 n}\binom{(1-\omega)n}{u_2 n}\right] =$$
$$= \omega h\Big(\frac{u_1}{\omega}\Big) + (1-\omega)h\Big(\frac{u_2}{1-\omega}\Big) + o(1) = g(t,p,\omega) + o(1), \tag{77}$$

где

$$g(t,p,\omega) = \omega h\Big(\frac{t+\omega-p}{2\omega}\Big) + (1-\omega)h\Big(\frac{t+p-\omega}{2(1-\omega)}\Big). \tag{78}$$

Также имеем

$$2g'_\omega(p,t,\omega) = -2\log\frac{1-\omega}{\omega} + \log\frac{(1-\omega)^2-(1-t-p)^2}{\omega^2-(t-p)^2},$$
$$2g'_t(p,t,\omega) = \log\frac{(1-t)^2-(1-\omega-p)^2}{t^2-(\omega-p)^2}, \quad g''_{tt}(p,t,\omega) < 0, \quad g''_{\omega\omega}(p,t,\omega) \leqslant 0.$$

Для корня ω_0 уравнения $g'_\omega(t,p,\omega) = 0$ имеем

$$\omega_0 = \frac{p-t}{1-2t}, \quad g(t,p,\omega_0) = h(t).$$

2. Функция $\boldsymbol{\mu(R,\alpha,\omega)}$. Введем функцию [14] ($0 \leqslant \tau \leqslant \alpha \leqslant 1/2$)

$$G(\alpha,\tau) = 2\frac{\alpha(1-\alpha)-\tau(1-\tau)}{1+2\sqrt{\tau(1-\tau)}} \geqslant 0. \tag{79}$$

Для α,τ, таких что $0 \leqslant \tau \leqslant \alpha \leqslant 1/2$ и $h_2(\alpha) - h_2(\tau) = 1-R$, введем функцию [16]

$$\mu(R,\alpha,\omega) = h_2(\alpha) - 2\int\limits_0^{\omega/2}\log\frac{P+\sqrt{P^2-4Qy^2}}{Q}\,dy - (1-\omega)h_2\Big(\frac{\alpha-\omega/2}{1-\omega}\Big), \tag{80}$$
$$P = \alpha(1-\alpha) - \tau(1-\tau) - y(1-2y), \quad Q = (\alpha-y)(1-\alpha-y).$$

Определим функцию $\delta_{GV}(R) \leqslant 1/2$ (граница Варшамова–Гилберта) как

$$1-R = h_2(\delta_{GV}(R)), \quad 0 \leqslant R \leqslant 1. \tag{81}$$

Важность функции $\mu(R,\alpha,\omega)$ и ее связь со спектром кода $\{B_i\}$ определяет следующий вариант теоремы 3 из [15].

Теорема 2 [15, теорема 3]. *Для любого (R,n)-кода и любого $\alpha \in [\delta_{GV}(R), 1/2]$ существуют $r_1(R,\alpha) > 0$ и ω, $0 < r_1(R,\alpha) \leqslant \omega \leqslant G(\alpha,\tau)$, где $h_2(\tau) = h_2(\alpha) - 1 + R$, а $G(\alpha,\tau)$ определено в* (79), *такие что*

$$n^{-1}\log B_{\omega n} \geqslant \mu(R,\alpha,\omega) + o(1), \quad n \to \infty. \tag{82}$$

Для $\mu(R,\alpha,\omega)$ из (80) *справедливо также неинтегральное представление* (83)–(85).

Замечание 8. Теорема 2 уточняет теорему 5 из [16] (см. также [12, теорема 2]). При $r_1 = 0$ теорема 2 переходит в теорему 5 из [16]. В [15, теорема 3] имеются оценки для $r_1(R,\alpha) > 0$.

Предложение 5 [11, предложение 3]. *Для функции* $\mu(R,\alpha,\omega)$ *справедливо представление*

$$\mu(R,\alpha,\omega)=(1-\omega)h_2\left(\frac{\alpha-\omega/2}{1-\omega}\right)-h_2(\alpha)+2h_2(\omega)+\omega\log\frac{2\omega}{e}-T(A,B,\omega), \quad (83)$$

где

$$\begin{aligned} T(A,B,\omega)&=\omega\log(v-1)-(1-\omega)\log\frac{v^2-A^2}{v^2-B^2}+\\ &+B\log\frac{v+B}{v-B}-A\log\frac{v+A}{v-A}-\frac{(v-1)(B^2-A^2)}{(v^2-B^2)\ln 2}, \quad (84)\\ v&=\frac{\sqrt{B^2\omega^2-2a_1\omega+a_1^2}+a_1}{\omega},\quad a_1=\frac{B^2-A^2}{2}, \end{aligned}$$

и

$$h_2(\alpha)-h_2(\tau)=1-R,\quad A=1-2\alpha,\quad B=1-2\tau,\quad 0\leqslant\tau\leqslant\alpha\leqslant 1/2. \quad (85)$$

Для любых $\alpha_0(R)\leqslant\alpha<1/2$ и $\omega>0$ имеем

$$\frac{d\mu(R,\alpha,\omega)}{d\alpha}>0,\quad \alpha_0(R)=h_2^{-1}(1-R).$$

Также для любых $\alpha>0$ и $R>0$ имеем $\mu(R,\alpha,0)=0$ и $\mu'_\omega(R,\alpha,\omega)|_{\omega=0}>0$. Кроме того, для любых $0\leqslant\tau\leqslant\alpha\leqslant 1/2$ и $0<\omega<G(\alpha,\tau)$

$$\mu''_{\omega^2}(R,\alpha,\omega)>0.$$

Для любого $\omega>0$ имеем $\mu(0,1/2,\omega)=0$.

Автор благодарит Ш. Ватанабе (Shun Watanabe) и рецензента за полезные обсуждения и конструктивные критические замечания, улучшившие статью.

СПИСОК ЛИТЕРАТУРЫ

1. *Бурнашев М.В., Амари Ш., Хан Т.С.* О некоторых задачах проверки гипотез с информационными ограничениями // Теория вероятн. и ее примен. 2000. Т. 45. № 4. С. 625–638.
2. *Бурнашев М.В., Хан Т.С., Амари Ш.* О некоторых задачах оценивания с информационными ограничениями // Теория вероятн. и ее примен. 2001. Т. 46. № 2. С. 233–246.
3. *Ahlswede R., Csiszár I.* Hypothesis Testing with Communication Constraints // IEEE Trans. Inform. Theory. 1986. V. 32. № 4. P. 533–542.
4. *Han T.S., Kobayashi K.* Exponential-type Error Probabilities for Multiterminal Hypothesis Testing // IEEE Trans. Inform. Theory. 1989. V. 35. № 1. P. 2–14.
5. *Ahlswede R., Burnashev M.V.* On Minimax Estimation in the Presence of Side Information about Remote Data // Ann. Statist. 1990. V. 18. № 1. P. 141–171.
6. *Han T.S., Amari S.* Statistical Inference under Multiterminal Data Compression // IEEE Trans. Inform. Theory. 1998. V. 44. № 6. P. 2300–2324.
7. *Shimokawa H., Han T.S., Amari S.* Error Bounds of Hypothesis Testing with Data Compression // Proc. 1994 IEEE Int. Sympos. on Information Theory (ISIT'94). Trondheim, Norway. June 27 – July 1, 1994. P. 114.
8. *Watanabe S.* Neyman–Pearson Test for Zero-Rate Multiterminal Hypothesis Testing // Proc. 2017 IEEE Int. Sympos. on Information Theory (ISIT'2017). Aachen, Germany. June 25–30, 2017. P. 116–120.

9. *Elias P.* Coding for Noisy Channels // IRE Conv. Rec. 1955. V. 4. P. 37–46. Reprinted in: Key Papers in the Development of Information Theory. New York: IEEE Press, 1974. P. 102–111.
10. *Gallager R.G.* Information Theory and Reliable Communication. New York: John Wiley & Sons, 1968.
11. *Бурнашев М.В.* О функции надежности ДСК: расширение области, где она известна в точности // Пробл. передачи информ. 2015. Т. 51. № 4. С. 3–22.
12. *Бурнашев М.В.* Спектр кода и функция надежности: двоичный симметричный канал // Пробл. передачи информ. 2006. Т. 42. № 4. С. 3–22.
13. *Бурнашев М.В.* Усиление оценки сверху для функции надежности двоичного симметричного канала // Пробл. передачи информ. 2005. Т. 41. № 4. С. 3–22.
14. *McEliece R.J., Rodemich E.R., Rumsey H., Jr., Welch L.R.* New Upper Bounds on the Rate of a Code via the Delsarte–MacWilliams Inequalities // IEEE Trans. Inform. Theory. 1977. V. 23. № 2. P. 157–166.
15. *Бурнашев М.В.* О границах снизу для спектра двоичного кода // Пробл. передачи информ. 2019. Т. 55. № 4. С. 76–85.
16. *Litsyn S.* New Bounds on Error Exponents // IEEE Trans. Inform. Theory. 1999. V. 45. № 2. P. 385–398.

Бурнашев Марат Валиевич
Институт проблем передачи информации
им. А.А. Харкевича РАН
burn@iitp.ru

Поступила в редакцию
10.04.2020
После доработки
15.05.2020
Принята к публикации
19.05.2020