

ОТЗЫВ

официального оппонента д.ф.-м.н. Э.Н. Гордеева
на диссертацию Драгунова Никиты Аркадьевича
«Поиск частых (нечастых) элементов декартова произведения конечных
частичных порядков и приложения»,
представленную на соискание ученой степени
кандидата физико-математических наук по специальности
1.2.3 - «Теоретическая информатика, кибернетика»

Актуальность темы. Диссертация посвящена задачам поиска частых и нечастых элементов в данных. Задачи такого рода возникли при поиске ассоциативных правил и наиболее полно изучены для случая, когда обрабатываемые данные являются элементами декартова произведения бинарных антицепей, то есть отношения порядка на множествах значений признаков отсутствуют. Между тем в прикладных задачах значения признаков обычно упорядочены естественным образом, что делает необходимой разработку методов поиска частых и нечастых элементов для данных, представленных декартовым произведением конечных частично упорядоченных множеств. С ростом размерности баз данных полное перечисление всех частых и нечастых элементов становится неэффективным ввиду экспоненциального роста их числа, поэтому практический интерес представляет перечисление лишь максимальных частых и минимальных нечастых элементов.

Эффективность перечислительных алгоритмов принято оценивать временем работы. Для монотонной дуализации алгоритмы с полиномиальной задержкой известны лишь в отдельных частных случаях, а известные способы совместного перечисления максимальных частых и минимальных нечастых элементов либо эффективны только при существенном различии мощностей этих множеств, либо носят инкрементальный характер, при котором сложность шага зависит от числа ранее найденных решений задачи. Тем самым актуальна разработка новых алгоритмов для исследуемых в работе классов задач.

Таким образом, тема диссертации актуальна.

Структура и содержание работы. Диссертация состоит из введения, трех глав, заключения и списка литературы из 75 наименований. В первой главе строится последовательно-совместный подход к перечислению максимальных частых и минимальных нечастых элементов, во второй главе он применяется для решения задачи расшифровки двузначной монотонной функции, третья глава посвящена применению методов поиска частых и нечастых элементов частичных порядков к задаче классификации по прецедентам.

Введение. Обоснована актуальность темы, сформулированы цель и задачи исследования, изложены научная новизна и значимость полученных результатов, приведены положения, выносимые на защиту, а также сведения о публикациях и апробации работы.

Глава 1. Рассматривается построение множеств максимальных частых и минимальных нечастых элементов декартова произведения конечных частичных порядков. Рассмотрены два известных способа: последовательный, при котором одно из множеств строится непосредственно, а второе получается дуализацией первого, и совместный, предложенный К. Эльбассиони, где оба

множества перечисляются одновременно, однако трудоемкость шага возрастает по мере накопления найденных решений.

Основным результатом главы является последовательно-совместная схема перечисления, использующая элементы обоих подходов; ее корректность доказана (утверждения 1–3). Эксперименты на произведении конечных цепей с использованием наиболее быстрого асимптотически оптимального алгоритма дуализации RUNC-M+ показали, что предложенная схема работоспособна и в ряде случаев предпочтительнее некоторых известных ранее.

Глава 2. Поставлена задача расшифровки двузначной монотонной функции, которая задана на произведении конечных частичных порядков и определена через оракула: требуется найти все нули и единицы функции при минимальном числе обращений к оракулу. Традиционные решения нацелены на создание оптимальных по Шеннону (эффективных в худшем случае) алгоритмов, характерным представителем которых служит алгоритм В. Б. Алексеева для функций, заданных на k -значном n -мерном кубе. Диссертант устанавливает взаимно однозначное соответствие между верхними нулями двузначной монотонной функцией и максимальными частыми элементами, а также между нижними единицами и минимальными нечастыми элементами. Указанное соответствие позволяет модифицировать последовательно-совместный алгоритм из первой главы к задаче расшифровки; корректность построенного алгоритма доказана (утверждения 4–6).

Проведено экспериментальное сравнение обоих алгоритмов. Показано, что предлагаемый в диссертации алгоритм в среднем работает быстрее и меньше обращается к оракулу при больших k и n .

Глава 3. Рассматривается задача классификации по прецедентам. Диссертантом предлагается применять методы поиска частых (нечастых) элементов в данных при поиске для каждого класса фрагментов описаний прецедентов специального вида, названных правильными представительными элементарными классификаторами (ЭК). Правильный представительный ЭК ранга r является r -частым элементом в множестве прецедентов класса K и 1-нечастым в множестве прецедентов из других классов. Для произведения антицепей построен классификатор REC, для произведения цепей - его модификация REC+.

Теоретическое обоснование эффективности предложенных классификаторов дают теоремы 1–3, в которых получены верхние асимптотические оценки типичного числа и типичного ранга правильных ЭК для антицепей и для цепей. Приведено сравнение полученных оценок с известными оценками числа минимальных 1-нечастых ЭК (теорема 4), на построении которых базируется построение традиционных моделей логических классификаторов. Показано, что полученные оценки позволяют определять оптимальный ранг искомых ЭК по формуле, а не подбирать его численным перебором.

Заключение. Сформулированы основные результаты работы и намечены направления дальнейших исследований.

Научная новизна результатов. Новыми являются следующие результаты.

1. Последовательно-совместный подход к перечислению максимальных частых и минимальных нечастых элементов частичных порядков, и доказательство ее корректности.

2. Сведение задачи расшифровки двузначной монотонной функции к задаче совместного перечисления максимальных частых и минимальных нечастых элементов частичных порядков и построенный на этой основе алгоритм расшифровки, нацеленный на типичный, а не на худший случай.

3. Алгоритмы логической классификации REC и REC+, основанные на первоначальном поиске правильных ЭК и реализующие новую схему обучения логического классификатора.

4. Асимптотические оценки типичного числа и типичного ранга правильных ЭК для произведения антицепей и для произведения цепей, развивающие технику оценивания метрических характеристик логических классификаторов.

Обоснованность и достоверность результатов. Все алгоритмические построения сопровождаются доказательствами (утверждения 1–6, леммы, теоремы 1–3). Алгоритмы реализованы программно на языках C++ и Python и проверены на модельных и реальных данных; результаты усреднены по независимым запускам. Выводы теоретической и экспериментальной частей согласуются между собой.

Теоретическая и практическая значимость. Теоретическая значимость определяется расширением класса задач, для которых построены эффективные в типичном случае перечислительные алгоритмы, а также развитием техники получения асимптотических оценок числа решений.

Практическая значимость состоит в том, что предложенные алгоритмы доведены до программной реализации, работают с частично упорядоченными данными и конкурентоспособны с существующими ранее.

Оценки типичного ранга позволяют настраивать параметры классификаторов без перебора, что существенно для задач большой размерности.

Результаты работы вошли в отчеты по проектам РФФИ 19-01-00430 и РНФ 24-21-00301.

Публикации и апробация. По теме диссертации опубликовано 11 работ, в том числе 5 статей в журналах перечня ВАК; 5 работ индексируются в Scopus, 2 статьи – в Web of Science Core Collection. Результаты доложены на шести всероссийских и международных конференциях в 2019–2025 гг. Публикации отражают содержание диссертации в полном объеме; автореферат соответствует содержанию работы.

Замечания.

1. Так как не приводятся теоретические оценки качества сравниваемых на стр. 29–30 алгоритмов и Depth Project, а методика их экспериментального сравнения лишь косвенно анализирует их работу на некоторых классах задач, то вывод, что первый алгоритм «существенно быстрее» второго не всегда убедителен.
2. Безусловно, последовательно-совместный алгоритм, анализируемый на стр. 40–43, выглядит более изощренным по сравнению с конкурентами, чьи слабости учитывались при построении этого алгоритма, но сравнить теоретически хотя бы на примере узкого частного класса задач было бы интересно.
3. На стр. 55–56 автор приводит рассуждения о сравнении своего алгоритма и классического алгоритма Алексеева 1976 года. Конечно, показано, что для которых классов задач на численных экспериментах предложенный автором алгоритм выглядит лучше. Но для

алгоритма Алексева есть теоретические оценки сложности, а у алгоритма, предложенного диссертантом, их нет.

4. На стр. 80–85 описаны эксперименты, связанные с самым интересным, с моей точки зрения алгоритмом работы REC. Как раз по нему вопросов с теоретическим анализом самого алгоритма нет. Но этот анализ проведен в рамках парадигмы построения самого алгоритма и не ясно, как он может использоваться для сравнения с другими. Эксперимент показал, что для взятых из репозитория 9 задач для некоторых из них в некоторых случаях REC выглядит лучше. То есть имеет право на существование и применение. А для других задач и других случаев вопрос открыт.

Заключение по диссертации

Диссертационная работа Н. А. Драгунова представляет собой законченное самостоятельное исследование, в котором получены новые результаты в области дискретной математики и ее приложений к расшифровке двужначных монотонных функций и классификации по прецедентам. Указанные замечания не затрагивают существа полученных результатов и не влияют на общую положительную оценку работы.

Тематика диссертации соответствует паспорту научной специальности 1.2.3 – «Теоретическая информатика, кибернетика», в частности пунктам 1, 3 и 17.

Диссертация удовлетворяет требованиям, предъявляемым к диссертациям на соискание ученой степени кандидата наук Положением о присуждении ученых степеней, утвержденным постановлением Правительства Российской Федерации от 24 сентября 2013 г. № 842, а ее автор, Драгунов Никита Аркадьевич, заслуживает присуждения ученой степени кандидата физико-математических наук по специальности 1.2.3 – «Теоретическая информатика, кибернетика».

Официальный оппонент,
профессор кафедры «Информационная безопасность»
ФГАОУ ВО «Московский государственный технический
университет имени Н. Э. Баумана
(национальный исследовательский университет)»,
д.ф.-м.н., профессор



Э. Н. Гордеев

Адрес: 105005, г. Москва, ул. 2-я Бауманская, д. 5, стр. 1

Тел.: +7 (499) 263-63-91

E-mail: werhorngord@gmail.com

21. 09. 2026

