ПРОБЛЕМЫ ПЕРЕДАЧИ ИНФОРМАЦИИ

Том 58 2022 Вып. 3

УДК 621.391 : 519.72

© 2022 г. И.В. Воробьев[1], В.С. Лебедев[2]

УЛУЧШЕНИЕ ВЕРХНИХ ГРАНИЦ СКОРОСТЕЙ РАЗДЕЛЯЮЩИХ И ПОЛНОСТЬЮ РАЗДЕЛЯЮЩИХ КОДОВ

Двоичный код называется (s,ℓ)-разделяющим кодом, если для любых двух непересекающихся наборов его слов мощности не более s и ℓ соответственно существует координата, в которой все слова из одного набора имеют символ 0, а все слова из другого набора имеют символ 1. Если же вдобавок для любых наборов существует вторая координата, в которой у первого набора во всех словах стоят 1, а у второго стоят 0, то такой код называется (s,ℓ)-полностью разделяющим кодом. В статье улучшаются верхние границы скоростей разделяющих и полностью разделяющих кодов.

Ключевые слова: разделяющие коды, полностью разделяющие коды, асимптотическая скорость, граница Плоткина.

DOI: 10.31857/S0555292322030044, **EDN:** EADNOC

§ 1. Введение

Впервые задача построения двоичных разделяющих систем возникла при исследовании асинхронных конечных автоматов. Для борьбы с критическими состояниями элементов памяти автомата при его переходе из одного устойчивого внутреннего состояния в другое Ю.Л. Сагаловичем было предложено использовать двоичные $(2,2)$-разделяющие коды [1]. В работе [2] было введено общее определение (s,ℓ)-разделяющих кодов. Отметим, что хотя изначально исследование разделяющих кодов было мотивировано задачами из теории автоматов, позднее они нашли применение при разработке способов защиты информации от нелегального копирования [3] и при построении хэш-функций [4].

Первая нижняя оценка скорости разделяющих кодов была получена с помощью случайного кодирования в работе [2]. В [5] с помощью случайного кодирования с выбрасыванием были получены нижние оценки скоростей $(2,2)$- и $(2,1)$-разделяющих и полностью разделяющих кодов, улучшающие оценки из [2] и совпадающие с оценкой скорости линейных $(2,2)$-разделяющих кодов из [6]. Отметим, что доказательство с выбрасыванием точно так же работает и для случая (s,ℓ)-разделяющих и полностью разделяющих кодов.

Также отметим неожиданное улучшение этих границ для $(2,1)$-разделяющих кодов [7].

Верхние границы для скорости $(2,2)$-разделяющих кодов впервые были получены Сагаловичем в [8] с помощью следующей идеи. Возьмем два кодовых слова на минимальном расстоянии (Хэмминга) d и ограничим исходный код на соответствующие d

[1] Исследование выполнено за счет гранта Российского научного фонда (номер проекта 22-41-02028).

[2] Работа выполнена при финансовой поддержке совместного проекта Российского фонда фундаментальных исследований и Национального научного фонда Болгарии (номер проекта 20-51-18002).

координат, предварительно удалив из кода эти два слова. Новый код длины d будет состоять из различных двоичных слов, и следовательно, мощность исходного кода не превосходит $2^d + 2$. Это позволяет оценить сверху скорость $(2,2)$-разделяющих кодов с помощью известных верхних оценок скорости кода (см. [9, 10]).

Идея, что этот подход можно обобщать на случай (s,ℓ)-кодов, высказывалась самим Ю.Л. Сагаловичем, а также Л.А. Бассалыго и Г.А. Кабатянским на семинарах ИППИ РАН по теории кодирования. Этот подход был реализован для хэш-кодов в [11], а для (s,ℓ)-свободных от перекрытий кодов, которые тесно связаны с разделяющими кодами, – в [12, 13].

В [14] были предложены рекуррентные верхние границы скоростей (s,ℓ)-разделяющих и полностью разделяющих кодов, выражающиеся через скорости кодов с параметрами $(s-1,\ell-1)$. Отметим, что в [14] (s,ℓ)-разделяющие коды сводились к $(s-1,\ell-1)$-полностью разделяющим кодам, что позволило получить более сильные оценки, чем если бы коды сводились к разделяющим.

В [15] получены рекуррентные неравенства, связывающие верхние границы для (s,ℓ)-разделяющих и полностью разделяющих кодов со скоростями кодов с параметрами $(s-u,\ell-v)$. Эти границы, в частности, позволили показать, что скорость t-IPP-кодов [16] экспоненциально мала по t (см. [17]).

В данной статье мы доказываем новые рекуррентные неравенства, которые связывают скорости разделяющих кодов и кодов, свободных от перекрытий. Эти неравенства позволяют улучшить верхние границы скоростей разделяющих кодов для многих параметров s и ℓ. Кроме того, мы улучшаем известную верхнюю границу для $(2,1)$-полностью разделяющих кодов. Так как наилучшие верхние границы выражаются через верхние границы скоростей кодов с меньшими параметрами, улучшение границы для $(2,1)$-полностью разделяющих кодов приводит к улучшению для широкого набора параметров.

§ 2. Определения и обозначения

Пусть N, M, s, ℓ – натуральные числа, символ $\triangleq$ обозначает равенство по определению, $|A|$ – мощность множества A, $[N] \triangleq \{1,2,\ldots,N\}$ – множество целых чисел от 1 до N. Произвольное подмножество булева куба $\{0,\ 1\}^N$ называется двоичным кодом длины N. Будем обозначать двоичную энтропию через

$$h(x) = -x\log_2(x) - (1-x)\log_2(1-x).$$

Определение 1 [2]. Двоичный код $\mathcal{C}$ называется (s,ℓ)-*разделяющим*, если для любых двух непересекающихся множеств кодовых слов $\mathcal{S}$ и $\mathcal{L}$, $|\mathcal{S}| \leqslant s$, $|\mathcal{L}| \leqslant \ell$, существует координата i, такая что:

$$\begin{aligned}&\text{либо}\quad x_i = 0 \text{ для любого } \boldsymbol{x} \in \mathcal{S} \quad \text{и} \quad y_i = 1 \text{ для любого } \boldsymbol{y} \in \mathcal{L},\\ &\text{либо}\quad x_i = 1 \text{ для любого } \boldsymbol{x} \in \mathcal{S} \quad \text{и} \quad y_i = 0 \text{ для любого } \boldsymbol{y} \in \mathcal{L}.\end{aligned}$$

Определение 2. Двоичный код $\mathcal{C}$ называется (s,ℓ)-*свободным от перекрытий*, если для любых двух непересекающихся множеств кодовых слов $\mathcal{S}$ и $\mathcal{L}$, $|\mathcal{S}| \leqslant s$, $|\mathcal{L}| \leqslant \ell$, существует координата i, для которой

$$x_i = 0 \text{ для любого } \boldsymbol{x} \in \mathcal{S} \quad \text{и} \quad y_i = 1 \text{ для любого } \boldsymbol{y} \in \mathcal{L}.$$

Определение 3. Двоичный код $\mathcal{C}$ называется (s,ℓ)-*полностью разделяющим*, если для любых двух непересекающихся множеств кодовых слов $\mathcal{S}$ и $\mathcal{L}$, $|\mathcal{S}| \leqslant s$, $|\mathcal{L}| \leqslant \ell$, существуют две координаты i и j, такие что:

$$\begin{aligned}&x_i = 0 \text{ для любого } \boldsymbol{x} \in \mathcal{S} \quad \text{и} \quad y_i = 1 \text{ для любого } \boldsymbol{y} \in \mathcal{L}, \quad \text{и}\\ &x_j = 1 \text{ для любого } \boldsymbol{x} \in \mathcal{S} \quad \text{и} \quad y_j = 0 \text{ для любого } \boldsymbol{y} \in \mathcal{L}.\end{aligned}$$

Отметим, что

определение 3 $\implies$ определение 2 $\implies$ определение 1,

а также что при $s = \ell$ определения (s,ℓ)-полностью разделяющих и (s,ℓ)-свободных от перекрытий кодов совпадают. Кроме того, напомним, что $(s,1)$-свободные от перекрытий коды известны как s-дизъюнктивные коды [18]. Полезно заметить, что если код $\mathcal{C}$ разделяющий, то и код $\mathcal{C}+\boldsymbol{a}$ тоже разделяющий для любого двоичного вектора $\boldsymbol{a}$. Если же код $\mathcal{C}$ полностью разделяющий, то и код $\mathcal{C}+\mathbf{1}$ обладает тем же свойством.

Принято думать, что (s,ℓ)-свободные от перекрытий коды были впервые определены в работе [19] в 1988 г., а полностью разделяющие системы – в работе [20] в 1973 г. На самом деле, в [20] были определены свободные от перекрытий коды (за 15 лет до работы [19]), но они были при этом названы полностью разделяющими.

Обозначим через $N_{\mathrm{s}}(M,s,\ell)$, $N_{\mathrm{cf}}(M,s,\ell)$ и $N_{\mathrm{cs}}(M,s,\ell)$ минимальную длину кодов из определений 1–3 при заданной мощности M. Определим асимптотические скорости кодов

$$R_{\mathrm{s}}(s,\ell) \triangleq \varlimsup_{M\to\infty} \frac{\log_2 M}{N_{\mathrm{s}}(M,s,\ell)}, \tag{1}$$

$$R_{\mathrm{cf}}(s,\ell) \triangleq \varlimsup_{M\to\infty} \frac{\log_2 M}{N_{\mathrm{cf}}(M,s,\ell)}, \tag{2}$$

$$R_{\mathrm{cs}}(s,\ell) \triangleq \varlimsup_{M\to\infty} \frac{\log_2 M}{N_{\mathrm{cs}}(M,s,\ell)}. \tag{3}$$

В этой статье мы улучшаем верхние границы скоростей R_{s} и R_{cs}. Заметим, что в силу симметрии определений 1 и 3 справедливы равенства $R_{\mathrm{s}}(s,\ell) = R_{\mathrm{s}}(\ell,s)$ и $R_{\mathrm{cs}}(s,\ell) = R_{\mathrm{cs}}(\ell,s)$.

§ 3. Известные результаты

Верхние границы скоростей разделяющих, полностью разделяющих и свободных от перекрытий кодов получаются из различных рекуррентных неравенств, связывающих скорости кодов для разных значений параметров s и ℓ.

Для $(s,1)$-свободных от перекрытий кодов, которые также называются s-дизъюнктивными кодами, известна [21] верхняя граница

$$R_{\mathrm{cf}}(s,1) \leqslant \overline{R}_{\mathrm{cf}}(s,1), \tag{4}$$

где последовательность $\overline{R}_{\mathrm{cf}}(s,1)$ определена следующим образом: $\overline{R}_{\mathrm{cf}}(1,1) \triangleq 1$, $\overline{R}_{\mathrm{cf}}(2,1) \triangleq \max\limits_{0<v<1} f_2(v)$, а $\overline{R}_{\mathrm{cf}}(s,1)$ при $s > 2$ является единственным решением уравнения

$$\overline{R}_{\mathrm{cf}}(s,1) = f_{\mathrm{s}}\left(1 - \frac{\overline{R}_{\mathrm{cf}}(s,1)}{\overline{R}_{\mathrm{cf}}(s-1,1)}\right),$$

где $f_{\mathrm{s}}(v) \triangleq h(v/s) - vh(1/s)$.

В [22] было доказано, что

$$R_{\mathrm{cf}}(s,\ell) \leqslant \left(\frac{1}{R_{\mathrm{cf}}(s,\ell-1)} + \frac{1}{R_{\mathrm{cf}}(s-1,\ell)}\right)^{-1}. \tag{5}$$

В [12] (см. также [23]) было доказано рекуррентное неравенство

$$R_{\mathrm{cf}}(s,\ell) \leqslant R_{\mathrm{cf}}(s-u,\ell-v)\frac{u^u v^v}{(u+v)^{u+v}}, \quad 1 \leqslant u \leqslant s-1, \quad 1 \leqslant v \leqslant \ell-1. \tag{6}$$

В [24] (см. также [13]) были доказаны более сильные неравенства

$$R_{\mathrm{cf}}(s,\ell) \leqslant \frac{R_{\mathrm{cf}}(s-u,\ell-v)}{R_{\mathrm{cf}}(s-u,\ell-v) + \dfrac{(u+v)^{u+v}}{u^u v^v}}, \tag{7}$$

$$R_{\mathrm{cf}}(s,\ell) \leqslant h\left(1/2 - \sqrt{\frac{2R_{\mathrm{cf}}(s,\ell)}{R_{\mathrm{cf}}(s-1,\ell-1)}\left(1 - \frac{2R_{\mathrm{cf}}(s,\ell)}{R_{\mathrm{cf}}(s-1,\ell-1)}\right)}\right). \tag{8}$$

В [14, 25] было доказано, что скорость $(s,1)$-разделяющих кодов удовлетворяет неравенству

$$R_{\mathrm{s}}(s,1) \leqslant \frac{1}{s}. \tag{9}$$

Для разделяющих и полностью разделяющих кодов в [14] были получены следующие результаты:

$$R_{\mathrm{s}}(s,\ell) \leqslant \widehat{R}\left(\frac{R_{\mathrm{s}}(s,\ell)}{R_{\mathrm{cs}}(s-1,\ell-1)}\right), \tag{10}$$

$$R_{\mathrm{cs}}(s,\ell) \leqslant \widehat{R}\left(\frac{2R_{\mathrm{cs}}(s,\ell)}{R_{\mathrm{cs}}(s-1,\ell-1)}\right), \tag{11}$$

где $\widehat{R}(\tau)$– произвольная верхняя асимптотическая оценка скорости кода с относительным расстоянием Хэмминга $\tau = d/n$.

В [15] для разделяющих кодов были доказаны рекуррентные неравенства, аналогичные неравенствам (6). А именно,

1. Для любых $u \in [s-1]$, $v \in [\ell-1]$

$$R_{\mathrm{s}}(s,\ell) \leqslant R_{\mathrm{s}}(s-u,\ell-v) \max_{0 \leqslant z \leqslant 1}\{z^u(1-z)^v + (1-z)^u z^v\}. \tag{12}$$

2. Для любого $v \in [\ell-1]$ и $u = v + s - \ell$, $1 \leqslant u \leqslant s-1$,

$$R_{\mathrm{s}}(s,\ell) \leqslant R_{\mathrm{cs}}(s-u,\ell-v) \max_{0 \leqslant z \leqslant 1}\{z^u(1-z)^v + (1-z)^u z^v\}. \tag{13}$$

3. Для любого $v \in [\min(s,\ell)-1]$

$$R_{\mathrm{cs}}(s,\ell) \leqslant R_{\mathrm{cs}}(s-v,\ell-v)\frac{1}{2^{2v}}. \tag{14}$$

Кроме того, было доказано неравенство

$$R_{\mathrm{s}}(s,\ell) \leqslant \min(R_{\mathrm{cf}}(s,\ell-1), R_{\mathrm{cf}}(s-1,\ell)). \tag{15}$$

§ 4. Новые неравенства

Мы начнем с улучшения верхней границы для скорости $(2,1)$-полностью разделяющих кодов. Отметим, что известная наилучшая верхняя оценка совпадала с верхней границей скорости $(2,1)$-свободных от перекрытий кодов и равнялась 0,321929.

Т е о р е м а 1. *Справедлива оценка*

$$R_{\rm cs}(2,1) = R_{\rm cs}(1,2) \leqslant h(0{,}25) - 0{,}5 = 0{,}311278\ldots \tag{16}$$

Д о к а з а т е л ь с т в о. Рассмотрим произвольный $(2,1)$-полностью разделяющий код $\mathcal{C}$ длины N и мощности M. Найдем вес w, такой что количество кодовых слов веса w максимально. Без ограничения общности можно считать, что $w \geqslant N/2$, так как в противном случае можно заменить код $\mathcal{C}$ на $\mathcal{C} + \mathbf{1}$.

Так как $(2,1)$-полностью разделяющий код является $(2,1)$-свободным от перекрытий, то можно применить известные оценки на количество слов фиксированного веса, доказанные в [21, 26]. Лемма 3 из [21] или теорема 1 из [26] позволяют оценить количество слов веса w как

$$4\frac{\binom{N}{\lfloor w/2\rfloor}}{\binom{2\lfloor w/2\rfloor}{\lfloor w/2\rfloor}}.$$

Тогда общее количество кодовых слов не превосходит

$$4N\frac{\binom{N}{\lfloor w/2\rfloor}}{\binom{2\lfloor w/2\rfloor}{\lfloor w/2\rfloor}},$$

что в силу хорошо известной асимптотики $\binom{N}{wN} = 2^{N(h(w)+o(1))}$ приводит к оценке

$$R_{\rm cs}(2,1) \leqslant \max_{0{,}5\leqslant w\leqslant 1}(h(w/2) - w) = h(1/4) - 1/2. \quad \blacktriangle$$

Т е о р е м а 2. *Пусть* $s, \ell \geqslant 2$. *Тогда*

$$R_{\rm cs}(s,\ell) \leqslant \frac{1}{2}\min\bigl(R_{\rm cf}(s,\ell-1), R_{\rm cf}(s-1,\ell)\bigr). \tag{17}$$

Д о к а з а т е л ь с т в о. Так как $R_{\rm s}(s,\ell) = R_{\rm s}(\ell,s)$ и $R_{\rm cf}(s,\ell) = R_{\rm cf}(\ell,s)$, то мы докажем только неравенство

$$R_{\rm cs}(s,\ell) \leqslant \frac{1}{2}R_{\rm cf}(s,\ell-1).$$

Рассмотрим произвольный (s,ℓ)-полностью разделяющий код $\mathcal{C}$ длины N и мощности M и некоторое его слово $\boldsymbol{c}$ веса w. Без ограничения общности можно считать, что $w \leqslant N/2$, ибо в противном случае мы рассмотрим код $\mathcal{C} + \mathbf{1}$. Построим новый код $\mathcal{C}'$, удалив все координаты, в которых выбранное слово $\boldsymbol{c}$ имеет нули. Удалим также и само слово $\boldsymbol{c}$. Проекция кода $\mathcal{C} \setminus \boldsymbol{c}$ на код $\mathcal{C}'$ инъективна. Действительно, пусть два слова $a \neq b$ из кода $\mathcal{C}\setminus\boldsymbol{c}$ при проекции на координаты, где вектор $\boldsymbol{c}$ имеет 1, совпали. Тогда в исходном коде нет координаты j, такой что $a_j = c_j = 1$ и $b_j = 0$, что противоречит тому, что исходный код был $(2,1)$-полностью разделяющим. Таким образом, длина кода $\mathcal{C}'$ равна $w \leqslant N/2$, а мощность равна $M-1$. Покажем, что код $\mathcal{C}'$ является $(s,\ell-1)$-свободным от перекрытий.

Действительно, рассмотрим произвольные непересекающиеся множества кодовых слов $\mathcal{S}$ и $\mathcal{L}$, $|\mathcal{S}| = s$, $|\mathcal{L}| = \ell - 1$, $\mathcal{S}, \mathcal{L} \subset \mathcal{C}'$. Этим множествам соответствуют множества $\widehat{\mathcal{S}}$ и $\widehat{\mathcal{L}}$ исходного кода $\mathcal{C}$. Так как код $\mathcal{C}$ является (s,ℓ)-полностью разделяющим, то для множеств $\widehat{\mathcal{S}}$ и $\widehat{\mathcal{L}} \cup \boldsymbol{c}$ найдется координата i, такая что

$$x_i = 0 \text{ для любого } \boldsymbol{x} \in \widehat{\mathcal{S}} \quad\text{и}\quad y_i = 1 \text{ для любого } \boldsymbol{y} \in \widehat{\mathcal{L}} \cup \boldsymbol{c}.$$

Так как $\boldsymbol{c}_i = 1$, то при построении кода $\mathcal{C}'$ эта координата не была удалена, и значит, в этой координате в коде $\mathcal{C}'$ выполнено

$$x_i = 0 \text{ для любого } \boldsymbol{x} \in \mathcal{S} \quad \text{и} \quad y_i = 1 \text{ для любого } \boldsymbol{y} \in \mathcal{L}.$$

А это и означает, что код $\mathcal{C}'$ является $(s, \ell - 1)$-свободным от перекрытий. Отсюда получаем искомое неравенство

$$R_{\text{cs}}(s, \ell) \leqslant \frac{1}{2} R_{\text{cf}}(s, \ell - 1). \quad \blacktriangle$$

Введем дополнительное определение. Для произвольных двух непересекающихся множеств кодовых слов $U, V \subset \mathcal{C}$ определим разделяющее "расстояние" $D(U, V)$ как количество координат, в которых все слова из одного множества имеют символ 0, а все слова из другого множества – символ 1. Будем говорить, что такие координаты разделяют множества U и V. Отметим, что $(1,1)$-разделяющее расстояние – это обычное расстояние Хэмминга, но при других параметрах u и v неравенство треугольника не выполняется.

Будем называть (u, v)-разделяющим расстоянием кода $\mathcal{C}$ величину

$$d_{uv}(\mathcal{C}) = \min_{\substack{U, V \subset \mathcal{C} \\ U \cap V = \varnothing \\ |U| = u,\ |V| = v}} D(U, V),$$

и будем называть ее просто разделяющим расстоянием $d(\mathcal{C})$ кода, когда параметры u и v ясны из контекста.

Определим величину

$$\Delta(u, v) = \max_{0 \leqslant z \leqslant 1} \{ z^u (1 - z)^v + (1 - z)^u z^v \}. \tag{18}$$

Следующее утверждение является обобщением классической границы Плоткина.

Лемма 1. *Для произвольного кода $\mathcal{C}$ длины N с (u, v)-разделяющим расстоянием d справедливо*

$$|\mathcal{C}| \leqslant \frac{u + v - 1}{1 - \left(\dfrac{N \Delta}{d} \right)^{1/(u+v-1)}},$$

если $d/N > \Delta = \Delta(u, v)$.

Доказательство. Рассмотрим код $\mathcal{C}$ длины N и мощности M и оценим сумму S всех (u, v)-разделяющих расстояний по всем парам непересекающихся кодовых подмножеств U, V таких, что $|U| = u, |V| = v$. Очевидно, что

$$S = \sum_{\substack{U, V \subset \mathcal{C} \\ |U| = u,\ |V| = v}} D(U, V) \geqslant \binom{M}{u + v} \binom{u + v}{u} d \geqslant \frac{M (M - u - v + 1)^{u+v-1}}{u!\, v!} d.$$

С другой стороны, если в i-й координате слов кода имеется A единиц (и $M - A$ нулей), то вклад этой координаты в S равен

$$\binom{A}{u} \binom{M - A}{v} + \binom{A}{v} \binom{M - A}{u},$$

и следовательно,

$$S \leqslant N \max_{0 \leqslant A \leqslant M} \left\{ \binom{A}{u} \binom{M - A}{v} + \binom{A}{v} \binom{M - A}{u} \right\} \leqslant N \frac{M^{u+v}}{u!\, v!} \Delta.$$

Объединяя верхнюю и нижнюю оценки величины S, получаем

$$(M-u-v+1)^{u+v-1}d \leqslant NM^{u+v-1}\Delta.$$

Отсюда следует

$$1-\frac{u+v-1}{M} \leqslant \left(\frac{N\Delta}{d}\right)^{1/(u+v-1)},$$

и так как $d/N > \Delta$ по условиям леммы, то

$$M \leqslant \frac{u+v-1}{1-\left(\frac{N\Delta}{d}\right)^{1/(u+v-1)}}. \quad \blacktriangle$$

Как уже отмечалось выше, $(1,1)$-разделяющее расстояние – это расстояние Хэмминга, $\Delta(1,1)=1/2$, и доказанная граница превращается в этом случае в классическую границу Плоткина: если $2d > N$, то $|\mathcal{C}| \leqslant \dfrac{2d}{2d-N}$.

Максимальную мощность кода длины N с (u,v)-разделяющим расстоянием d будем обозначать через $M_{u,v}(N,d)$. Следующая лемма является аналогом асимптотической границы Плоткина.

Л е м м а 2. *Для любого* τ, $0<\tau<\Delta=\Delta(u,v)$, *справедливо*

$$\log_2 M_{u,v}(N,\lfloor \tau N\rfloor) \leqslant N\left(1-\frac{\tau}{\Delta}+o(1)\right). \qquad (19)$$

Д о к а з а т е л ь с т в о. Сначала докажем неравенство

$$M_{u,v}(N,d) \leqslant 2M_{u,v}(N-1,d). \qquad (20)$$

Пусть $\mathcal{C}$ – код максимальной мощности $M_{u,v}(N,d)$ длины N с (u,v)-разделяющим расстоянием d. Выберем произвольную координату и без ограничения общности будем считать, что в словах кода $\mathcal{C}$ в ней больше нулей, чем единиц. Удалим эту координату и все кодовые слова, имеющие единицу в этой координате. У полученного кода $\mathcal{C}'$ его (u,v)-разделяющее расстояние не уменьшится, длина кода уменьшится на 1, а число слов уменьшится не более чем в два раза, что и доказывает неравенство (20).

Применяя неравенство (20) i раз, получаем

$$\log_2 M_{u,v}(N,d) \leqslant i+\log_2 M_{u,v}(N-i,d). \qquad (21)$$

Теперь возьмем минимальное целое i, такое что $\dfrac{d}{N-i} \geqslant \dfrac{\Delta}{1-\varepsilon}$ для некоторого $\varepsilon>0$, т.е. $i=N-\left\lfloor\dfrac{d(1-\varepsilon)}{\Delta}\right\rfloor$. Тогда из леммы 1 получаем

$$M_{u,v}(N-i,d) \leqslant \frac{u+v-1}{1-\left(\frac{(N-i)\Delta}{d}\right)^{1/(u+v)}} \leqslant \frac{u+v-1}{1-(1-\varepsilon)^{1/(u+v-1)}}.$$

Полагая $\varepsilon=1/N$, получаем, что

$$M(N-i,d) \leqslant c(u,v)N+o(N),$$

где $c(u,v)$ – некоторая константа, зависящая от u и v, но не зависящая от N. Теперь оценку (21) можно переписать в виде

$$\log_2 M_{u,v}(N,d) \leqslant N - \frac{d(1-\varepsilon)}{\Delta} + \log_2((c(u,v)+o(1))N) = \\ = N\left(1 - \frac{d}{N\Delta}\right) + o(N). \quad \blacktriangle \tag{22}$$

Замечание 1. Из леммы, в частности, следует, что не существует разделяющих кодов с положительной асимптотической скоростью и относительным разделяющим расстоянием $\geqslant \Delta$. Несложно доказать, что для любого $\tau < \Delta$ разделяющие коды с положительной асимптотической скоростью и относительным разделяющим расстоянием τ существуют, а значит, Δ является критической точкой. Доказательство этого факта мы приводим в Приложении.

Теперь мы готовы доказать основную теорему статьи. Доказательство основывается на обобщении упомянутой во введении идеи Ю.Л. Сагаловича на случай произвольных (s,ℓ)-разделяющих кодов и на доказанной выше асимптотической границе Плоткина для разделяющих кодов.

Теорема 3. *Для любых* $u \in [s-1]$, $v \in [\ell-1]$

$$R_{\text{s}}(s,\ell) \leqslant \frac{R_{\text{cf}}(s-u,\ell-v)\cdot}{R_{\text{cf}}(s-u,\ell-v)+\Delta^{-1}(u,v)}. \tag{23}$$

Доказательство. Докажем, что у любого (s,ℓ)-разделяющего кода $\mathcal{C}$ его минимальное (u,v)-разделяющее расстояние $d = d_{u,v}(C)$ достаточно велико, а именно

$$d_{u,v}(C) \geqslant N_{\text{cf}}(|C|-u-v, s-u, \ell-v). \tag{24}$$

Возьмем два непересекающихся множества кодовых слов $U, V \subset C$, $|U| = u$, $|V| = v$, на которых достигается минимальное (u,v)-разделяющее расстояние d кода $\mathcal{C}$. Обозначим через $I \subset [N]$ множество координат, разделяющих U и V. Как уже отмечалось, свойство разделимости инвариантно относительно сдвига кода на произвольный вектор, и следовательно, можно считать, что во всех координатах из множества I слова из U имеют нули, а слова из V – единицы. Рассмотрим код $\mathcal{C}'$ длины d, полученный из кода $\mathcal{C}\setminus\{U\cup V\}$ его укорочением (проекцией) на множество координат I. Покажем, что код $\mathcal{C}'$ является $(s-u,\ell-v)$-свободным от перекрытий кодом мощности $|C|-u-v$.

Действительно, рассмотрим произвольные непересекающиеся множества кодовых слов $\mathcal{S}', \mathcal{L}' \subset C'$, такие что $|\mathcal{S}'| = s-u$, $|\mathcal{L}'| = \ell - v$, и соответствующие им множества $\mathcal{S}$ и $\mathcal{L}$ исходного кода $\mathcal{C}$. Рассмотрим множества $\widehat{\mathcal{S}} = \mathcal{S}\cup U$ и $\widehat{\mathcal{L}} = \mathcal{L}\cup V$. Так как код $\mathcal{C}$ является (s,ℓ)-разделяющим, то найдется координата $i \in [N]$, которая разделяет множества $\widehat{\mathcal{S}}$ и $\widehat{\mathcal{L}}$, а значит, она разделяет и множества U и V, а следовательно, $i \in I$. Так как во всех координатах из множества I слова из U имеют нули, а слова из V – единицы, то код $\mathcal{C}'$ является $(s-u,\ell-v)$-свободным от перекрытий. Отсюда, в частности, следует, что при укорочении кода $\mathcal{C}\setminus\{U\cup V\}$ никакие два слова не могли совпасть, т.е. $|\mathcal{C}'| = |\mathcal{C}|-u-v$, и неравенство (24) доказано.

В силу неравенства (24) перепишем асимптотическую границу Плоткина (19) при $\tau < \Delta = \Delta(u,v)$ в виде

$$\log_2 M_{u,v}(N,\lfloor\tau N\rfloor) \leqslant N\left(1 - \frac{\tau}{\Delta} + o(1)\right) \leqslant \\ \leqslant N\left(1 - \frac{N_{\text{cf}}(M_{u,v}(N,\lfloor\tau N\rfloor)-u-v, s-u, \ell-v)}{N\Delta} + o(1)\right).$$

Таблица 1

Верхние границы скоростей разделяющих кодов

$s \diagdown \ell$	1	2	3	4	5	6	7	8
1	1,000000	0,500000	0,321929	0,199282	0,140457	0,105641	0,083000	0,067305
2	0,500000	0,283477	0,116879	0,066265	0,038684	0,024305	0,016336	0,011569
3	0,321929	0,116879	0,066265	0,028695	0,015326	0,008215	0,005271	0,003270
4	0,199282	0,066265	0,028695	0,015326	0,007088	0,003703	0,001912	0,001090
5	0,140457	0,038684	0,015326	0,007088	0,003703	0,001761	0,000912	0,000463
6	0,105641	0,024305	0,008215	0,003703	0,001761	0,000912	0,000439	0,000226
7	0,083000	0,016336	0,005271	0,001912	0,000912	0,000439	0,000226	0,000110
8	0,067305	0,011569	0,003270	0,001090	0,000463	0,000226	0,000110	0,000056

Таблица 2

Способ получения верхних границ скоростей кодов из табл. 1

$s \diagdown \ell$	1	2	3	4	5	6	7	8
1		Н9	Н15	Н15	Н15	Н15	Н15	Н15
2	Н9	Н10	Н10 + Т1	Н10 + Т2	Т3 $(1,3)$	Т3 $(1,3)$	Т3 $(1,4)$	Т3 $(1,4)$
3	Н15	Н10 + Т1	Н15	Н10 + Т1	Н15	Т3 $(1,3)$	Т3 $(2,5)$	Т3 $(2,5)$
4	Н15	Н10 + Т2	Н10 + Т1	Н15	Н10 + Т1	Н15	Т3 $(1,3)$	Т3 $(2,5)$
5	Н15	Т3 $(3,1)$	Н15	Н10 + Т1	Н15	Н10 + Т1	Н15	Т3 $(1,3)$
6	Н15	Т3 $(3,1)$	Т3 $(3,1)$	Н15	Н10 + Т1	Н15	Н10 + Т1	Н15
7	Н15	Т3 $(4,1)$	Т3 $(5,2)$	Т3 $(3,1)$	Н15	Н10 + Т1	Н15	Н10 + Т1
8	Н15	Т3 $(4,1)$	Т3 $(5,2)$	Т3 $(5,2)$	Т3 $(3,1)$	Н15	Н10 + Т1	Н15

Подставив туда очевидное соотношение $N_{\rm cf}(M,a,b) \geqslant \log_2 M(R_{\rm cf}+o(1))^{-1}$, получим, что

$$\log_2 M_{u,v}(N, \lfloor \tau N \rfloor) \leqslant N\left(1 - \frac{\log_2 M_{u,v}(N,\lfloor \tau N \rfloor)}{N\Delta(R_{\rm cf}+o(1))} + o(1)\right),$$

или, что равносильно,

$$\log_2 M_{u,v}(N, \lfloor \tau N \rfloor)\left(1 + \frac{1}{R_{\rm cf}\Delta} + o(1)\right) \leqslant N(1+o(1)). \quad \blacktriangle$$

Подчеркнем, что эта теорема ограничивает скорость разделяющих кодов через скорость свободных от перекрытий кодов, в отличие от неравенства (12), где в правой части неравенства присутствует скорость разделяющих кодов. В неравенстве (13) скорость разделяющих кодов ограничивается через скорость полностью разделяющих кодов, однако неравенство выполняется лишь для $u = v + s - \ell$, тогда как теорема 3 выполняется для всех u и v.

§ 5. Сравнения и таблицы

В этом параграфе мы приводим численные значения верхних границ асимптотической скорости разделяющих, полностью разделяющих и свободных от перекрытий кодов. Дополнительно мы предоставляем таблицы, где указано, с помощью какой именно теоремы (Т) или неравенства (Н) был получен тот или иной результат.

Результаты для разделяющих кодов приведены в табл. 1, 2. В табл. 2 в скобках указаны значения параметров u и v из теоремы 3, дающие наилучшие значения. Из таблиц видно, что новые теоремы позволяют улучшить верхние оценки скорости для многих значений параметров s и ℓ.

Таблица 3

Верхние границы скоростей полностью разделяющих кодов

$s\diagdown\ell$	1	2	3	4	5	6	7	8
1	1,000000	0,311278	0,160964	0,099641	0,070228	0,052820	0,041500	0,033652
2	0,311278	0,160964	0,064317	0,033133	0,021507	0,014338	0,010192	0,007299
3	0,160964	0,064317	0,033133	0,014895	0,007663	0,004861	0,003166	0,002115
4	0,099641	0,033133	0,014895	0,007663	0,003601	0,001852	0,001133	0,000719
5	0,070228	0,021507	0,007663	0,003601	0,001852	0,000887	0,000456	0,000265
6	0,052820	0,014338	0,004861	0,001852	0,000887	0,000456	0,000220	0,000113
7	0,041500	0,010192	0,003166	0,001133	0,000456	0,000220	0,000113	0,000055
8	0,033652	0,007299	0,002115	0,000719	0,000265	0,000113	0,000055	0,000028

Таблица 4

Способ получения верхних границ скоростей кодов из табл. 3

$s\diagdown\ell$	1	2	3	4	5	6	7	8
1		T1	T2	T2	T2	T2	T2	T2
2	T1	CF	H11 + T1	T2	T2	T2	T2	T2
3	T2	H11 + T1	CF	H11 + T1	T2	T2	T2	T2
4	T2	T2	H11 + T1	CF	H11 + T1	T2	T2	T2
5	T2	T2	T2	H11 + T1	CF	H11 + T1	T2	T2
6	T2	T2	T2	T2	H11 + T1	CF	H11 + T1	T2
7	T2	T2	T2	T2	T2	H11 + T1	CF	H11 + T1
8	T2	T2	T2	T2	T2	T2	H11 + T1	CF

Для полностью разделяющих кодов (табл. 3, 4) мы улучшаем все значения, кроме диагональных, где границы совпадают с границами свободных от перекрытий кодов.

Для свободных от перекрытий кодов (табл. 5, 6) мы не получаем новых результатов, все оценки получены с помощью ранее известных теорем. Но, как отмечалось ранее, верхние оценки скоростей свободных от перекрытий кодов используются для получения оценок скоростей разделяющих и полностью разделяющих кодов. Насколько нам известно, ранее такие таблицы, учитывающие все известные теоремы, не публиковались.

Как и в случае разделяющих кодов, в табл. 6 значения в скобках показывают параметры u и v, использующиеся для получения оптимальных границ с помощью неравенства (7).

ПРИЛОЖЕНИЕ: НИЖНЯЯ ГРАНИЦА СКОРОСТИ КОДОВ С РАЗДЕЛЯЮЩИМ РАССТОЯНИЕМ

Теорема 4. *Максимальная мощность $M_{u,v}(N,d)$ кода с (u,v)-разделяющим расстоянием $d = \lfloor \tau N \rfloor$ удовлетворяет неравенству*

$$\log_2 M_{u,v}(N,d) \geqslant N \frac{-h(\tau) - \tau \log_2 \Delta(u,v) - (1-\tau)\log_2(1-\Delta(u,v)) + o(1)}{u+v-1} \qquad (25)$$

при $0 \leqslant \tau < \Delta(u,v)$, где величина $\Delta(u,v)$ определена в (18).

Отметим, что правая часть положительна при $\tau < \Delta(u,v)$ и стремится к нулю при $\tau \to \Delta(u,v)$. При $u = v = 1$ точка $\Delta(1,1) = 1/2$, и наша нижняя граница превращается в границу Варшамова – Гильберта.

Таблица 5

Верхние границы скоростей свободных от перекрытий кодов

$s \diagdown \ell$	1	2	3	4	5	6	7	8
1	1,000000	0,321929	0,199282	0,140457	0,105641	0,083000	0,067305	0,055905
2	0,321929	0,160964	0,066265	0,043015	0,028677	0,020384	0,014598	0,011019
3	0,199282	0,066265	0,033133	0,015326	0,009722	0,006332	0,004230	0,003011
4	0,140457	0,043015	0,015326	0,007663	0,003703	0,002265	0,001438	0,000937
5	0,105641	0,028677	0,009722	0,003703	0,001852	0,000912	0,000529	0,000333
6	0,083000	0,020384	0,006332	0,002265	0,000912	0,000456	0,000226	0,000128
7	0,067305	0,014598	0,004230	0,001438	0,000529	0,000226	0,000113	0,000056
8	0,055905	0,011019	0,003011	0,000937	0,000333	0,000128	0,000056	0,000028

Таблица 6

Номер неравенства, из которого получены значения в табл. 5

$s \diagdown \ell$	1	2	3	4	5	6	7	8
1		Н4	Н4	Н4	Н4	Н4	Н4	Н4
2	Н4	Н5	Н8	Н8	Н7 $(1,2)$	Н7 $(1,2)$	Н7 $(1,3)$	Н7 $(1,3)$
3	Н4	Н8	Н5	Н8	Н7 $(1,2)$	Н7 $(1,2)$	Н7 $(1,2)$	Н7 $(1,2)$
4	Н4	Н8	Н8	Н5	Н8	Н7 $(1,2)$	Н7 $(1,2)$	Н7 $(1,2)$
5	Н4	Н7 $(2,1)$	Н7 $(2,1)$	Н8	Н5	Н8	Н7 $(2,3)$	Н7 $(3,5)$
6	Н4	Н7 $(2,1)$	Н7 $(2,1)$	Н7 $(2,1)$	Н8	Н5	Н8	Н7 $(2,3)$
7	Н4	Н7 $(3,1)$	Н7 $(2,1)$	Н7 $(2,1)$	Н7 $(3,2)$	Н8	Н5	Н8
8	Н4	Н7 $(3,1)$	Н7 $(2,1)$	Н7 $(2,1)$	Н7 $(5,3)$	Н7 $(3,2)$	Н8	Н5

Доказательство. Рассмотрим случайный код длины N и мощности M, где каждый элемент каждого кодового слова выбирается независимо и равен 1 с вероятностью p.

Вероятность того, что фиксированная координата разделяет два набора кодовых слов размера u и v равна

$$q = p^u(1-p)^v + p^v(1-p)^u.$$

Параметр p мы выберем так, чтобы максимизировать q. Отметим, что максимум вероятности разделения равен в точности величине $\Delta(u,v)$, определенной в формуле (18). Число ξ координат, разделяющих два фиксированных набора кодовых слов, имеет биномиальное распределение с параметрами N и $\Delta = \Delta(u,v)$. Оценим вероятность того, что $\xi < d$ в предположении $\Delta(u,v) > \tau$:

$$\mathcal{P} = \sum_{k=0}^{d-1} \binom{N}{k} \Delta^k (1-\Delta)^{N-k} \leqslant N 2^{N(h(\tau)+\tau\log_2\Delta+(1-\tau)\log_2(1-\Delta)+o(1))}.$$

Таким образом,

$$N^{-1}\log_2\mathcal{P} = h(\tau) + \tau\log_2\Delta + (1-\tau)\log_2(1-\Delta) + o(1).$$

Математическое ожидание количества наборов кодовых слов, разделяющее расстояние между которыми меньше d, не превосходит $\mathcal{P}\cdot M^{u+v}$. Стандартное рассуждение с выбрасыванием приводит к оценке

$$\log_2 M_{u,v}(N,d) \geqslant N\frac{-h(\tau) - \tau\log_2\Delta - (1-\tau)\log_2(1-\Delta) + o(1)}{u+v-1},$$

что и требовалось доказать. ▲

СПИСОК ЛИТЕРАТУРЫ

1. *Сагалович Ю.Л.* Метод повышения надежности конечного автомата // Пробл. передачи информ. 1965. Т. 1. № 2. С. 27–35. http://mi.mathnet.ru/ppi734
2. *Friedman A.D., Graham R.L., Ullman J.D.* Universal Single Transition Time Asynchronous State Assignments // IEEE Trans. Comput. 1969. V. 18. № 6. P. 541–547. https://doi.org/10.1109/T-C.1969.222707
3. *Barg A., Blakley G.R., Kabatiansky G.A.* Digital Fingerprinting Codes: Problem Statements, Constructions, Identification of Traitors // IEEE Trans. Inform. Theory. 2003. V. 49. № 4. P. 852–865. https://doi.org/10.1109/TIT.2003.809570
4. *Stinson D.R., Wei R., Chen K.* On Generalized Separating Hash Families // J. Combin. Theory Ser. A. 2008. V. 115. № 1. P. 105–120. https://doi.org/10.1016/j.jcta.2007.04.005
5. *Сагалович Ю.Л.* Полностью разделяющие системы // Пробл. передачи информ. 1982. Т. 18. № 2. С. 74–82. http://mi.mathnet.ru/ppi1227
6. *Пинскер М.С., Сагалович Ю.Л.* Нижняя граница мощности кода состояний автомата // Пробл. передачи информ. 1972. Т. 8. № 3. С. 58–66. http://mi.mathnet.ru/ppi854
7. *Randriambololona H.* $(2,1)$-Separating Systems beyond the Probabilistic Bound // Israel J. Math. 2013. V. 195. № 1. P. 171–186. https://doi.org/10.1007/s11856-012-0126-9
8. *Сагалович Ю.Л.* Верхняя граница мощности кода состояний автомата // Пробл. передачи информ. 1973. Т. 9. № 1. С. 73–83. http://mi.mathnet.ru/ppi884
9. *Körner J., Simonyi G.* Separating Partition Systems and Locally Different Sequences // SIAM J. Discrete Math. 1988. V. 1. № 3. P. 355–359. https://doi.org/10.1137/0401035
10. *Сагалович Ю.Л.* Новые верхние границы мощности разделяющих систем // Пробл. передачи информ. 1993. Т. 29. № 2. С. 109–111. http://mi.mathnet.ru/ppi182
11. *Bassalygo L.A., Burmester M., Dyachkov A., Kabatianskii G.* Hash Codes // Proc. 1997 IEEE Int. Sympos. on Information Theory (ISIT'97). Ulm, Germany. June 29 – July 4, 1997. P. 174. https://doi.org/10.1109/ISIT.1997.613089
12. *D'yachkov A.G., Vilenkin P.A., Yekhanin S.M.* Upper Bounds on the Rate of Superimposed (s,ℓ)-Codes Based on Engel's Inequality // Proc. 8th Int. Workshop on Algebraic and Combinatorial Coding Theory (ACCT-8). Tsarskoe Selo, Russia. Sept. 8–14, 2002. P. 95–99.
13. *Лебедев В.С.* Асимптотическая верхняя граница для скорости кодов, свободных от (w,r)-перекрытий // Пробл. передачи информ. 2003. Т. 39. № 4. С. 3–9. http://mi.mathnet.ru/ppi311
14. *Cohen G.D., Schaathun H.G.* Asymptotic Overview on Separating Codes // Tech. Rep. № 248. Dept. of Informatics, Univ. of Bergen. Bergen, Norway, 2003. Available at http://www.ii.uib.no/~georg/sci/inf/coding/hyperpdf/cs03rep.pdf.
15. *Воробьев И.В.* Границы скоростей разделяющих кодов // Пробл. передачи информ. 2017. Т. 53. № 1. С. 34–46. http://mi.mathnet.ru/ppi2225
16. *Hollmann H.D.L., van Lint J.H., Linnartz J.-P., Tolhuizen L.M.G.M.* On Codes with the Identifiable Parent Property // J. Combin. Theory Ser. A. 1998. V. 82. № 2. P. 121–133. https://doi.org/10.1006/jcta.1997.2851
17. *Кабатянский Г.А.* Идентифицирующие коды и их обобщения // Пробл. передачи информ. 2019. Т. 55. № 3. С. 93–105. https://doi.org/10.1134/S0555292319030070
18. *Kautz W., Singleton R.* Nonrandom Binary Superimposed Codes // IEEE Trans. Inform. Theory. 1964. V. 10. № 4. P. 363–377. https://doi.org/10.1109/TIT.1964.1053689
19. *Mitchell C.J., Piper F.C.* Key Storage in Secure Networks // Discrete Appl. Math. 1988. V. 21. № 3. P. 215–228. https://doi.org/10.1016/0166-218X(88)90068-6
20. *Magó G.* Monotone Functions in Sequential Circuits // IEEE Trans. Comput. 1973. V. 22. № 10. P. 928–933. https://doi.org/10.1109/T-C.1973.223620
21. *Дьячков А.Г., Рыков В.В.* Границы длины дизъюнктивных кодов // Пробл. передачи информ. 1982. Т. 18. № 3. С. 7–13. http://mi.mathnet.ru/ppi1232

22. *Stinson D.R., Wei R., Zhu L.* Some New Bounds for Cover-Free Families // J. Combin. Theory Ser. A. 2000. V. 90. № 1. P. 224–234. https://doi.org/10.1006/jcta.1999.3036
23. *D'yachkov A., Vilenkin P., Macula A., Torney D.* Families of Finite Sets in Which No Intersection of ℓ Sets Is Covered by the Union of s Others // J. Combin. Theory Ser. A. 2002. V. 99. № 2. P. 195–218. https://doi.org/10.1006/jcta.2002.3257
24. *Lebedev V.S.* Some Tables for (w, r)-Superimposed Codes // Proc. 8th Int. Workshop on Algebraic and Combinatorial Coding Theory (ACCT-8). Tsarskoe Selo, Russia. Sept. 8–14, 2002. P. 185–189.
25. *Blackburn S.R.* Frameproof Codes // SIAM J. Discrete Math. 2003. V. 16. № 3. P. 499–510. https://doi.org/10.1137/S0895480101384633
26. *Erdős P., Frankl P., Füredi Z.* Families of Finite Sets in Which No Set Is Covered by the Union of Two Others // J. Combin. Theory Ser. A. 1982. V. 33. № 2. P. 158–166. https://doi.org/10.1016/0097-3165(82)90004-8

Воробьев Илья Викторович
Сколковский институт науки и технологий (Сколтех), Москва
vorobyev.i.v@yandex.ru
Лебедев Владимир Сергеевич
Институт проблем передачи информации
им. А.А. Харкевича РАН, Москва
lebedev37@mail.ru

Поступила в редакцию
14.04.2022
После доработки
28.07.2022
Принята к публикации
30.07.2022