ОТЗЫВ ОФИЦИАЛЬНОГО ОППОНЕНТА

о диссертационной работе Владислава Юрьевича Щукина «Дизъюнктивные коды со списочным декодированием» на соискание ученой степени кандидата физико-математических наук по специальности 01.01.05 — теория вероятностей и математическая статистика

Диссертация В. Ю. Щукина посвящена разработке вероятностных методов построения нижних границ асимптотических скоростей для ряда обобщений дизъюнктивных кодов и некоторым смежным вопросам.

Дизъюнктивный (s-дизъюнктивный) код — это множество двоичнных слов (одинаковой длины), в котором дизъюнкция (поразрядная) любых s слов не покрывает никакого другого его слова. Дизъюнктивные коды широко применяются в прикладной математике и информатике. Они используются в каналах множественного доступа, в задачах группового тестирования, в криптографии при распределении ключей, в задачах поиска файлов в системах хранения, имеют многочисленные другие применения.

Дизъюнктивные коды введены в 1964 г. У. Каутсом и Р. Синглетоном и интенсивно изучаются вплоть до настоящего времени. Наилучшие к данному моменту нижняя и верхняя границы скорости $R(s)$ дизъюнктивных кодов имеют вид $R(s) \gtrsim \dfrac{1}{s^2 \log e}$ и $R(s) \lesssim \dfrac{2 \log s}{s^2}$, соответственно (здесь и дальше все логарифмы двоичные). Нижняя установлена А. Г. Дьячковым, В. В. Рыковым и А. М. Рашадом в 1989 г. методом случайного кодирования для равновесного ансамбля кодов, верхняя получена А. Г. Дьячковым и В. В. Рыковым в 1982 г. на основе некоторой рекуррентной процедуры.

Повышение скоростей может быть достигнуто применением некоторых обобщений дизъюнктивных кодов. Одну из таких возможностей предоставляют дизъюнктивные коды со списочным декодированием. Другой способ повышения скорости дает использование кодов, обеспечивающих восстановление s-ок кодовых слов с малой вероятностью ошибки. Такими являются почти дизъюнктивные коды и почти дизъюнктивные коды со списочным декодированием. Тематика, связанная с оценкой характеристик дизъюнктивных кодов и их обобщений, привликла внимание ряда известных математиков. Помимо уже упомянутых У. Каутса, Р. Синглетона и А.Г. Дьячкова, к ним относятся П. Эрдеш, Л. А. Бассалыго, М. Б.

Малютов и некоторые другие.

Все указанные выше коды используют операцию дизъюнкции и потому могут быть только двоичными. Преодолеть это ограничение позволяют гиперкоды (и гиперкоды со списочным декодированием). Используемые в них q-ичные символы интерпретируются как одноэлементные множества и вместо дизъюнкции применяется теоретико-множественная операция объединения. Такие семейства кодов исследовались (под разными названиями) рядом авторов, в числе которых Д. Боне, Дж. Шоу, С. Блэкберн, Д. Стинсон, Ж. Коэн и другие.

Роль дизъюнктивных кодов и их обобщений в приложениях информатики, а также значительный интерес к ним со стороны математиков свидетельствуют об актуальности и важности этой проблематики.

Остановимся более подробно на содержании диссертации. Она включает введение и четыре главы. Во введении приведен обзор литературы по теме диссертации, изложено ее краткое содержание, перечислены основные результаты.

В первой главе рассматриваются дизъюнктивные коды со списочным декодированием — СД (s, L)-коды. Такой код представляет собой множество двоичных слов некоторой длины N, $N \geq s + L - 1$, в котором дизъюнкция любых s слов покрывает менее L других. Асимптотической скоростью СД (s, L)-кодов называется величина $R_L(s) = \varlimsup\limits_{N\to\infty} \dfrac{\log t(s, L, N)}{N}$, где $t(s, L, N)$ — максимальный объем СД (s, L)-кодов длины N.

В главе методом случайного кодирования на равновесном ансамбле кодов получена нижняя граница $\underline{R}_L(s)$ скорости $R_L(s)$ для любых s и L. Выражение для $\underline{R}_L(s)$ имеет весьма сложный вид: включает максимизацию по некоторому параметру, содержит величину, заданную неявно как корень некоторого уравнения, использует в своем описании достаточно сложные функции. Оно не позволяет составить представление о поведении этой границы. Используя нетривиальную технику, автор извлек из него явную асимптотику функции $\underline{R}_L(s)$ при $L = \text{const}$, $s \to \infty$ и получил асимптотическую оценку $R_L(s) \gtrsim \dfrac{L}{s^2 \log e}$, которая совпадает с наилучшей известной нижней границей, найденной в 2003 г. А. Г. Дьячковым другим методом. Отметим, что наилучшая к данному моменту верхняя граница, установленная в 2014 г. И. В. Воробьевым, имеет вид $R_L(s) \lesssim \dfrac{2L \log s}{s^2}$.

Метод, развитый в данной главе, давая ту же асимптотику нижней границы скорости $R_L(s)$, что и метод Дьячкова, имеет некоторое преимущество перед последним, поскольку он позволил также получить нижнюю границу максимальной скорости двухступенчатых групповых проверок, которая при $s > 2$ является наилучшей в данный момент. Она имеет вид $\log\left[\frac{(s-1)^{s-1}}{s^s}+1\right]$ и найдена путем вычисления значения $\lim\limits_{L\to\infty}\underline{R}_L(s)$.

Вторая глава посвящена почти дизъюнктивным кодам со списочным декодированием — СД (s, L, ϵ)-кодам. Они характеризуются тем, что доля s-ок кодовых слов, дизъюнкция которых покрывает более $L-1$ кодовых слов, не превышает ϵ. Автор нашел нижнюю границу для экспоненты ошибки $\mathbf{E}(s, L, R)$, которая представляет собой максимальный показатель экспоненциального убывания ошибки ϵ при фиксированной скорости R и возрастающей длине СД (s, L, ϵ)-кода. Для вывода этой границы использован метод случайного кодирования на ансамбле двоичных равновесных кодов, подобный методу, развитому в первой главе. Граница представлена в виде довольно сложного выражения и далее в диссертации проведено ее качественное исследование. Нижняя граница для экспоненты ошибки почти дизъюнктивных СД кодов получена впервые. В заключение главы чисто комбинаторным методом найдена верхняя оценка пропускной способности $C(s, L)$, которая для почти дизъюнктивных кодов со списочным декодированием имеет смысл некоторого аналога асимптотической скорости СД (s, L)-кодов и определяется как верхняя грань скорости СД (s, L, ϵ)-кодов при возрастающей длине кода и убывающей к нулю ошибке ϵ. Оценка имеет вид $C(s, L) \le 1/s$ и отличается от нижней оценки, установленной И. В. Воробьевым в 2015 г. в $\log e$ раз.

В третьей главе диссертации рассматривается новая постановка задачи о групповом тестировании, когда число s_{un} дефектных элементов не известно и не известна какая-либо его оценка. Требуется по результатам тестов определить, не превышает ли это число заданную константу s. В случае безошибочного принятия решения достаточно простые соображения показывают, что оптимальная процедура групповых проверок задается дизъюнктивным s-кодом. Однако при ненулевой вероятности ошибки применение почти дизъюнктивных кодов уже не является оптимальным. Более продуктивным оказывается другой подход, предложенный автором диссертации. Для проверки гипотезы $\{H_0 : s_{un} \le s\}$

против $\{H_1 : s_{un} \geq s+1\}$ вводится T-пороговый критерий, в соответствии с которым, если m^+ означает количество тестов с положительным исходом, то при $m^+ \leq T$ принимается гипотеза H_0, а при $m^+ \geq T+1$ — гипотеза H_1. Методом случайного кодирования на ансамбле равновесных кодов В. Ю. Щукин построил нижнюю границу для экспоненты ошибки принятия решения на основе T-порогового критерия и установил, что для максимального при варьировании T значения $\underline{\mathbf{E}}_{\mathrm{Thr}}(s)$ этой границы справедлива асимптотическая оценка $\underline{\mathbf{E}}_{\mathrm{Thr}}(s) \gtrsim \dfrac{\log e}{4s^2}$. На основе этой оценки и оценки пропускной способности из главы 2 автор делает вывод о превосходстве подхода, использующего пороговый критерий, над более традиционным подходом, основанным на почти дизъюнктивных кодах. Ранее пороговый критерий для задач группового тестирования не применялся.

Четвертая глава работы посвящена исследованию границ асимптотической скорости гиперкодов со списочным декодированием. Считается, что множество M q-ичных слов подчиняет слово v, если все буквы слова v содержатся в объединении букв, присутствующих в словах множества M. Множество q-ичных слов образует СД (s, L)-гиперкод, если любое его s-подмножество подчиняет менее L слов этого множества. Методом случайного кодирования на ансамбле кодов с независимыми одинаково распределенными q-ичными компонентами в главе получена нижняя граница асимптотической скорости $R_L^{(q)}(s)$ q-ичных СД (s, L)-гиперкодов. На ее основе автор установил асимптотическую оценку $R_L^{(q)}(s) \gtrsim \dfrac{L(q-1)}{s^2 \log q \log e}$ при $q, L = \mathrm{const}$, $s \to \infty$. Также получена верхняя оценка скорости $R_L^{(q)}(s) \lesssim \dfrac{2L(q-1)\log s}{s^2 \log q}$ при $q, L = \mathrm{const}$, $s \to \infty$. В заключение главы проведено подробное сравнение новых границ для $R_L^{(q)}(s)$ с ранее известными. Наибольший интерес вызывает улучшение как нижних, так и верхних границ, полученных относительно недавно в 2014 г. Ч. Шенгуэном и Дж. Же с их соавторами.

Результаты диссертации являются новыми. Они получены с помощью достаточно сложной техники, сочетающей вероятностные, комбинаторные, теоретико-информационные методы, методы оптимизации и асимптотического анализа.

Остановлюсь на замеченных недостатках.

Утверждение о том, что двоичные СД гиперкоды не имеют преиму-

ществ перед дизъюнктивными СД кодами, сформулированное как следствие теоремы 4.3.1, из этой теоремы не вытекает и его нельзя считать доказанным. Отмечу, что данное утверждение не используется при доказательстве других результатов.

Неясно, почему пороговый метод проверки гипотез, которому посвящена глава 3, назван «пороговым декодированием», ибо декодирования, как такового, там нет.

В определении s-активности присутствует условие $s \ll t$, делающее определение непроверяемым. Это условие следует удалить из определения и ввести там, где оно применяется.

Во многих местах диссертации автор утверждает, что полученные оценки являются наилучшими. Следовало бы уточнить, что они наилучшие из известных, ибо в обычном понимании термин «наилучшая» в применении к оценке означает ее неулучшаемость.

В верхней части страницы 38 сообщается как очевидный факт, что функция во второй строке не убывает по R. Для меня этот факт не очевиден и требует пояснений.

Предложение автора диссертации использовать для кодирования недоопределенных данных вместо дизъюнктивных кодов дизъюнктивные СД коды позволяет сократить длину кодов, но при этом нарушается условие, что отношение инцидентности между основными и недоопределенными символами должно переноситься на их кодирования.

Указанные замечания не очень принципиальны и не оказывают заметного влияния на общую оценку работы.

Результаты диссертации опубликованы в 13 научных работах, среди которых 6 работ в журналах из перечня ВАК, остальные 7 — в рецензируемых трудах международных конференций. Автореферат диссертации правильно и полно отражает ее содержание. Работа полностью соответствует паспорту специальности 01.01.05.

Диссертация В. Ю. Щукина является научно-квалификационной работой, в которой предложено значительное усовершенствование техники оценивания больших уклонений в методе случайного кодирования на пространстве двоичных равновесных кодов, что позволило построить наилучшие к данному моменту нижние границы скорости для некоторых обобщений дизъюнктивных кодов, впервые оценить снизу экспоненту ошибки почти дизъюнктивного СД кода, обосновать предложенный в диссертации новый для теории группового тестирования пороговый метод проверки гипотезы о том, что число дефектов ограничено заданной

величиной. Эти результаты имеют существенное значение для области вероятностных методов оценки информационных характеристик кодирования.

Представленная работа удовлетворяет требованиям ВАК к кандидатским диссертациям по математике, а ее автор, Владислав Юрьевич Шукин, заслуживает присуждения ученой степени кандидата физико-математических наук по специальности 01.01.05 — теория вероятностей и математическая статистика.

Официальный оппонент:
Шоломов Лев Абрамович,
главный научный сотрудник
лаборатории «Математические методы анализа и синтеза сложных систем»
Института системного анализа ФИЦ УИ РАН,
доктор физико-математических наук,
профессор

Адрес: 117312, г. Москва, пр-т 60-летия Октября, 9
Телефон: 8 (499) 135-54-44,
Адрес электронной почты: sholomov@isa.ru

22.02.2017
Шоломов Лев Абрамович

Подпись Л.А. Шоломова заверяю
Ученый секретарь ФИЦ ИУ РАН
к.ф.-м.н. Захаров
2017 г.