ОТЗЫВ

официального оппонента на диссертацию Лебедева Владимира Сергеевича
«Коды для каналов множественного доступа и задачи комбинаторного поиска»
на соискание ученой степени доктора физико-математических наук
по специальности 05.13.17 – теоретические основы информатики

Целью диссертационного исследования является разработка методов получения оценок характеристик кодов, используемых для передачи и защиты информации, в том числе обслуживающих каналы множественного доступа.

Актуальность темы диссертации объясняется широкими приложениями рассматриваемых в работе математических моделей защиты и передачи информации. В частности, часть моделей, соответствующих каналам множественного доступа, применимы и к другим задачам, которые возникают в производственном процессе, включая задачи поиска дефектных элементов.

В диссертации получены следующие новые научные результаты:

- Предложен алгоритм передачи информации по q-ичному каналу с ошибками при наличии безошибочной и бесплатной обратной связи. Показано, что алгоритм оптимален при большой (больше 1/q от числа передаваемых символов) доле ошибок в канале.

- Предложен адаптивный алгоритм поиска двух дефектных элементов при помощи проверок на наличие дефектных элементов в выбранном тестовом множестве. Предложенный алгоритм выигрывает у известных ранее по числу тестов.

- Предложены конструкции кодов, свободных от (w,r) перекрытий, доказана оптимальность некоторых из построенных кодов.

- Введено понятие композиционного расстояния, получена верхняя граница на скорость кодов с таким расстоянием. При помощи нее получена верхняя граница на скорость кодов, свободных от (w,r) перекрытий.

- Получена асимптотическая формула для метрической размерности пространсв Хэмминга над троичным и четверичным алфавитом.

- Получено оптимальное решение задачи поиска одного из d дефектных элементов для классической (когда определяется только наличие дефектных элементов в тестируемом множестве) и суммирующей (определяется число дефектных элементов в тестируемом множестве).

Структура диссертации. Диссертация состоит из введения, четырех глав, заключения и списка литературы из 175 источников. Общий объем диссертации составляет 201 стр.

Во введении обоснована актуальность исследований, сформулирована их цель и дана общая характеристика работы.

В первой главе диссертации рассматривается задача передачи информации по q-ичному каналу с безошибочной обратной связью. Подразумевается, что канал обратной связи не только безошибочный, но и пренебрежимо дешевый. По сути это означает, что передающая сторона знает, что получено на принимающей стороне, и в частности какие искажения

произошли в результате помех. Такая ситуация в реальности достаточно часто встречается, например, когда на передающей стороне устройство с ограниченной мощностью и дорогой энергией, а на принимающей – база, которая может транслировать обратно гораздо более мощный сигнал. Данная задача рассматривалась Берлекампом для двоичного канала. В диссертации предложены методы кодирования, отличные от метода Берлекампа и основанные на том, что при q больше 2 один из символов можно использовать как служебный. Рассмотренные методы являются обобщениями алгоритма 1-удаления, в котором служебный символ (обозначаемый через 0) используется для удаления последнего принятого символа, в случае когда передающая сторона увидела, что последний символ был принят с ошибкой. Во втором параграфе главы описывается алгоритм r-удаления. В этом варианте служебный символ 0 может использоваться как информационный для передачи основного сообщения, а для затирания используется последовательность из r нулей. Таким образом, для передачи таким способом сообщение должно быть закодировано при помощи последовательности символов, не содержащих r подряд идущих нулей, то есть избегающих служебной последовательности, использующейся для затирания. К слову, такой алгоритм применим и в двоичном случае. Основным результатом параграфа является оценка пропускной способности канала, которая показывает, что метод r-удаления оптимален при большой (больше 1/q) доле ошибок в канале. Для оценок скорости канала необходимо знать астмптотику числа последовательностей с такими запретами. Как отмечено в тексте диссертации, результаты в этом направлении известны, но автором диссертации предложено обобщение данной подзадачи, которое также может иметь значение для прикладных исследований, в том числе связанных с биоинформатикой. А именно, в третьем параграфе приведены оценки числа последовательностей данной длины над алфавитом из q элементов, содержащих вхождение двух подряд идущих нулей ровно r раз. В четвертом параграфе главы 2 рассмотрено другое обобщение алгоритма 1-удаления. Этот обобщенный алгоритм 1-удаления более сложный, чем алгоритм r-удаления, но в то же время более гибкий, полкольку передаваемая последовательность ограничена только по числу информационных нулей, но не по их расположению. Доказывается, что обобщенный алгоритм 1-удаления улучшает предыдущий алгоритм при меньшей доле ошибок в канале.

Во второй главе диссертации рассмотрена задача группового тестирования, которая также может быть описана в терминах дизъюнктивного канала множественного доступа (часто разные прикладные задачи описываются одной математической моделью). В общем случае имеется множество из N элементов, D из которых дефектно (это D может быть известно, а может быть неизвестно, но лежать в некоторых границах), необходимо найти некоторое число m дефектных элементов (не обязательно все, m может быть меньше чем D). Тестом является выбор некоторого множества тестируемых элементов, а результат теста, 0 или 1, зависит от числа дефектных элементов в выбранном наборе (эта зависимость определяется тестовой функцией). В первом параграфе описано решение задачи поиска двух дефектных элементов, что соответствует дизъюнктивному каналу множественного доступа с двумя активными пользователями. Рассматривается адаптивный алгоритм, когда за один шаг мы узнаем, есть ли дефектный элемент среди выбранного множества, а при выборе тестируемого множества можем учитывать результаты тестирования

предшествующих шагов. Предложен алгоритм решения данной задачи и доказано, что он оптимален для почти всех значений N (исходное число элементов). Далее рассмотрена задача поиска хотя бы одного дефектного элемента среди N элементов, среди которых D дефектных. При этом для тестирования используется более общая пороговая 0,1-значная функция, она дает значение 1, если число дефектных элементов в тестируемом множестве превышает некоторый порог (он может задаваться некоторой константой, а может зависеть от числа тестируемых элементов, составлять некоторую долю от этого числа). Во втором параграфе доказана нижняя оценка на необходимое число тестов для произвольной тестовой функции. В третьем параграфе предложен алгоритм, позволяющий достичь этой границы при «классической» тестовой функции (с порогом 1). В четвертом параграфе этот результат обобщен на случай, когда порог равен произвольному числу, и нужно найти m дефектных элементов. В пятом параграфе решается задача поиска в случае пороговой функции с порогом, пропорциональным числу тестируемых элементов.

В третьей главе диссертации рассмотрена проблема построения кодов, свободных от (w,r) перекрытий. Эта задача также мотивируется практическими приложениями, но, в отличие от предыдущих глав, в данной главе речь не о поиске оптимального адаптивного алгоритма, а о построении множества кодовых слов (или соответствующей кодовой матрицы), удовлетворяющих заданным условиям. Двоичная матрица с N строками и T столбцами называется кодом, свободным от (w,r) перекрытий, если для любых непересекающихся наборов W, R из w столбцов и из r столбцов соответственно найдется строка, содержащая единицы в позициях из W и нули в позициях из R. Целью является построение матрицы с наибольшим T при данном N или с наименьшем N при данном T. Есть два дуальных способа определить, что является кодовыми словами для данной задачи. Первый способ – считать кодом множество строк, элементов пространства Хэмминга размерности T. При этом в каждом хэммингове подпространстве, получаемом фиксацией нулями некоторых R координат и единицами некоторых W координат, обязана встретиться хотя бы одна кодовая вершина. Поиск таких совершенных кодов (когда условие «хотя бы одна» заменяется на «ровно одна») соответствует задаче о точном покрытии, которая на практике решается при помощи алгоритма «икс» Д. Кнута, что могло бы помочь для поиска небольших кодов при помощи вычислительной техники. К сожалению, данный подход не упомянут (хотя иногда он присутствует неявно, например, в определении тривиального кода). Для исследования используется другое представление, где кодовыми словами являются столбцы матрицы. Как показано автором диссертации, этот подход позволяет использовать методы теории кодов, исправляющих ошибки, для чего в работе вводится специальное композиционное расстояние. В этой главе несколько основных результатов, что отражено также в количестве печатных работ, в которых они опубликованы. Среди наиболее новаторских направлений следует отметить введение композиционного расстояния для исследования асимптотики и обобщение на q-ичный случай. В третьем и шестом параграфе доказывается оптимальность некоторых кодов, свободных от перекрытий (в третьем параграфе – некоторых конкретных кодов, в шестом – тривиальных кодов, при выполнении некоторых условий). В четвертом параграфе доказывается единственность, с точностью до эквивалентности, трех оптимальных кодов. В пятом

параграфе доказываются асимптотические границы на скорость (отношение логарифма мощности кода к длине кодовых слов) кодов, свободных от перекрытий. В седьмом параграфе рассмотрено обобщение задачи, когда наборы столбцов W и R из определения могут быть выбраны не произвольным образом, а из заданного множества (эти наборы в практических приложениях соответствуют коалициям пользователей, и довольно естественно, что некоторые группы пользователей не могут образовывать коалиции). Доказано несколько соотношений, из которых, как показано на примере, можно сделать вывод об оптимальности некоторых кодов в этом случае. В восьмом параграфе рассматриваются связи кодов, свободных от перекрытий, с разделяющими кодами и с задачами поиска дефектных элементов. В девятом параграфе доказываются асимптотические границы на скорость обобщенных q-ичных кодов, свободных от перекрытий (автор называет их окрашенными кодами).

В четвертой главе диссертации рассмотрена задача о метрической размерности пространств Хэмминга, которая имеет отношение к суммирующему каналу множественного доступа (в терминах задачи о тестировании, случай суммирующей тестовой функции). Метрическая размерность пространства – минимальное число вершин, по расстоянию до которых можно однозначно идентифицировать любую точку пространства. Подобный вопрос был рассмотрен в 1963 году Эрдешом и Реньи, которые доказали асимптотические оценки для метрической размерности двоичного пространства Хэмминга. В последствии для двоичного случая была установлена асимптотика этой величины. В первом параграфе главы доказывается асимптотика метрической размерности троичного и четверичного пространства Хэмминга. В трех последующих параграфах рассматривается поиск одного или нескольких дефектных элементов при помощи суммирующей функции. Исследования показывают, что суммирующая модель не дает преимущества для поиска только одного дефектного элемента.

Значимость полученных результатов. Теоретическая значимость исследования состоит в получении ответов на фундаментальные вопросы теории кодирования и комбинаторики, связанные с математическими моделями обработки информации, описываемыми каналом множественного доступа. Получены важные асимптотические результаты для скорости кодов и сложности тестирующих алгоритмов, для чего проведены как конструктивные исследования, заключающиеся в построении кодов и алгоритмов, так и теоретическое обоснование эффективности алгоритмов и оптимальности кодов. Спектр решенных проблем характеризуют как натуральные постановки, диктуемые практическими задачами (что, несомненно, означает высокую значимость для прикладных исследований; стоит также отметить, что одним и тем же математическим моделям соответствуют разные практические задачи), так и естественные вопросы, которые важны для фундаментальной математики безотносительно применения на практике (коды без перекрытий, метрическая размерность).

Автореферат правильно и полно отражает содержание диссертации.

Основные результаты, составляющие новизну диссертационной работы, получены лично автором или в нераздельном соавторстве. Они прошли широкую апробацию,

докладывались на многочисленных международных и российских научных конференциях и семинарах. Основные результаты диссертации Лебедева В.С. опубликованы: всего по теме диссертации опубликовано более 30 работ, в том числе 16 статей – в журналах, рекомендованных ВАК.

Замечания.

В тексте довольно много уделено внимания смысловому описанию подходов и алгоритмов, при этом формальные определения не всегда самодостаточны и даже корректны, их бывает трудно понять без окружающего контекста. Например, в определении кода, исправляющего ошибки в канале с безошибочной обратной связью, говорится, что декодер имеет систему множеств D, с помощью которой по полученной последовательности он может восстановить исходное сообщение. Не говорится, как эти множества связаны с кодирующей функцией (на самом деле каждое из них – это множество сообщений, которые декодер может получить при кодировании одного и того же сообщения данной функцией, если произошло не более чем t ошибок), а без этого определение не полно. При чтении этого можно и не заметить, поскольку в других местах подробно объясняется, как происходит кодирование и декодирование, но формальное определение некорректно. Это не значит, что в изложении отсутствует математическая строгость, но она «размазана» по тексту и иногда то, чего не хватает в какой-то формулировке, приходится искать в других местах или вспоминать. Другой пример – утверждение 5: «Обобщенный алгоритм 1-удаления асимптотически улучшает предыдущий алгоритм...» Какой «предыдущий», почему нельзя произносить его имя? То же самое про «классическую» тестовую функцию, в разделе, посвященном ей, уместно напомнить, что это слово значит.

В определении тактической конфигурации на с. 105 вместо «любой возможный на длине l набор единиц» должно быть «любой возможный на длине n набор из l единиц».

Довольно странно (хотя формально корректно) выглядит определение 10 разделяющего множества. Оно бы выглядело натурально для q-ичного случая, а для двоичного, когда упомянутые множества одноэлементны, проще определить как и написано в следующем параграфе после слов «таким образом». Возможно, лучше было оставить произвольное q в определении.

Обозначения [N], [i,j] определяются в разделе 2.2, а используются гораздо раньше.

Имеются отдельные несогласования падежей (например с. 90 «минимальное число тестов, требуемой для ...» и потери управления (например с. 143, «код, свободный от перекрытий размера T'» – на самом деле имеется ввиду размер кода, а не перекрытия), погрешности набора.

В целом указанные замечания никак не умаляют значимости для теории и практики результатов диссертационной работы Лебедева В.С.

Считаю, что результаты диссертационного исследования Лебедева В.С. соответствуют паспорту научной специальности 05.13.17 – теоретические основы информатики.

Диссертация Лебедева Владимира Сергеевича на тему «Коды для каналов множественного доступа и задачи комбинаторного поиска» имеет внутреннее единство,

обладает новизной и является завершенной научно-квалификационной работой, в которой на основании выполненных автором исследований разработаны теоретические положения, совокупность которых можно квалифицировать как научное достижение в области теоретических основ информатики.

Диссертация соответствует критериям п. 9 «Положения о присуждении учёных степеней», утвержденного постановлением Правительства РФ № 842 от 24.09.2013 г. (ред. от 01.10.2018 г.), а ее автор, Лебедев Владимир Сергеевич, заслуживает присуждения ученой степени доктора физико-математических наук по специальности 05.13.17 – теоретические основы информатики.

Официальный оппонент,
доктор физ.-мат. наук, профессор РАН 21.09.2021 Кротов Денис Станиславович,
главный научный сотрудник Федерального государственного бюджетного учреждения науки Институт математики им. С. Л. Соболева Сибирского отделения Российской академии наук, Новосибирск, 630090, проспект Академика Коптюга, 4.
E-mail: krotov@math.nsc.ru
Тел. +7(383)3297542

Подпись Кротова Д. С.
удостоверяю
Зав. орготделом Голубкина Л. А.
ИМ СО РАН
«21» 09 2021 г.