ФГБОУ ВО
Московский государственный университет
имени М. В. Ломоносова

На правах рукописи

ВОРОБЬЕВ Илья Викторович

Разделяющие коды

01.01.05 — теория вероятностей и математическая статистика

АВТОРЕФЕРАТ
диссертации на соискание ученой степени
кандидата физико–математических наук

Москва — 2016

Работа выполнена на кафедре теории вероятностей
механико–математического факультета ФГБОУ ВО
«Московский государственный университет имени М.В. Ломоносова»

Научный руководитель:
доктор физико-математических наук, профессор
Дьячков Аркадий Георгиевич

Официальные оппоненты:
доктор физико-математических наук, профессор
Кабатянский Григорий Анатольевич,
советник ректора по науке АНОО ВПО
“Сколковский институт науки и технологий”.

кандидат физико-математических наук
Жуковский Максим Евгеньевич,
доцент кафедры дискретной математики
факультета инноваций и высоких технологий
ФГАОУ ВПО "Московский физико-технический
институт (государственный университет)".

Ведущая организация:
ФГБУН Институт вычислительных технологий Сибирского отделения РАН

Защита диссертации состоится 7 апреля 2017 г. в 16^{45} на заседании диссертационного совета Д 501.001.85 на базе МГУ имени М.В. Ломоносова по адресу: 119234, Москва, ГСП–1, Ленинские горы, д. 1, МГУ имени М. В. Ломоносова, механико-математический факультет, аудитория 16–24.

С диссертацией можно ознакомиться в Фундаментальной библиотеке ФГБОУ ВО «Московский государственный университет имени М.В. Ломоносова» (Москва, Ломоносовский проспект, д. 27, сектор А, $8^{\text{й}}$ этаж), и на сайте http://mech.math.msu.ru/∼snark/index.cgi.

Автореферат разослан « » 2017 г.

Ученый секретарь диссертационного
совета Д 501.001.85 на базе МГУ,
доктор физико–математических наук,
профессор Власов
Виктор Валентинович

Актуальность темы и история вопроса

В настоящей диссертации рассматриваются задачи, лежащие на стыке теории вероятностей, теории информации и комбинаторной теории кодирования. Первая глава диссертации посвящена исследованию разделяющих кодов. Коду X мощности t и длины N сопоставим матрицу размера $N \times t$, столбцами которой служат кодовые слова кода X. Код X называется *разделяющим* (s, ℓ)-*кодом*, если в соответствующей ему матрице для любых двух непересекающихся множеств столбцов $\mathcal{S}$ и $\mathcal{L}$, $|\mathcal{S}| \leqslant s$, $|\mathcal{L}| \leqslant \ell$, существует такая строка (координата) i, что в ней множества символов из столбцов $\mathcal{S}$ и $\mathcal{L}$ не пересекаются. Будем говорить, что такая координата i *разделяет* множества слов $\mathcal{S}$ и $\mathcal{L}$.

В случае двоичных разделяющих кодов определение может быть переформулировано следующим образом. Код X называется *двоичным разделяющим* (s, ℓ)-*кодом*, если для любых двух непересекающихся множеств $\mathcal{S}$ и $\mathcal{L}$, $|\mathcal{S}| \leqslant s$, $|\mathcal{L}| \leqslant \ell$, существует такая координата i, для которой выполнено одно из следующих условий

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

или

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

Скоростью q-ичного кода X длины N и мощности t будем называть величину $R = \frac{\log_q t}{N}$. Определим *асимптотическую верхнюю и нижнюю скорости* разделяющих (s, ℓ)-кодов как

$$\overline{R}_s^{(q)}(s, \ell) = \varlimsup_{t \to \infty} \frac{\log_q t}{N^{(q)}(t, s, \ell)}, \qquad \underline{R}_s^{(q)}(s, \ell) = \varliminf_{t \to \infty} \frac{\log_q t}{N^{(q)}(t, s, \ell)}, \tag{1}$$

где число $N^{(q)}(t, s, \ell)$ равно минимальной длине q-ичного разделяющего (s, ℓ)-кода мощности t. Асимптотические скорости двоичных кодов будем обозначать просто $\overline{R}_s(s, \ell)$ и $\underline{R}_s(s, \ell)$.

Двоичные разделяющие (2, 2)-коды были впервые введены Ю.Л. Сагаловичем в 1965 году[1]. Начало исследованиям по разделяющим кодам положили проблемы противогоночного кодирования состояний дискретных автоматов. В 1969 году в работе А. Фридмана и др.[2] понятие двоичных разделяющих

[1] Сагалович Ю.Л. Метод повышения надежности конечного автомата // *Пробл. передачи информ.*, **1**:2 (1965), 27–35.

[2] Friedman A.D., Graham R.L., Ullman J.D. Universal single transition time asynchronous state assignments // *IEEE Trans. Comput.*, **18**:6 (1969), 541–547.

(2, 2)-кодов было обобщено до двоичных разделяющих (s, ℓ)-кодов. В этой же работе была получена первая нижняя граница

$$\underline{R}_s(s, \ell) \geqslant \frac{-\log_2(1 - 2^{1-s-\ell})}{s + \ell}. \tag{2}$$

В работе 1972-го года[3] М.С. Пинскер и Ю.Л. Сагалович стали рассматривать q-ичные разделяющие коды и доказали следующую нижнюю оценку асимптотической скорости линейных разделяющих (2, 2)-кодов

$$\underline{R}_s^{(q)} \geqslant \frac{-\log_q(1 - \beta)}{3}, \quad \beta = 1 - 4/q + 6/q^2 - 3/q^3. \tag{3}$$

Позднее Ю.Л. Сагаловичем[4] было показано, что эта граница верна не только для линейных, но и для произвольных кодов.

В статьях Ю.Л. Сагаловича[5] была доказана верхняя оценка асимптотической скорости

$$\overline{R}_s(2, 2) < 0.3045. \tag{4}$$

Эта оценка опирается на верхнюю границу $\overline{R}(\delta)$ скоростей кодов с фиксированным отношением расстояния и длины. Приведенное выше значение было получено с помощью границы Элайеса. Дж. Кернер и Г. Симоньи[6] повторили рассуждение Ю.Л. Сагаловича из указанных выше работ[5] и, воспользовавшись наилучшей известной границей для величины $\overline{R}(\delta)$ Р. Макэлиса и др.[7], получили оценку

$$\overline{R}_s(2, 2) < 0.2835, \tag{5}$$

которая не улучшена до сих пор.

Новая волна интереса к разделяющим кодам возникла благодаря статье Д. Боне и Д. Шоу 1998 года[8], посвященной задаче защиты авторских прав, где могут быть применены разделяющие $(s, 1)$-коды и (s, s)-коды.

Наилучшие границы скорости для многих значений параметров s и ℓ доказаны с помощью неравенств, связывающих скорости разделяющих, свободных от перекрытий и полностью разделяющих кодов.

---

[3]Пинскер М.С., Сагалович Ю.Л. Нижняя граница мощности кода состояний автомата // *Пробл. передачи информ.*, **8**:3 (1972), 59–66.

[4]Сагалович Ю.Л. Полностью разделяющие системы // *Пробл. передачи информ.*, **18**:2 (1982), 74–82.

[5]Сагалович Ю.Л. Верхняя граница мощности кода состояний автомата // *Пробл. передачи информ.*, **9**:1 (1973), 73–83. Сагалович Ю.Л. Кодирование состояний и надежность автоматов // *М.: Связь*, 1975.

[6]Korner J., Simonyi G. Separating Partition Systems and Locally Different Sequences // *SIAM J. Discrete Math.*, **1**:3 (1988), 355–359.

[7] McEliece R.J., Rodemich E.R., Rumsey H.C., Welch L.R. New Upper Bounds on the Rate Of Code Via Delsarte-MacWilliams Inequalities // *IEEE Trans. Inform. Theory*, **23**:2 (1977), 157–166.

[8]Boneh D., Shaw J. Collusion-secure fingerprinting for digital data // *IEEE Trans. Inform. Theory*, **44**:5 (1998), 1897–1905.

Код X называется *полностью разделяющим* (s,ℓ)*-кодом*, если для любых двух непересекающихся множеств $\mathcal{S}$ и $\mathcal{L}$, $|\mathcal{S}| \leqslant s$, $|\mathcal{L}| \leqslant \ell$, существуют две координаты i и j, для которых выполнены следующие условия

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

и

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

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

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

Полностью разделяющие коды впервые появились в 1973 году в работе Г. Маго[9]. Эти коды, как и разделяющие, используются в теории автоматов. Свободные от перекрытий коды были введены К. Митчеллом и Ф. Пайпером[10] в 1988 году в связи с криптографической задачей распределения ключей, описание которой можно также найти в статье В.С. Лебедева[11] или статье В.М. Сидельникова и О.Ю. Приходова[12]. Асимптотические скорости полностью разделяющих и свободных от перекрытий кодов определяются аналогично скорости разделяющих кодов

$$\overline{R}_{cs}(s,\ell) = \varlimsup_{t\to\infty} \frac{\log_2 t}{N_{cs}(t,s,\ell)}, \qquad \underline{R}_{cs}(s,\ell) = \varliminf_{t\to\infty} \frac{\log_2 t}{N_{cs}(t,s,\ell)},$$

$$\overline{R}_{cf}(s,\ell) = \varlimsup_{t\to\infty} \frac{\log_2 t}{N_{cf}(t,s,\ell)}, \qquad \underline{R}_{cf}(s,\ell) = \varliminf_{t\to\infty} \frac{\log_2 t}{N_{cf}(t,s,\ell)},$$

где числа $N_{cs}(t,s,\ell)$ и $N_{cf}(t,s,\ell)$ равны соответственно минимальным длинам полностью разделяющего и свободного от перекрытий кода мощности t.

Непосредственно из определений вытекают следующие неравенства

$$\begin{aligned} \overline{R}_{s}(s,\ell)/2 \leqslant \overline{R}_{cs}(s,\ell) \leqslant \overline{R}_{cf}(s,\ell) \leqslant \overline{R}_{s}(s,\ell), \\ \underline{R}_{s}(s,\ell)/2 \leqslant \underline{R}_{cs}(s,\ell) \leqslant \underline{R}_{cf}(s,\ell) \leqslant \underline{R}_{s}(s,\ell). \end{aligned} \tag{6}$$

В работе Г. Коэна и Г. Шаатуна 2003 года[13] представлены неравенства

$$\overline{R}_{cs}(s,\ell) \leqslant \overline{R}\left(\frac{2\overline{R}_{cs}(s,\ell)}{\overline{R}_{cs}(s-1,\ell-1)}\right),$$

[9]Mago G. Monotone Functions in Sequential Circuits // *IEEE Trans. Comput.*, **22**:10 (1973), 928–933.

[10]Mitchell C.J., Piper F.C. Key Storage in Secure Networks // *Discrete Appl. Math.*, **21**:3 (1988), 215–228.

[11]Лебедев В.С. Асимптотическая верхняя граница для скорости кодов, свободных от (w,r)-перекрытий // *Пробл. передачи информ.*, **39**:4 (2003), 3–9.

[12]Сидельников В.М., Приходов О.Ю. О построении кодов, свободных от (w,r)-перекрытий // *Пробл. передачи информ.*, **45**:1 (2009), 36–40.

[13]Cohen G.D., Schaathun H.G. Asymptotic overview on separating codes // *Tech. Report 248*, Department of Informatics, University of Bergen, Bergen, Norway, 2003.

$$\overline{R}_s(s,\ell) \leqslant \overline{R}\left(\frac{\overline{R}_s(s,\ell)}{\overline{R}_{cs}(s-1,\ell-1)}\right), \tag{7}$$

где $\overline{R}(\delta)$ — оценка асимптотической скорости произвольного кода с фиксированным отношением кодового расстояния к длине кода из уже названной работы Р. Макэлиса и др. Рекурсивное применение указанных неравенств позволило авторам получить верхние оценки скоростей двоичных разделяющих и полностью разделяющих кодов для малых значений параметров s и ℓ; многие из этих оценок остаются наилучшими до сих пор.

Асимптотическое поведение величин $\overline{R}_s(s,\ell)$ и $\underline{R}_s(s,\ell)$ при $s \to \infty$ и фиксированном ℓ следует из границ для свободных от перекрытий кодов и неравенства (2)

$$\frac{\ell^\ell \log_2 e}{e^\ell}\frac{1}{s^{\ell+1}}(1+o(1)) \leqslant \underline{R}_{cf}(s,\ell),$$

$$\overline{R}_{cf}(s,\ell) \leqslant \frac{(\ell+1)^{\ell+1}}{2e^{\ell-1}}\frac{\log_2 s}{s^{\ell+1}}(1+o(1)), \quad s \to \infty. \tag{8}$$

Верхняя оценка была доказана А.Г. Дьячковым и др.[14] с помощью рекуррентных неравенств, установленных ранее В.С. Лебедевым

$$\overline{R}_{cf}(s,\ell) \leqslant \frac{\overline{R}_{cf}(s-i,\ell-j)}{\overline{R}_{cf}(s-i,\ell-j)+\frac{(i+j)^{i+j}}{i^i\cdot j^j}}, \quad i \in [s-1],\, j \in [\ell-1].$$

Нижняя оценка была получена Н. Куонгом и Т. Зейселем[15] в 1988 году с помощью метода случайного кодирования на некотором специальном ансамбле с независимыми двоичными равновесными словами.

Множество результатов было получено специально для наиболее важных для задачи защиты авторских прав разделяющих $(s,1)$-кодов и (s,s)-кодов, причем зачастую авторов интересовали результаты для конечных длин, а не асимптотическая скорость. В работе Дж. Стаддон и др.[16] были установлены границы

$$t \leqslant s(q^{\left\lceil \frac{N(t,s,1)}{s} \right\rceil} - 1),$$

$$t \leqslant q^{\left\lceil \frac{N(t,s,s)}{s} \right\rceil} + 2s - 2,$$

откуда следуют оценки для асимптотической скорости

$$\overline{R}_s^{(q)}(s,1) \leqslant \frac{1}{s}, \quad \overline{R}_s^{(q)}(s,s) \leqslant \frac{1}{s}.$$

[14] 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. of ACCT-8*, Tsarskoe Selo, (2002), 95–99.

[15] Nguyen Quang A., Zeisel T. Bounds on Constant Weight Binary Superimposed Codes // *Problems of Control and Inform. Theory.*, **17**:4 (1988), 223–230.

[16] Staddon J.N., Stinson D.R., Wei R. Combinatorial properties of frameproof and traceability codes // *IEEE Trans. Inform. Theory*, **47**:3 (2001), 1042–1049.

В работе Д. Стинсона и др.[17] было показано, что

$$t \leqslant 4q^{\left\lceil \frac{N(t,2,2)}{3} \right\rceil} - 3,$$

а позднее Д. Стинсоном и Г. Заверухой[18] этот результат был обобщен до

$$t \leqslant q^{\left\lceil \frac{N(t,s,s)}{2s-1} \right\rceil} + (s-1)(2s-1)(q^{\left\lceil \frac{N(t,s,s)}{2s-1} \right\rceil} - 1).$$

Для общего случая разделяющих (s,ℓ)-кодов Д. Стинсоном и Г. Заверухой[19] была доказана граница

$$t \leqslant (2s\ell - w)q^{\left\lceil \frac{N(t,s,\ell)}{s+\ell-1} \right\rceil} - 2s\ell + s + 1,$$

что дает следующую оценку для асимптотической скорости

$$\overline{R}_s^{(q)}(s,\ell) \leqslant \frac{1}{s+\ell-1}.$$

Вместе с достаточно очевидной случайной нижней границей

$$\lim_{q\to\infty} \underline{R}_s^{(q)}(s,\ell) \geqslant \frac{1}{s+\ell-1}$$

это дает нам равенство

$$\lim_{q\to\infty} \underline{R}_s^{(q)}(s,\ell) = \lim_{q\to\infty} \overline{R}_s^{(q)}(s,\ell) = \frac{1}{s+\ell-1}.$$

В недавней работе Ч. Шэнгуэн и др.[20] построили новую границу для разделяющих $(s,1)$-кодов

$$t \leqslant q^{\left\lceil \frac{N(t,s,1)(q-1)}{s(s-1)/2} \right\rceil \log_q(eqs(s-1)/(2(q-1)))} + s,$$

которая дает следующую оценку для асимптотической скорости

$$\overline{R}_s^{(q)}(s,1) \leqslant \frac{4(q-1)\log_q s}{s^2}(1+o(1)), \qquad s\to\infty.$$

В этой же работе была доказана нижняя граница

$$\underline{R}_s^{(q)}(s,1) \geqslant \frac{q-1}{s^2 e \ln q}(1+o(1)), \qquad s\to\infty.$$

---

[17] Stinson D. R., Wei R., Chen K. On generalized separating hash families // *Journal of Combinatorial Theory, Series A*, **115**:1 (2008), 105–120.

[18] Stinson D.R., Zaverucha G.M. New bounds for generalized separating hash families // *Technical Report 2007-21, Center for Applied Cryptographic Research*, University of Waterloo, 2007.

[19] Stinson D.R., Zaverucha G.M. Some improved bounds for secure frameproof codes and related separating hash families // *IEEE Transactions on Information Theory*, **54**:6 (2008), 2508–2514.

[20] Shangguan C., Wang X., Ge G., Miao Y. New Bounds For Frameproof Codes // *arXiv preprint arXiv:1411.5782*, 2014.

В следующих двух главах настоящей диссертации исследуются два обобщения классических дизъюнктивных кодов. Дизъюнктивные коды являются частным случаем свободных от перекрытий кодов при $\ell = 1$. Дизъюнктивной суммой множества двоичных слов называется двоичное слово, у которого в i-ой позиции стоит 0 в том и только в том случае, если у всех складываемых слов в этой позиции стоит 0. Говорят, что слово $\mathbf{u}$ покрывает слово $\mathbf{v}$, если их дизъюнктивная сумма совпадает со словом $\mathbf{u}$. Двоичный код X называется *дизъюнктивным кодом силы* s, если произвольная дизъюнктивная сумма s слов не покрывает никакое другое слово кода. Ключевое свойство дизъюнктивных кодов заключается в том, что, имея дизъюнктивную сумму не более чем s слов, мы можем эти слова распознать.

Дизъюнктивные коды были введены У. Каутсом и Р. Синглтоном в 1964 году[21]. Эти коды имеют множество приложений, обзор которых можно найти в книге Д. Ду и Ф. Хвонга[22] или монографии Ф. Чикалеза[23]. Наиболее известное приложение – это применение дизъюнктивных кодов в задаче группового тестирования. Асимптотические нижнюю и верхнюю скорости дизъюнктивных кодов $\underline{R}_{cf}(s,1)$ и $\overline{R}_{cf}(s,1)$ будем для краткости обозначать просто $\underline{R}_{cf}(s)$ и $\overline{R}_{cf}(s)$. Из работ А.Г. Дьячкова с В.В. Рыковым[24] и с А.М. Рашадом[25] известно, что выполняются неравенства

$$\frac{\ln 2}{s^2}(1+o(1)) \leqslant \underline{R}_{cf}(s) \leqslant \overline{R}_{cf}(s) \leqslant \frac{2\log_2 s}{s^2}(1+o(1)), \quad s \to \infty.$$

Вторая глава диссертации посвящена дизъюнктивным кодам со списочным декодированием. Двоичный код называется *дизъюнктивным кодом со списочным декодированием силы s и размером списка L* (СД s_L-кодом), если дизъюнктивная сумма любых s кодовых слов покрывает не более $L-1$ других кодовых слов. СД s_L-коды были введены А.Г. Дьячковым и В.В. Рыковым в статье 1981 года[26], где рассматривалось использование таких кодов при передаче информации через канал множественного доступа в системе связи

[21]Kautz W.H., Singleton R.C. Nonrandom Binary Superimposed Codes // *IEEE Trans. Inform. Theory.*, **10**:4 (1964), 363–377.

[22]Du D.Z., Hwang F.K. Combinatorial Group Testing and Its Applications, 2nd ed. // *Series on Applied Mathematics*, **12**, 2000.

[23]Cicalese F. Fault-Tolerant Search Algorithms // *Monographs in Theoretical Computer Science–An EATCS Series*, Springer-Verlag, **15**, 2013.

[24]Дьячков А.Г., Рыков В.В. Границы длины дизъюнктивных кодов // *Пробл. передачи информ.*, **18**:3 (1982), 7–13.

[25]D'yachkov A.G., Rashad A.M. Universal Decoding for Random Design of Screening Experiments // *Microelectronics and Reliability*, **29**:6 (1989), 965–971.

[26]Дьячков А.Г., Рыков В.В. Применение кодов для канала с множественным доступом в системе связи АЛОХА // *Тр. VI Всесоюзн. школы-семинара по вычислительным сетям*, Москва - Винница, **4** (1981), 18–24.

АЛОХА. В работе П.А. Виленкина 1998 года[27] приведены некоторые конструкции СД s_L-кодов, а также рассмотрено их применение при построении двухступенчатых процедур групповых проверок.

Определим асимптотические верхнюю и нижнюю скорости СД s_L-кодов стандартным образом:

$$\overline{R}_L(s) = \overline{\lim_{t\to\infty}} \frac{\log_2 t}{N(t,s,L)}, \qquad \underline{R}_L(s) = \varliminf_{t\to\infty} \frac{\log_2 t}{N(t,s,L)},$$

где число $N(t,s,L)$ равно минимальной длине СД s_L-кода мощности t. В 1983 году А.Г. Дьячковым и В.В. Рыковым[28] были получены границы для асимптотической скорости $R_L(s)$, из которых вытекают неравенства

$$\overline{R}_L(s) \leqslant \frac{1}{s}, \qquad \text{при любых натуральных } s \text{ и } L,$$

$$\frac{L\log_2 e}{es^2}(1+o(1)) \leqslant \underline{R}_L(s) \leqslant \overline{R}_L(s) \leqslant \frac{2L^2\log_2 s}{s^2}(1+o(1)), \qquad \text{при } s\to\infty.$$

В работе 2005 года А. Де Бонис и др.[29] была улучшена верхняя граница скорости:

$$\overline{R}_L(s) \leqslant \frac{8L\log_2 s}{s^2}(1+o(1)), \qquad s\to\infty.$$

В третьей главе диссертации исследуются почти дизъюнктивные коды, которые являются вероятностным обобщением понятия дизъюнктивных кодов. Впервые они были определены Э. Макулой и др. в 2004-ом году[30]. Будем называть код X *почти дизъюнктивным* (s,ε)*-кодом*, если доля множеств из s кодовых слов, чья дизъюнктивная сумма покрывает хотя бы одно другое слово, не превосходит ε. Параметр ε может интерпретироваться как вероятность ошибки декодирования (восстановления s слагаемых, дающих исходную дизъюнктивную сумму).

В работе 2013-го года Л.А. Бассалыго и В.В. Рыкова[31] для некоторого семейства кодов была посчитана вероятность ошибки ε и доказано существование почти дизъюнктивных кодов с параметрами

$$\frac{\log_2 t}{N} = \frac{\ln 2}{s}(1+o(1)), \quad s^2 = \Theta(N), \quad \varepsilon\to 0, \quad N\to\infty. \tag{9}$$

[27]Vilenkin P.A. On Constructions of List-Decoding Superimposed Codes // *Proc. of ACCT-6*, Pskov, (1998), 228–231.

[28]D'yachkov A.G., Rykov V.V. A Survey of Superimposed Code Theory // *Problems of Control and Inform. Theory*, **12**:4 (1983), 229–242.

[29]De Bonis A., Gasieniec L., Vaccaro U. Optimal two-stage algorithms for group testing problems // *SIAM J. Comp.*, **34**:5 (2005), 1253–1270.

[30]Macula A.J., Rykov V.V., Yekhanin S., Trivial two-stage group testing for complexes using almost disjunct matrices // *Discrete Applied Mathematics*, **137**:1 (2004), 97–107.

[31]Бассалыго Л.А., Рыков В.В. Гиперканал множественного доступа // *Пробл. передачи информ.*, **49**:4 (2013), 3–12.

В данном исследовании нас интересует пропускная способность почти дизъюнктивных кодов при константном s и $N \to \infty$. *Пропускной способностью* $C(s)$ почти дизъюнктивных кодов будем называть точную верхнюю грань скорости кодов, для которых вероятность ошибки ε убывает экспоненциально с ростом длины кода.

Близким к понятию дизъюнктивных кодов является понятие дизъюнктивных планов поиска, которое накладывает несколько более слабые ограничения на код. Код называется *дизъюнктивным планом поиска силы* s, если нет двух разных множеств мощности s с одинаковой дизъюнктивной суммой. Дизъюнктивный план тоже позволяет восстанавливать множество слов по их дизъюнктивной сумме, но для этого нам может потребоваться перебирать все множества слов мощности s. Будем называть код X *почти дизъюнктивным* (s, ε)*-планом*, если доля множеств из s кодовых слов, чья дизъюнктивная сумма совпадает с дизъюнктивной суммой хотя бы одного другого множества мощности s, не превосходит ε. *Пропускной способностью* $C_p(s)$ почти дизъюнктивных s-планов будем называть точную верхнюю грань скорости кодов, для которых вероятность ошибки ε убывает экспоненциально с ростом длины кода. Пропускная величина $C_p(s)$ дизъюнктивных планов поиска была подсчитана М.Б. Малютовым и В.Л. Фрейдлиной[32]:

$$C_p(s) = \frac{1}{s}.$$

Последняя часть диссертации посвящена многоступенчатому поиску дефектов с помощью групповых проверок. Имеется два основных типа алгоритмов поиска: адаптивные и неадаптивные. В неадаптивных алгоритмах все тесты определены заранее и могут проводиться одновременно, а в адаптивных алгоритмах тесты проводятся последовательно, и при планировании следующего могут быть использованы результаты предыдущих. Многоступенчатые алгоритмы являются промежуточным вариантом между адаптивными и неадаптивными. В них выделяется несколько этапов (ступеней), внутри которых тесты могут проводиться параллельно, но при планировании следующего этапа учитываются результаты предыдущих. Верхней(нижней) асимптотической скоростью поиска $\overline{R}^{(p)}(s)(\underline{R}^{(p)}(s))$ мы будем называть верхний(нижний) предел отношения двоичного логарифма количества объектов к минимальному необходимому числу тестов для нахождения не более s дефектов за p ступеней. Очевидно, что при росте количества ступеней растет и скорость поиска, минимальную скорость имеют неадаптивные алгоритмы, а максимальную – адаптивные.

Теоретико-информационная граница дает верхнюю оценку $1/s$ на скорость поиска s дефектов для алгоритма с произвольным числом ступеней.

[32]Малютов М.Б., Фрейдлина В.Л. О применении теории информации к одной задаче выделения значимых факторов // *Теория вероятностей и ее применения*, **18**:2 (1973), 432—444.

Известно, что адаптивные алгоритмы достигают этой границы, а неадаптивные алгоритмы ее не достигают для $s = 2$ (см. работу Д. Копперсмита и Дж. Ширера[33]) и $s \geqslant 11$ (см. уже упоминавшуюся статью А.Г. Дьячкова и В.В. Рыкова 1982 года). Отметим работу П. Дамашке и др.[34], где предложен новый подход к многоступенчатым алгоритмам поиска, основанный на гиперграфах, и доказано существование двухступенчатого алгоритма поиска со скоростью 1/2.44, а также работу П. Дамашке и Ш. Мухаммада,[35] в которой был предложен явный алгоритм со скоростью 0.4.

Цель работы

Цель диссертационной работы — построение новых оценок асимптотических скоростей разделяющих и списочных дизъюнктивных кодов, пропускной способности дизъюнктивных кодов, а также исследование многоступенчатых алгоритмов поиска малого числа дефектов.

Научная новизна работы

Результаты работы являются новыми и заключаются в следующем:

1. Доказаны новые верхние и нижние границы для асимптотической скорости q-ичных разделяющих кодов.

2. Доказаны новые верхние границы асимптотической скорости списочных дизъюнктивных кодов.

3. Впервые получены границы снизу для пропускной способности $C(s)$ почти дизъюнктивных кодов.

4. Впервые построен алгоритм поиска двух дефектов с конечным числом ступеней, достигающий информационной границы.

5. Впервые явно построен алгоритм поиска произвольного числа дефектов с конечным числом ступеней и ненулевой скоростью.

[33]Coppersmith D., Shearer J. New Bounds for Union-free Families of Sets // *Journal of Combinatorics*, **5** (1998), 581–596.

[34]Damaschke P., Sheikh Muhammad A., Wiener G. Strict group testing and the set basis problem // *Journal of Combinatorial Theory, Series A*, **126** (2014), 70–91.

[35]Damaschke P., Sheikh Muhammad A. A toolbox for provably optimal multistage strict group testing strategies // *International Computing and Combinatorics Conference*, Springer Berlin Heidelberg, (2013), 446–457.

Основные методы исследования

В работе используются вероятностные методы, в частности метод случайного кодирования для ансамбля равновесных кодов и для ансамбля с независимыми компонентами, методы оценки вероятностей больших уклонений. Также используются аналитических методы и методы комбинаторной теории кодирования.

Практическая и теоретическая значимость работы

Результаты диссертационной работы носят теоретический характер. Они могут быть полезны специалистам, работающим в области теории информации и комбинаторной теории кодирования.

Апробация диссертации

Результаты исследования неоднократно докладывались автором на следующих научно-исследовательских семинарах:

1. Семинар по теории кодирования под рук. Л.А. Бассалыго в 2013–2016 гг., Институт проблем передачи информации им. А.А. Харкевича РАН.

2. Семинар «Проблемы современной теории информации» в 2013–2016 гг. под рук. А.Г. Дьячкова, кафедра теории вероятностей, механико-математический факультет, Московский государственный университет им. М.В. Ломоносова.

3. Семинар «Экстремальная комбинаторика и случайные структуры» под рук. Д.А. Шабанова в 2016 году, кафедра теории вероятностей, механико-математический факультет, Московский государственный университет им. М.В. Ломоносова.

4. Семинар по дискретной математике под рук. М.В. Вялого и С.П. Тарасова в 2016 г., Вычислительный центр им. А.А. Дородницына РАН.

Результаты исследования докладывались на следующих конференциях:

1. Международная научная конференция студентов, аспирантов и молодых учёных «*Ломоносов-2013*», Москва, Россия, 2013.

2. 14th International Workshop «*Algebraic and Combinatorial Coding Theory*», Svetlogorsk, Russia, 2014.

3. IEEE International Symposium on Information Theory, Honolulu, USA, 2014.

4. Ninth International Workshop on Coding and Cryptography, Paris, France, 2015.

5. IEEE International Symposium on Information Theory, Hong Kong, China, 2015.

6. Международная научная конференция студентов, аспирантов и молодых учёных «*Ломоносов-2016*», Москва, Россия, 2016.

7. 15th International Workshop «*Algebraic and Combinatorial Coding Theory*», Albena, Bulgaria, 2016.

8. IEEE International Symposium on Information Theory, Barcelona, Spain, 2016.

Публикации

Результаты автора по теме диссертации опубликованы в 13 работах, список которых приведен в конце автореферата. Среди них 5 работ [1]-[5] в журналах из перечня ВАК и 8 работ [6]-[13] в рецензируемых трудах международных конференций.

Структура и объем диссертации

Диссертация состоит из введения, четырех глав, заключения и списка литературы, который включает 64 наименования. Объем диссертации составляет 80 страниц.

Краткое содержание диссертации

Во **введении** сформулированы основные объекты исследования, дан краткий исторический обзор предыдущих результатов, а также приведено основное содержание работы и ее апробация.

Первая глава состоит из семи разделов, посвященных границам асимптотических скоростей разделяющих кодов.

В первом разделе первой главы вводятся основные обозначения и даются определения исследуемых объектов.

Во втором разделе указаны основные приложения разделяющих кодов. Подробно описано применение разделяющих $(s,1)$ и (s,s)-кодов в задаче защиты авторских прав (в англоязычной литературе эта задача известна под названием *digital fingerprinting*).

В третьем разделе даются вспомогательные определения и формулируется первый новый результат, связывающий асимптотические скорости двоичных разделяющих кодов и свободных от перекрытий кодов.

Теорема 1. *Для скоростей двоичных разделяющих кодов и свободных от перекрытий кодов справедливы следующие неравенства*

$$\underline{R}_{cf}(s,\ell) \leqslant \underline{R}_{s}(s,\ell) \leqslant \overline{R}_{s}(s,\ell) \leqslant \overline{R}_{cf}(s-1,\ell). \tag{10}$$

Доказанные в этой теореме неравенства позволяют улучшить верхние границы скорости двоичных разделяющих кодов для многих значений параметров s и ℓ, а также получить новую верхнюю границу для случая фиксированного ℓ и $s \to \infty$, использовав известную оценку для скорости свободных от перекрытых кодов (2).

В четвертом разделе первой главы рассматриваются границы скоростей q-ичных разделяющих кодов. Следующая теорема позволяет улучшить верхние оценки скорости для многих значений параметров q, s и ℓ.

Теорема 2. *Введем обозначение* $m = \min\{\max\{q-s,1\},\ell\}$*.Для любого* $q \geqslant 2$ *скорости* q*-ичных разделяющих кодов* $\overline{R}_{s}^{(q)}(s,\ell)$ *ограничены следующим образом*

$$\overline{R}_{s}^{(q)}(s,\ell) \leqslant \frac{(2^{q-1}-1)\cdot \overline{R}_{s}(s,\ell)}{\log_2 q}, \tag{11}$$

$$\overline{R}_{s}^{(q)}(s,\ell) \leqslant \frac{\sum\limits_{k=m}^{\ell} \binom{q-1}{k} \cdot \overline{R}_{cf}(s-1,\ell)}{\log_2 q} \quad \text{для } s \geqslant 2, \tag{12}$$

$$\overline{R}_{s}^{(q)}(s,\ell) \leqslant \frac{\sum\limits_{k=m}^{\ell} \binom{q-1}{k-1} \cdot \overline{R}_{cf}(s,\ell-1)}{\log_2 q} \quad \text{для } \ell \geqslant 2. \tag{13}$$

Нижнюю границу дает следующая теорема:

Теорема 3. *При фиксированных* $q \geqslant 2$, $\ell \geqslant 1$ *и* $s \to \infty$ *для скоростей* $\underline{R}^{(q)}(s,\ell)$ q*-ичных разделяющих кодов справедливо следующее неравенство*

$$\underline{R}_{s}^{(q)}(s,\ell) \geqslant \frac{(q-1)^{\ell}}{e^{\ell}\ln q}\frac{1}{s^{\ell+1}}(1+o(1)). \tag{14}$$

Отметим, что отношение верхней и нижней границ, получаемых из теорем 2 и 3, ограничено не зависящей от q величиной, которая при $q > 2\ell$ и $s \to \infty$ может быть записана как

$$\frac{e(\ell+1)^{\ell+1}}{2(\ell-1)!}\ln s(1+o(1)).$$

В пятом разделе в виде теоремы 4 сформулированы рекуррентные неравенства, связывающие скорости разделяющих и полностью разделяющих кодов с разными параметрами s и ℓ.

Теорема 4. 1) *Для любых* $u \in [s-1]$, $v \in [\ell-1]$

$$\overline{R}_s(s,\ell) \leqslant \overline{R}_s(s-u,\ell-v)\cdot \max_{0\leqslant z\leqslant 1}\{z^u(1-z)^v + (1-z)^u z^v\}. \qquad (15)$$

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

$$\overline{R}_s(s,\ell) \leqslant \overline{R}_{cs}(s-u,\ell-v)\cdot \max_{0\leqslant z\leqslant 1}\{z^u(1-z)^v + (1-z)^u z^v\}. \qquad (16)$$

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

$$\overline{R}_s(s,\ell)/2 \leqslant \overline{R}_{cs}(s,\ell) \leqslant \overline{R}_{cs}(s-v,\ell-v)\frac{1}{2^{2v}}. \qquad (17)$$

Теорема позволяет улучшить верхние границы скоростей $\overline{R}(s,\ell)$ для многих значений параметров s и ℓ. Утверждение 3 дает следующую оценку для скорости разделяющих кодов

$$\overline{R}_s(s,s) = O\left(\frac{1}{2^{2s}}\right),$$

что довольно близко к нижней оценке

$$\underline{R}_s(s,s) \geqslant \frac{-\log_2(1-2^{-(2s-1)})}{2s-1} \sim \frac{\log_2 e}{s2^{2s}}, \quad s\to\infty,$$

которая доказывается с помощью стандартной техники случайного кодирования.

Доказательство теоремы опирается на несколько лемм. Первая лемма является аналогом границы Плоткина для разделяющих кодов со специально введенным “расстоянием”. Дадим вспомогательные определения.

Обозначим за X произвольный двоичный код мощности t и длины N. Пусть $\mathcal{U}$ и $\mathcal{V}$ - два непересекающихся множества кодовых слов кода X мощности u и v соответственно. Множество координат i кода X, разделяющих

$\mathcal{U}$ и $\mathcal{V}$, обозначим за $D_{u,v}(\mathcal{U},\mathcal{V},X)$. Определим среднюю по всем возможным выборам упорядоченной пары множеств $\mathcal{U}$ и $\mathcal{V}$ мощность $D_{u,v}(\mathcal{U},\mathcal{V},X)$

$$\overline{D}_{u,v}(X) \triangleq \sum_{\substack{\mathcal{U}\in\mathcal{P}_u(t),\,\mathcal{V}\in\mathcal{P}_v(t),\\ \mathcal{U}\cap\mathcal{V}=\varnothing}} \frac{|D_{u,v}(\mathcal{U},\mathcal{V},X)|}{\binom{t}{u+v}\cdot\binom{u+v}{u}},$$

и максимальную среднюю мощность

$$\overline{D}_{u,v}(t,N) = \max_X \overline{D}_{u,v}(X)$$

по всем кодам X фиксированной мощности t и длины N.

Лемма 1. (Граница Плоткина). *Выполнено следующее асимптотическое неравенство*

$$\varlimsup_{t\to\infty} \frac{\overline{D}_{u,v}(t,N(t))}{N(t)} \leqslant \max_{0\leqslant z\leqslant 1}\{z^u(1-z)^v + (1-z)^u z^v\}, \tag{18}$$

где $N(t)$ - произвольная функция.

Следующие три леммы с помощью величины $\overline{D}_{u,v}$ связывают минимальные длины разделяющих и полностью разделяющих кодов с разными параметрами.

Лемма 2. *Для любых $u\in[s-1]$ и $v\in[\ell-1]$ минимальная длина разделяющего $(s-u,\ell-v)$-кода мощности $t-(u+v)$ удовлетворяет неравенству*

$$N_s(t-(u+v),s-u,\ell-v) \leqslant \overline{D}_{u,v}(t,N_s(t,s,\ell)). \tag{19}$$

Лемма 3. *Для любого $v\in[\ell-1]$ и $u=v+s-\ell$, $s-(\ell-1)\leqslant u\leqslant s-1$ минимальная длина полностью разделяющего $(s-u,\ell-v)$-кода мощности $t-(u+v)$ удовлетворяет неравенству*

$$N_{cs}(t-(u+v),s-u,\ell-v) \leqslant \overline{D}_{u,v}(t,N_s(t,s,\ell)).$$

Лемма 4. *Для любого $v\in[\ell-1]$ минимальная длина полностью разделяющего $(s-v,\ell-v)$-кода мощности $t-2v$ удовлетворяет неравенству*

$$2N_{cs}(t-2v,s-v,\ell-v) \leqslant \overline{D}_{v,v}(t,N_{cs}(t,s,\ell)).$$

В шестом разделе приведены таблицы с численными значениями верхних границ скорости двоичных и троичных разделяющих кодов. Новые оценки, полученные в представленной диссертации, сравниваются с наилучшими ранее известными.

В седьмом разделе содержатся доказательства теорем и лемм.

Вторая глава состоит из пяти разделов, посвященных верхним границам асимптотических скоростей классических дизъюнктивных кодов и дизъюнктивных кодов со списочным декодированием.

В первом разделе второй главы даются основные определения и вводятся обозначения, используемые в дальнейшем.

Во втором разделе мы напоминаем известные нижние и верхние границы асимптотических скоростей $\overline{R}_{cf}(s)$ и $\underline{R}_{cf}(s)$ классических двоичных дизъюнктивных кодов. Для доказательства верхней границы А.Г. Дьячковым и В.В. Рыковым[36] была построена рекуррентная последовательность $\overline{R}_{DR}(s)$, ограничивающая сверху скорость дизъюнктивных кодов, и было показано, что

$$\overline{R}_{DR}(s) \leqslant \frac{2\log_2 s}{s^2}(1+o(1)), \qquad s \to \infty,$$

однако нижней границы для $\overline{R}_{DR}(s)$ получено не было. В этом разделе формулируется теорема, устанавливающая нижнюю границу для $\overline{R}_{DR}(s)$.

Теорема 5. *Если $s \geqslant 8$, то рекуррентная последовательность $\overline{R}_{DR}(s)$ удовлетворяет неравенству*

$$\overline{R}_{DR}(s) \;\geqslant\; \frac{2\log_2[(s+1)/8]}{(s+1)^2}, \qquad s \geqslant 8. \tag{20}$$

Теорема позволяет утверждать, что

$$\overline{R}_{DR}(s) \;=\; \frac{2\log_2 s}{s^2}(1+o(1)), \qquad s \to \infty, \tag{21}$$

то есть сделанная ранее оценка рекуррентной функции $\overline{R}_{DR}(s)$ была асимптотически оптимальна, и нельзя улучшить верхнюю границу скорости $\overline{R}_{cf}(s)$ с помощью более аккуратных оценок последовательности $\overline{R}_{DR}(s)$.

В третьем разделе речь идет о верхних границах скоростей списочных дизъюнктивных кодов, которые являются обобщением классических дизъюнктивных кодов. Формулируется следующая теорема.

Теорема 6. (Верхняя рекуррентная граница $r_L(s)$). *Имеют место следующие три утверждения.*

1. Для любого фиксированного $L \geqslant 1$ скорость СД s_L-кодов удовлетворяет неравенству

$$\overline{R}_L(s) \leqslant \min\left\{\frac{1}{s}, r_L(s)\right\}, \qquad s = 1, 2, \ldots,$$

в правой части которого последовательность $r_L(s)$, $s = 1, 2, \ldots$, определяется рекуррентно:

[36] Дьячков А.Г., Рыков В.В. Границы длины дизъюнктивных кодов // *Пробл. передачи информ.*, **18**:3 (1982), 7–13.

- *если* $1 \leqslant s \leqslant L$, *то*

$$r_L(s) \triangleq 1/s, \qquad s = 1, 2, \ldots, L\,; \tag{22}$$

- *если* $s \geqslant L+1$, *то* $r_L(s)$ *является единственным решением уравнения*

$$r_L(s) \triangleq \max_{(26)} f_{\lfloor s/L \rfloor}(v), \qquad s = L+1, L+2, \ldots, \tag{23}$$

в котором при $n = 1, 2, \ldots$ *функция* $f_n(v)$ *параметра* v, $0 < v < 1$, *определена равенствами*

$$h(v) \triangleq -v \log_2 v - (1-v)\log_2(1-v), \qquad 0 < v < 1, \tag{24}$$

и

$$f_s(v) \triangleq h(v/s) - v\, h(1/s), \quad 0 < v < 1, \quad s = 1, 2, \ldots, \tag{25}$$

а максимум берется по всем v, *удовлетворяющим условию*

$$0 < v < 1 - \frac{r_L(s)}{\min\left\{\frac{1}{s-1}, r_L(s-1)\right\}}; \tag{26}$$

- *если* $s > 2L$, *то уравнение* (23) *можно записать в виде равенства*

$$r_L(s) = f_{\lfloor s/L \rfloor}\left(1 - \frac{r_L(s)}{\min\left\{\frac{1}{s-1}, r_L(s-1)\right\}}\right), \quad L \geqslant 1, \quad s > 2L. \tag{27}$$

2. Для любого $L \geqslant 1$ *существует целое число* $s(L) \geqslant 2$, *такое, что*

$$r_L(s) = \begin{cases} \geqslant 1/s, & \text{если} \quad s = s(L) - 1, \\ < 1/s, & \text{если} \quad s \geqslant s(L), \end{cases}$$

и $s(L) = L \log_2 L(1 + o(1))$ *при* $L \to \infty$.

3. Если $L \geqslant 1$ *фиксировано и* $s \to \infty$, *то*

$$r_L(s) = \frac{2L \log_2 s}{s^2}(1 + o(1)). \tag{28}$$

Рекуррентная граница $r_L(s)$, определяемая (22)-(27), а также ее асимптотика (28) являются обобщением рекуррентной границы $\overline{R}_{DR}(s)$ и асимптотики (21). Второе утверждение теоремы отвечает на вопрос, с какого момента новая оценка становится лучше тривиальной оценки $1/s$.

В четвертом разделе дается определение дизъюнктивных планов поиска и описывается их связь со списочными дизъюнктивными кодами. Приведена

таблица с численными значениями верхних границ скорости дизъюнктивных планов поиска, получаемыми из теоремы 6.

В пятом разделе приводятся доказательства теорем 5 и 6.

Третья глава состоит из трех разделов, в которых исследуется пропускная способность почти дизъюнктивных кодов.

В первом разделе даны основные определения и установлены некоторые простые свойства, в том числе доказана верхняя граница для пропускной способности почти дизъюнктивных кодов $C(s) \leqslant 1/s$.

Во втором разделе сформулирована основная теорема главы, дающая нижнюю границу для пропускной способности почти дизъюнктивных кодов.

Теорема 7. *Справедливы следующие два утверждения.*

1. *Величина $C(s)$ удовлетворяет неравенствам:*

$$C(s) \geqslant \underline{C}(s) \triangleq \max_{0<Q<1} C(s,Q) = C(s,Q(s)), \qquad s \geqslant 1, \tag{29}$$

$$C(s,Q) \triangleq h(Q) - [1-(1-Q)^s]\, h\left(\frac{Q}{1-(1-Q)^s}\right), \quad s \geqslant 1, \quad 0 < Q < 1, \tag{30}$$

2. *При $s \to \infty$ асимптотика границы случайного кодирования $\underline{C}(s)$, задаваемой (29)–(30), и асимптотика оптимального значения $Q(s)$ в (29) имеют вид:*

$$\underline{C}(s) = \frac{\ln 2}{s}(1+o(1)), \quad Q(s) = \frac{\ln 2}{s}(1+o(1)). \tag{31}$$

Для доказательства теоремы рассматривается ансамбль $E(N,t,Q)$ двоичных $(N \times t)$-матриц X с N строками и t столбцами, где столбцы выбираются независимо и равновероятно из множества столбцов фиксированного веса $\lfloor QN \rfloor$, $0 \leqslant Q \leqslant 1$. Оценивается матожидание мощности множества $\mathbf{B}(s,X)$ плохих совокупностей из s столбцов, объединение которых покрывает посторонний столбец. Для оценки матожидания рассматривается условная вероятность $\Pr\left\{ \mathcal{S} \in \mathbf{B}(s,X) \,\middle|\, \left|\bigvee_{i\in\mathcal{S}} \boldsymbol{x}(i)\right| = k \right\}$ того, что некая совокупность является плохой при условии, что вес ее объединения равен k. Далее эта вероятность оценивается сверху как минимум из произведения количества посторонних столбцов на вероятность покрытия одного постороннего и единицы. Следующая лемма устанавливает, что эта оценка отличается от настоящего значения не более чем в константу раз, а это значит, что оценка пропускной способности на ансамбле точна.

Лемма 5. *Пусть* $\lfloor QN \rfloor \leqslant k \leqslant \min\{N, s\lfloor QN \rfloor\}$. *Для условной вероятности* $\Pr\left\{ \mathcal{S} \in \mathbf{B}(s, X) \,\middle|\, \left|\bigvee_{i\in\mathcal{S}} \boldsymbol{x}(i)\right| = k \right\}$ *выполнена следующая оценка*

$$\Pr\left\{ \mathcal{S} \in \mathbf{B}(s, X) \,\middle|\, \left|\bigvee_{i\in\mathcal{S}} \boldsymbol{x}(i)\right| = k \right\} \geqslant \min\left\{ 1/4;\ \frac{t-s}{2}\frac{\binom{k}{\lfloor QN\rfloor}}{\binom{N}{\lfloor QN\rfloor}} \right\}. \qquad (32)$$

Также отметим, что в статье [2] доказан и более сильный результат, а именно то, что на рассматриваемом ансамбле нельзя улучшить нижнюю оценку пропускной способности из теоремы 7 и для более общего случая почти дизъюнктивных кодов со списочным декодированием, т.е. пропускная способность на ансамбле не зависит от длины списка.

Для завершения доказательства теоремы нужно разобраться с поведением вероятности $\Pr\left\{\left|\bigvee_{i\in\mathcal{S}} \boldsymbol{x}(i)\right| = k\right\}$ в зависимости от параметра k, в чем помогает следующая лемма.

Лемма 6. *Пусть* X_n *и* Y_n *- две последовательности случайных величин, i-ая случайная величина принимает целые значения из отрезка* $[0, i]$, *причем для любого* $\varepsilon > 0$ *существуют* $\delta_1(\varepsilon) > 0$ *и* $\delta_2(\varepsilon) > 0$, *такие, что*

$$\Pr\left(\left|\frac{X_n}{n} - q_1\right| > \varepsilon\right) < 2^{-\delta_1(\varepsilon)n}(1 + o(1)), \qquad n \to \infty,$$

$$\Pr\left(\left|\frac{Y_n}{n} - q_2\right| > \varepsilon\right) < 2^{-\delta_2(\varepsilon)n}(1 + o(1)), \qquad n \to \infty.$$

Пусть случайная величина Z_n *равна весу объединения двух случайных столбцов веса* X_n *и* Y_n *соответственно. Тогда для любого* $\varepsilon > 0$ *существует* $\delta(\varepsilon) > 0$, *такая, что*

$$\Pr\left(\left|\frac{Z_n}{n} - (q_1 + q_2 - q_1q_2)\right| > \varepsilon\right) < 2^{-\delta(\varepsilon)n}(1 + o(1)), \qquad n \to \infty. \qquad (33)$$

В последнем разделе содержатся доказательства теоремы 7 и лемм 5, 6.

Четвертая глава состоит из пяти разделов, в которых обсуждается задача многоступенчатого поиска дефектов.

В первом разделе формулируется рассматриваемая задача, вводятся определения и приводятся некоторые ранее известные результаты по этой теме. Задача заключается в поиске малого числа дефектов элементов среди большого числа объектов с помощью тестов специального вида.

Во втором разделе задача многоступенчатого поиска дефектов описывается на языке гиперграфов. Каждому объекту соответствует вершина, а каждому гиперребру – возможное множество дефектов. Изначально в гиперграфе

проведены все гиперребра, чей размер не превосходит максимального возможного числа дефектов s. После проведения тестов мы узнаем, что некоторые вершины уже не могут быть дефектными, поэтому гиперребра, содержащие эти вершины, стираются. Используя свойства гиперграфа, порожденного проведенными на текущий момент тестами, мы можем придумывать тесты для последующих ступеней алгоритма. Общая идея алгоритма заключается в том, что мы стараемся разбить вершины на множества, в которых будет по одному дефекту, а потом отдельно разобраться с каждым из этих множеств.

В третьем разделе описанная ранее процедура применяется для двух дефектов вместе с возможными в этом частном случае оптимизациями. Предлагается такая первая ступень поиска, после которой порожденный граф имеет небольшое хроматическое число. Хроматические классы этого графа и будут теми множествами, которые содержат не более одного дефекта. В итоге получается четырехступенчатый алгоритм, чья скорость поиска достигает теоретико-информационной границы, т.е. доказана следующая теорема:

Теорема 8. *Асимптотическая скорость поиска двух дефектов четырехступенчатым алгоритмом равна 1/2,*

$$\overline{R}^{(4)}(2) = \underline{R}^{(4)}(2) = \frac{1}{2}.$$

В четвертом разделе рассмотрен общий случай произвольного числа дефектов и доказана следующая теорема

Теорема 9. *Скорость поиска s дефектов $2s-1$-ступенчатым алгоритмом удовлетворяет неравенству*

$$\frac{1}{2s-1} \leqslant \underline{R}^{(2s-1)}(s) \leqslant \overline{R}^{(2s-1)}(s) \leqslant \frac{1}{s}.$$

В пятом разделе построены таблицы, в которых приведено минимальное количество тестов, за которое предлагаемый алгоритм находит 2 дефектных элемента. При построении таблиц были также использованы некие дополнительные оптимизации, не влияющие на асимптотическую скорость поиска, но позволяющие улучшить результат для конечного количества размера множества, в котором производится поиск.

В **заключении** перечислены основные результаты и возможные направления дальнейших исследований.

В настоящей диссертационной работе были получены новые нижние и верхние границы для асимптотических скоростей $\overline{R}_s^{(q)}(s,\ell)$ и $\underline{R}_s^{(q)}(s,\ell)$ разделяющих кодов. Тем не менее как в случае фиксированного ℓ и $s \to \infty$, так и в случае $s = \ell$, $s \to \infty$ между верхними и нижними границами остается зазор, по порядку равный логарифму от самих оценок.

Также в диссертации были установлены верхние оценки асимптотической скорости дизъюнктивных кодов со списочным декодированием, обобщающие ранее доказанные границы для классических дизъюнктивных кодов. Представляет интерес порядок главного члена асимптотики скорости списочных дизъюнктивных кодов $R_L(s)$ при $s \to \infty$. На текущий момент существуют две гипотезы, которые следуют из доказанных верхней и нижней границ: L/s^2 и $L \ln s/s^2$.

Кроме того, в диссертации доказаны новые оценки для пропускной способности $C(s)$ почти дизъюнктивных кодов. В дальнейшем интересно было бы найти константу в асимптотике пропускной способности почти дизъюнктивных кодов $C(s)$. Сейчас известно только, что $(1 + o(1)) \ln 2/s \leqslant C(s) \leqslant 1/s$.

Наконец, были рассмотрены многоступенчатые алгоритмы поиска дефектов в задаче группового тестирования. В частности, был предложен четырехступенчатый алгоритм поиска двух дефектов, достигающий теоретико-информационной границы сложности. Дальнейшее исследование темы может быть связано с поиском достигающих теоретико-информационной границы алгоритмов с ограниченным числом ступеней для произвольного числа дефектов.

Благодарности

Автор глубоко благодарен и признателен своему научному руководителю профессору Аркадию Георгиевичу Дьячкову за постановку интересных задач, обсуждение результатов и постоянное внимание к работе, а также слушателям и докладчикам семинара по теории кодирования в ИППИ РАН за полезные замечания и предложения.

Работы автора по теме диссертации

[1] Дьячков А. Г., Воробьев И. В., Полянский Н. А., Щукин В. Ю., Границы скорости дизъюнктивных кодов // *Пробл. передачи информ.*, **50**:1 (2014), 31–63. [Дьячкову А. Г. принадлежат постановка задачи, теорема 3, предложения 1–3; Воробьеву И. В. — теоремы 1 и 6; Полянскому Н. А. — теоремы 2, 4 и 5; Щукину В. Ю. — теорема 7]

[2] Дьячков А. Г., Воробьев И. В., Полянский Н. А., Щукин В. Ю., Почти дизъюнктивные коды со списочным декодированием // *Пробл. передачи информ.*, **51**:2 (2015), 27–49. [Дьячкову А. Г. принадлежат постановка задачи и предложение 1; Воробьеву И. В. — предложение 2, пункты 1 (о пропускной способности) и 2 теоремы 4; Полянскому Н. А. — пример 1,

предложение 3 и теорема 1; Щукину В. Ю. — теорема 2, пункты 1 (об экспоненте ошибки) и 3 теоремы 4]

[3] D'yachkov A. G., Vorobyev I. V., Polyanskii N. A., Shchukin V. Yu. Cover-free codes and separating system codes // *Designs, Codes and Cryptography*, (2016), doi:10.1007/s10623-016-0265-9.[Дьячкову А. Г. принадлежит постановка задачи; Воробьеву И. В. — теорема 1, пункты 1, 2, 4 теоремы 2, теорема 3, леммы 3 и 4, пункты 2 и 3 теоремы 4; Полянскому Н. А. — леммы 1 и 2, пункт 1 теоремы 4; Щукину В. Ю. — пункты 3 и 5 теоремы 2]

[4] D'yachkov A. G., Vorobyev I. V., Polyanskii N. A., Shchukin V. Yu. Almost cover-free codes and designs // *Designs, Codes and Cryptography*, (2016), doi:10.1007/s10623-016-0279-3.[Дьячкову А. Г. принадлежит постановка задачи и предложение 1; Воробьеву И. В. — пример 1, теорема 1; Полянскому Н. А. — теоремы 2 и 3; Щукину В. Ю. — пример 2, предложения 2 и 3]

[5] D'yachkov A. G., Vorobyev I. V., Polyanskii N. A., Shchukin V. Yu. Symmetric disjunctive list-decoding codes // *Designs, Codes and Cryptography*, (2016), doi:10.1007/s10623-016-0278-4.[Дьячкову А. Г. принадлежит постановка задачи и предложение 4; Воробьеву И. В. — пункт 1 теоремы 1 и следствие 3; Полянскому Н. А. — предложение 1; Щукину В. Ю. — следствия 4, 5, пункты 2 и 3 теоремы 1, теорема 2, теорема 3]

[6] D'yachkov A. G., Vorobyev I. V., Polyanskii N. A., Shchukin V. Yu. Almost Disjunctive List-Decoding Codes // *Proc. of ACCT-14*, Svetlogorsk, (2014), 115–126. [Дьячкову А. Г. принадлежит постановка задачи; Воробьеву И. В. — пункты 1 (о пропускной способности) и 2 теоремы 2, лемма 1; Полянскому Н. А. — раздел 3 (о конструкциях); Щукину В. Ю. — пункты 1 (об экспоненте ошибки) и 3 теоремы 2, лемма 2]

[7] D'yachkov A. G., Vorobyev I. V., Polyanskii N. A., Shchukin V. Yu. Bounds on the Rate of Superimposed Codes // *Proc. IEEE Int'l Symp. Inf. Theory*, Honolulu, (2014), 2341–2345. [Дьячкову А. Г. принадлежат постановка задачи, теорема 3 и предложения 1–3; Воробьеву И. В. — теоремы 1 и 6; Полянскому Н. А. — теоремы 2, 4 и 5; Щукину В. Ю. — теорема 7]

[8] D'yachkov A. G., Vorobyev I. V., Polyanskii N. A., Shchukin V. Yu. Almost Cover-Free Codes and Designs // *Proc. IEEE Int'l Symp. Inf. Theory*, Hong Kong, (2015), 2899–2903. [Дьячкову А. Г. принадлежат постановка задачи и предложение 1; Воробьеву И. В. — пример 1 и теоремы 1; Полянскому Н. А. — теоремы 2 и 3; Щукину В. Ю. — пример 2, предложение 2 и 3]

[9] D'yachkov A. G., Vorobyev I. V., Polyanskii N. A., Shchukin V. Yu. Cover-Free Codes and Separating System Codes // *Proc. IEEE Int'l Symp. Inf. Theory*, Hong Kong, (2015), 2894–2898. [Дьячкову А. Г. принадлежат постановка задачи и предложения 1–2; Воробьеву И. В. — предложение 3, теоремы 1 и 1', пункты 1–3 теоремы 2, лемма 1; Полянскому Н. А. — лемма 2, пункты 4 и 6 теоремы 2; Щукину В. Ю. — лемма 3, пункт 5 теоремы 2]

[10] D'yachkov A. G., Vorobyev I. V., Polyanskii N. A., Shchukin V. Yu. Symmetric Disjunctive List-Decoding Codes // *Proc. IEEE Int'l Symp. Inf. Theory*, Hong Kong, (2015), 2236–2240. [Дьячкову А. Г. принадлежат постановка задачи и предложение 3; Воробьеву И. В. — пункт 1 теоремы 2, следствие 1'; Полянскому Н. А. — теорема 1; Щукину В. Ю. — следствия 2' и 3', пункты 2 и 3 теоремы 2, теорема 3]

[11] D'yachkov A. G., Vorobyev I. V., Polyanskii N. A., Shchukin V. Yu. Threshold Decoding for Disjunctive Group Testing // *Proc. of ACCT-15*, Albena, 2016. [Дьячкову А. Г. принадлежит постановка задачи; Воробьеву И. В. – раздел 3 (моделирование); Полянскому Н. А. – таблица 1; Щукину В. Ю. – теорема 3]

[12] D'yachkov A. G., Vorobyev I. V., Polyanskii N. A., Shchukin V. Yu. On Multistage Learning a Hidden Hypergraph // *Proc. IEEE Int'l Symp. Inf. Theory*, Barcelona, 2016. [Дьячкову А. Г. принадлежат постановка задачи, разделы 1 и 2; Воробьеву И. В. — принадлежат разделы 4 и 5 (общий алгоритм и алгоритм поиска 2 дефектов); Полянскому Н. А. — раздел 3 (поиск s дефектов); Щукину В. Ю. — раздел 6 (оптимизация алгоритма)]

[13] D'yachkov A. G., Vorobyev I. V., Polyanskii N. A., Shchukin V. Yu. On a Hypergraph Approach to Multistage Group Testing Problems // *Proc. IEEE Int'l Symp. Inf. Theory*, Barcelona, 2016. [Дьячкову А. Г. принадлежит постановка задачи; Воробьеву И. В. — теорема 2; Полянскому Н. А. — теорема 3; Щукину В. Ю. — теорема 1]