ОТЗЫВ

научного руководителя о диссертации
Воробьева Ильи Викторовича
«Разделяющие коды» (Separating Codes),
представленной на соискание ученой степени кандидата
физико - математических наук по специальности
01.01.05 - «теория вероятностей и математическая статистика».

Двоичным *разделяющим* (s,ℓ)-*кодом* (separating code) длины N и мощности t называется двоичная $N\times t$-матрица инцидентности X семейства из t подмножеств N-элементного множества, что для любых двух непересекающихся наборов $\mathcal{S}$ и $\mathcal{L}$ мощности $\le s$ и $\le \ell$, составленных из подмножеств семейства, найдется элемент исходного N-элементного множества, принадлежащий всем подмножествам из $\mathcal{S}$ и не принадлежащий никакому подмножеству из $\mathcal{L}$, или наоборот, принадлежащий всем подмножествам из $\mathcal{L}$ и не принадлежащий никакому подмножеству из $\mathcal{S}$. Двоичные разделяющие коды и их обобщения, называемые q-ичными, $q=2,3,\dots$ разделяющими (s,ℓ)-кодами, были мотивированы важными прикладными вопросами к теории информации и кодирования, возникшими в теории автоматов, в задачах, связанных с хэш-функциями, и задачах защиты авторских прав на цифровую продукцию. С понятием разделяющего кода оказалось тесно связано также имеющее свой широкий круг приложений понятие *свободного от перекрытий* (cover-free) (s,ℓ)-кода (кратко, СП (s,ℓ)-кода) длины N и объема t, которое определяется как двоичная $(N\times t)$-матрица инцидентности X семейства множеств, состоящего из t подмножеств конечного N-множества, для которого пересечение любых ℓ, $1\le\ell<t$, членов семейства подмножеств не покрывается объединением (не принадлежит объединению) любых других s, $1\le s<t$, членов семейства.

Диссертация И.В. Воробьева состоит из введения и четырех глав. Во вводной части диссертации приводится достаточно подробный обзор наиболее значимых предшествующих работ по разделяющим кодам, где описываются их приложения и обсуждаются результаты, полученные в этих работах. Дается также краткий обзор результатов всех 4 глав диссертационной работы.

Первая глава диссертации, которую И.В. Воробьев рассматривает как центральную и определяющую название его работы, посвящена выводу новых нижних и верхних асимптотических $(N\to\infty)$ границ максимальной мощности $t_s^{(q)}(N,s,\ell)$ q-ичных разделяющих (s,ℓ)-кодов длины N. Полученные автором верхние границы величины $t_s^{(q)}(N,s,\ell)$ основаны на известных результатах по исследованию логарифмической асимптотики $(N\to\infty)$ верхних границ максимального объема $t_{cf}(N,s,\ell)$ СП (s,ℓ)-кодов длины N. Поэтому целью главы 1 стало изучение связи между функциями целочисленных параметров $s,\ell=1,2,\dots$:

$$R_s^{(q)}(s,\ell)=\varlimsup_{N\to\infty}\frac{\log_q t_s^{(q)}(s,\ell,N)}{N},\qquad R_{cf}(s,\ell)=\varlimsup_{N\to\infty}\frac{\log_2 t_{cf}(N,s,\ell)}{N}$$

называемыми, следуя традиции комбинаторной теории кодирования, *скоростями* соответствующих кодов.

Прежде всего, отмечу очень важный результат, установивший связь между скоростью двоичных разделяющих кодов и скоростью кодов, свободных от перекрытий, который представляет собой неравенство $R_s^{(2)}(s,\ell) \le R_{cf}(s-1,\ell)$. Данное соотношение, несмотря на его простоту, ранее не было замечено многочисленными другими авторами. Эта оценка и установленная в 2002 году мной совместно с П.А. Виленкиным верхняя граница скорости СП (s,ℓ)-кодов позволяет диссертанту получить наилучшую в настоящее время верхнюю асимптотическую (ℓ-фиксировано, а $s \to \infty$) границу скорости двоичных разделяющих кодов, совпадающую с верхней границей для СП (s,ℓ)-кодов.

Далее И.В. Воробьев строит новые нетривиальные неасимптотические верхние оценки для скорости $R_s^{(q)}(s,\ell)$ q-ичных, $q>2$, разделяющих кодов через скорости СП (s,ℓ)-кодов и скорости $R_s^{(2)}(s,\ell)$ двоичных разделяющих кодов. Такие оценки позволяют получить наилучшую в настоящее время верхнюю асимптотическую ($\ell \ge 1$, $q \ge 2$ фиксированы, а $s \to \infty$) границу скорости q-ичных разделяющих кодов. Эту асимптотическую верхнюю границу диссертант сравнивает с нижней границей скорости

$$R_s^{(q)}(s,\ell) \ge \frac{(q-1)^\ell}{e^\ell \ln q}\,\frac{1}{s^{\ell+1}}\,(1+o(1)). \tag{1}$$

которая была известна ранее лишь при $\ell = 1$. В общем случае, т.е. при $\ell \ge 1$, автор доказывает нижнюю асимптотическую границу (1) с помощью разработанного им специально для разделяющих кодов метода случайного кодирования, и аналитически проверяет, что граница (1) в рассматриваемой асимптотике при достаточно больших значениях q, $q > 2\ell$, отличается от верхней границы на множитель не превышающий величину $e(\ell+1)^{\ell+1}\ln s/2(\ell-1)!\,(1+o(1))$, которая не зависит от объема алфавита q.

Несомненно важными результатами первой главы также являются полученные в ней неасимптотические верхние оценки скорости $R_s^{(2)}(s,\ell)$ двоичных разделяющих кодов, которые для многих конкретных значений параметров (s,ℓ) улучшают соответствующие значения в ранее построенных неасимптотических верхних границах. Данное улучшение, показанное в приводимых числовых таблицах, достигается с помощью применения нетривиальных рекуррентных неравенств для скорости $R_s^{(2)}(s,\ell)$, которые автор доказал существенно развивая и обобщая идеи вывода аналогичных неравенств для скорости $R_{cf}(s,\ell)$ СП (s,ℓ)-кодов и скорости двоичных разделяющим $(2,2)$-кодов, установленных В.С. Лебедевым в 2003 году и Ю.Л. Сагаловичем в 1994 году.

Вторая глава диссертации И.В. Воробьева посвящена исследованию асимптотики рекуррентной верхней границы для скорости классических двоичных дизъюнктивных s-кодов, а также уточнению и анализу асимптотики обобщающей ее рекуррентной верхней границы для скорости двоичных дизъюнктивных кодов со списочным декодированием (кратко, СД s_L-кодов), построенных в статьях 1982 и 1983 годов мной совместно с В.В. Рыковым. Предложенное в 1964 году У. Каутсом и Р. Синглтоном понятие двоичного дизъюнктивного s-кода (или СД s_1-кода) совпадает с частным случаем определения СП (s,ℓ)-кода при $\ell = 1$, т.е. СП $(s,1)$-кода. Введенное мной и В.В. Рыковым в 1983 году определение СД s_L-кода X означает, что дизъюнкция, т.е. покомпонентная булева сумма, любых s кодовых слов (столбцов) кода (матрицы) X может покрывать $\le L-1$ посторонних кодовых слов. В данных работах мы с В.В. Рыковым также ввели понятие дизъюнктивного s-плана, т.е. двоичной $(N \times t)$-матрицы (кода) X для которой дизъюнкция любых s столбцов (кодовых слов) отличается от дизъюнкции любых других s кодов слов.

Главным достижением диссертанта в главе 2 является давно ожидаемое мной от учеников построение верхней рекуррентной границы скорости СД s_L-кодов и вычисление главного члена ее асимптотики ($L \geq 1-$ фиксировано, а $s \to \infty$), который имеет вид $2L \log_2 s/s^2$, что здесь впервые доказано даже для наиболее важного частного случая $L = 1$. Из данного результата для частном случае СД $(s-1)_2$-кодов вытекает верхняя асимптотическая граница скорости дизъюнктивных s-планов, которая имеет вид $4 \log_2 s/s^2$. В качестве следствия построенной автором неасимптотической верхней границы скорости СД $(s-1)_2$-кодов устанавливается важное утверждение о том, что при $s \geq 11$ скорость дизъюнктивных s-планов $< 1/s$.

В **третьей главе** диссертации И.В. Воробьев рассматривает дизъюнктивную модель из N групповых проверок при статическом (неадаптивном) поиске неизвестного множества дефектов $\mathcal{S}$, $\mathcal{S} \subset [t] = \{1, 2, \ldots, t\}$, для которого параметр s, $1 \leq s < t$, определяет ограничение на его объем $|\mathcal{S}|$, т.е. предполагается, что $|\mathcal{S}| = s$. План статического поиска задается двоичной $(N \times t)$-матрицей X, а результаты N групповых проверок определяются как компоненты двоичного (из 0, 1) столбца Σ длины N, равного дизъюнкции столбцов X, номера которых составляют множество $\mathcal{S}$. При декодировании полным перебором (кратко, (bf)-декодировании, следуя англоязычной терминологии "brute force decoding"), т.е. когда столбец Σ сравнивается на совпадение с каждой из $\binom{t}{s}$ дизъюнкций, составленных из s-наборов столбцов X, неизвестное множество $\mathcal{S}$ восстанавливается однозначно тогда и только тогда, когда X является дизъюнктивным s-планом. При декодировании, которое состоит в проверке на покрытие ("check covering") столбцом Σ каждого из t столбцов кода X (кратко, (cc)-декодировании), неизвестное множество $\mathcal{S}$ восстанавливается однозначно тогда и только тогда, когда X является дизъюнктивным s-кодом. Отметим, что сложность (bf)-декодирования, равная $\Theta\left(\binom{t}{s} \log t\right)$, существенно превосходит сложность (cc)-декодирования, равную $\Theta(t \log t)$.

В главе 3 автор исследует введенный в 1975 году М.Б. Малютовым важный теоретико-вероятностный подход для дизъюнктивной модели статического поиска s дефектов во множестве $[t]$, когда неизвестное искомое множество дефектов $\mathcal{S}$ интерпретируется как случайная величина, принимающая равновероятные значения на множестве всех $\binom{t}{s}$ s-подмножеств множества $[t]$. При этом в качестве естественной характеристики такого поиска рассматривается вероятность ошибки, которая определяется как доля числа s-подмножеств среди всех $\binom{t}{s}$ s-подмножеств множества $[t]$, появление которых в качестве дефектных s-наборов приводит при (bf)-декодировании к неоднозначному решению.

В работе 1979 года я для ансамбля $(N \times t)$-матриц (кодов) X с независимыми компонентами кодовых слов и для ансамбля $(N \times t)$-матриц X с независимыми равновесными словами исследовал логарифмическую асимптотику средней по ансамблю вероятности ошибки, когда $t, N \to \infty$ и фиксирован параметр $R \sim \log_2 t/N$, $0 < R < 1$, называемый скоростью кода X. Я показал что при (bf)-декодировании для каждого ансамбля в диапазоне скоростей R , $0 < R < 1/s$, соответствующая средняя вероятность ошибки $\mathcal{P}_N^{(bf)}(s, t)$ дизъюнктивной модели поиска s дефектов во множестве $[t]$ экспоненциально стремится к 0 и имеет вид:

$$\mathcal{P}_N^{(bf)}(s, t) = \exp\{-N[E^{(bf)}(s, R) + o(1)]\}, \quad 0 < R < C_s^{(bf)} = 1/s. \tag{2}$$

Для ансамбля с равновесными словами экспонента $E^{(bf)}(s, R) > 0$ является монотонно убывающей функцией параметра R , $0 < R < 1/s$, и не меньше соответствующей положительной экспоненты для ансамбля с независимыми компонентами кодовых слов. Отсюда, в частности, следует доказанная в 1975 году М.Б. Малютовым и В.Л. Фрейдлиной тео-

рема Шеннона о том, что для стремящейся к 0 вероятности ошибки пропускная способность дизъюнктивных s-планов при (bf)-декодировании равна $C_s^{(bf)} = 1/s$. Отметим, что результат главы 2 показывает, что при $s \geq 11$ для нулевой ошибки при (bf)-декодировании пропускная способность дизъюнктивных s-планов $< C_s^{(bf)} = 1/s$.

Цель главы 3 диссертации – для ансамбля равновесных кодов и (cc)-декодирования вычислить логарифмическую асимптотику средней по ансамблю вероятности ошибки $\mathcal{P}_N^{(cc)}(s,t)$ для дизъюнктивной модели поиска s дефектов во множестве $[t]$. Другими словами, автор поставил задачу найти аналог формулы (2) для средней по ансамблю вероятности ошибки, если вместо (bf)- декодирования применяется существенно более простое (cc)-декодирование. И.В. Воробьев получил важный результат доказав, что

$$\mathcal{P}_N^{(cc)}(s,t) = \exp\{-N[E^{(cc)}(s,R) + o(1)]\}, \quad 0 < R < C_s^{(cc)} \sim \frac{\ln 2}{s}, \ s \to \infty, \tag{3}$$

где экспонента вероятности ошибки $E^{(cc)}(s,R) > 0$ является монотонно убывающей функцией параметра R , $0 < R < C_s^{(cc)}$. Построенную методом случайного кодирования функцию $C_s^{(cc)} < 1/s$, аргумента s, $s = 1, 2, \ldots$, можно назвать нижней границей пропускной способности дизъюнктивной модели поиска s дефектов, при использовании (cc)-декодирования. Экспоненту $E^{(cc)}(s,R)$ и экспоненту $E^{(bf)}(s,R)$, где $E^{(cc)}(s,R) < E^{(bf)}(s,R)$, можно интерпретировать как критерии увеличения вероятности ошибки при переходе от (bf)-декодирования к (cc)-декодированию.

В **четвертой главе** диссертации И.В. Воробьев занимается исследованием некоторых специальных адаптивных (последовательных) процедур поиска $\leq s$ дефектов, во множестве $[t]$ для дизъюнктивной модели с числом N групповых проверок, после проведение которых однозначно восстанавливается любой среди $\sum\limits_{i=0}^{s} \binom{t}{i}$ возможных наборов дефектов. Изучается асимптотическая ситуация, когда $t, N \to \infty$, фиксирован параметр s, задающий ограничение на число дефектов, и нужно максимизировать параметр $R \sim \log_2 t/N$, $0 < R < 1$, называемый скоростью поиска. Диссертант рассматривает адаптивные многоступенчатые процедуры, где N проверок разбиты на конечное (не зависящее от t и N) число p, $p \geq 1$, ступеней, а внутри каждой ступени проводятся статические проверки, построение которых может зависеть лишь от результатов проверок, проведенных на предыдущих ступенях. Пусть $R^{(p)}(s)$ обозначает максимальную скорость R для поиска $\leq s$ дефектов, которую можно достичь на p-ступенчатых процедурах. Отметим верхнюю границу $R^{(p)}(s) \leq 1/s$, которая, очевидно, достигается на адаптивных стратегиях поиска.

Впервые общую задачу исследования многоступенчатых стратегий поиска в статье 2009 года предложил П. Дамашке, где получил, в частности, первые нетривиальные результаты по нижним границам для скорости $R^{(p)}(s)$. Для $s = p = 2$ построением легко реализуемой 2-ступенчатой стратегии поиска ≤ 2 дефектов он показал, что $R^{(2)}(2) \geq 2/5 = 0.40$, а применяя стратегию, существование которой доказано методом случайного кодирования для дизъюнктивных 2-кодов, установил, что $R^{(2)}(2) \geq 1/2.44 = 0.41$.

Указанные результаты П. Дамашке можно сопоставить с результатами четвертой главы диссертации, где с помощью построения достаточно легко реализуемых процедур поиска И.В. Воробьев для частного случая поиска ≤ 2 дефектов с помощью 4-ступенчатой процедуры показывает, что $R^{(4)}(2) = 1/2$, а в общем случае поиска $\leq s$ дефектов устанавливает нижнюю границу $R^{(2s-1)}(s) \geq 1/(2s-1)$.

Перечень результатов диссертационной работы и мой комментарий показывает весьма широкий охват актуальных направлений комбинаторной теории кодирования, рассматриваемых И.В. Воробьевым. Общей является только классическая постановка задачи основателя теории информации К. Шеннона по исследованию логарифмической асимптотики границ объема и вероятности ошибки оптимальных кодов. Конкретные задачи, решенные автором, отражают его свободное владение различными комбинаторными, вероятностными и аналитическими методами. На мой взгляд, результаты И.В. Воробьева по границам для разделяющих кодов, представленные в основной первой главе, определившей название диссертации, и использующие методы и технику доказательств, разработанные в теории дизъюнктивных кодов, существенно улучшают аналогичные результаты других достаточно известных в теории кодирования авторов. Это является критерием высокой значимости достижений самого диссертанта.

Диссертация И.В Воробьева несомненно удовлетворяет всем требованиям «Положения о порядке присуждения ученых степеней» Высшей аттестационной комиссии Министерства образования и науки Российской Федерации, а ее автор, Воробьев Илья Викторович, заслуживает присуждения ему ученой степени кандидата физико - математических наук по специальности 01.01.05 - «теория вероятностей и математическая статистика».

Научный руководитель:
доктор физико - математических наук
по специальности 01.01.05,
профессор кафедры теории вероятностей
механико-математического факультета
ФГБОУ ВО «Московский государственный
университет им. М.В. Ломоносова» Дьячков Аркадий Георгиевич
тел. +7(495)9312119,
электронная почта: agd-msu@yandex.ru

20.09.2016

Подпись профессора А.Г. Дьячкова заверяю
и.о. декана механико - математического факультета МГУ,
доктор физико - математических наук,
профессор Чубариков Владимир Николаевич