

ОТЗЫВ

официального оппонента доктора технических наук, старшего научного
сотрудника Фархадова Маиса Паша Оглы
на диссертацию Монгуш Чодураа Михайловны «Разработка метода и средств
фрагментации и дефрагментации формальных контекстов»
по специальности 05.13.17 – Теоретические основы информатики
на соискание ученой степени кандидата физико-математических наук

Актуальность. В настоящее время объемы данных, в том числе, текстовых, постоянно увеличиваются. Этот факт обуславливает необходимость исследований и разработок, направленных на создание аппаратных и программных средств обработки текстовых данных. Методы анализа текстов классифицируются по двум направлениям: методы, использующие статистики встречаемости слов в текстах; методы, основанные на применении семантических моделей текста. Данная диссертационная работа относится ко второму направлению. В ней используются методы анализа формальных понятий. Преимуществом методов анализа формальных понятий является универсальность решения типовых задач анализа данных. Однако при больших исходных данных возможности их практического применения ограничены. Диссертационная работа Монгуш Ч.М. посвящена преодолению этой проблемы.

В связи с вышеизложенным, выполненные Монгуш Ч.М. исследования по снижению размерности объектно-признаковой таблицы, описывающей некоторую предметную область, а также разработанные ею метод и средства фрагментации и дефрагментации формальных контекстов являются актуальными.

Содержание работы.

Диссертация состоит из введения, трех разделов, заключения, списка литературы. Общий объем диссертации составляет 105 страниц. Список использованных источников содержит 127 наименований.

Характеристика диссертационной работы по главам.

В первой главе вводятся основные положения и обозначения, используемые в работе, исследуется $\#P$ -полная задача нахождения всех формальных понятий формального контекста, приводится обзор существующих алгоритмов и программных комплексов решения поставленной задачи. Глава носит, в основном, обзорный характер.

Во второй главе подробно описывается метод «неискажающей» декомпозиции формального контекста, который позволяет снизить трудоемкость процесса нахождения всех формальных понятий исходного формального контекста. Для эффективной организации процесса декомпозиции определяются фрагменты формального контекста, исследуются свойства фрагментов для установления правил остановки итерационного процесса декомпозиции и восстановления искомого решения. Доказывается, что разложение формального контекста на фрагменты является «неискажающим» относительно формальных понятий, т. е. при декомпозиции ни одно формальное понятие не теряется и не появляются новые формальные понятия.

Приводится описание новых алгоритмов: формирования системы фрагментов, восстановления решетки формальных понятий, реализации запросов на извлечение знаний из построенной решетки формальных понятий. Для дополнительного снижения времени вычисления всех формальных понятий исходного контекста рассматриваются процедуры предобработки формального контекста. Приводится анализ результативности разработанных метода, алгоритмов и процедур.

В третьей главе описывается комплекс программ FSACorpus для проведения экспериментальных исследований на реальных данных, оценки результативности предложенных в диссертационной работе средств и включения FSACorpus в корпус тувинского языка.

Новизна результатов проведенных исследований. Новизна результатов диссертационной работы Монгуш Ч.М. заключается в том, что:

- разработан новый метод декомпозиции формального контекста на фрагменты без потери формальных понятий. В отличие от существующих методов анализа формальных понятий, предложенный метод позволяет уменьшить размерность формального контекста с сохранением всех искомым формальных понятий и тем самым повысить производительность известных алгоритмов нахождения всех формальных понятий формального контекста;

- впервые разработан алгоритм реализации предложенного метода «неискажающей» декомпозиции формального контекста. Алгоритм отличается от ранее существующих алгоритмов тем, что разлагает исходный формальный контекст без потери формальных понятий и восстанавливает решение поставленной задачи исходя из решений, полученных для подзадач.

Достоверность полученных результатов подтверждается строгими математическими выкладками, проведенными в работе с использованием математического аппарата анализа формальных понятий и теории решеток.

Теоретическая и практическая значимость работы заключается в развитии направления анализа формальных понятий и комбинаторной оптимизации. Описанный в диссертации метод существенно расширяет круг решаемых задач, позволяя использовать результаты диссертационной работы при решении задач определения всех максимально полных подматриц бинарной матрицы.

Разработанные программные средства успешно применены для задач анализа текста на реальных данных в рамках корпуса тувинского языка.

Полнота опубликования результатов работы.

Все основные результаты диссертации были опубликованы в 14 работах, в том числе имеется 4 научных публикации в журналах, рекомендованных ВАК. Основные результаты диссертации неоднократно докладывались на международных и всероссийских конференциях, получены 2 свидетельства о государственной регистрации программ для ЭВМ.

Содержание автореферата достаточно полно отражает диссертационную работу.

Замечания.

В ходе ознакомления с диссертацией возникли следующие замечания:

1. Целью работы указано повышение производительности алгоритмов. В работе нигде не представлено определение производительности алгоритмов и не указано отличие производительности алгоритмов от эффективности алгоритмов.
2. На стр. 31 идет речь о декомпозиции как о разбиении исходной задачи на подзадачи. Почему избегается употребление термина "динамическое программирование", хотя оно в точности подходит под заявленные действия? Возможно ли добавить существующие приемы динамического программирования к предложенным методам?
3. На стр. 63 в формуле для расчёта плотности контекста присутствует опечатка: в знаменателе должен быть n^2 . Хотя окончательный результат расчётов верен.

Указанные недостатки не являются определяющими и не оказывают существенного влияния на научную и практическую значимость полученных в диссертации результатов.

Общее заключение.

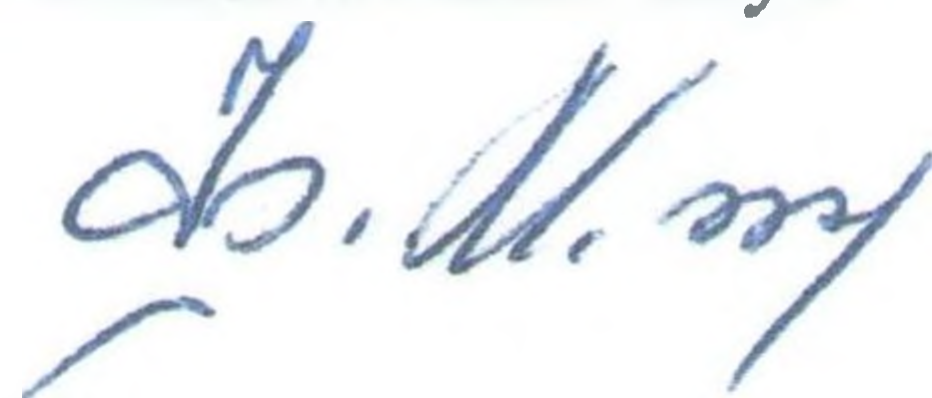
Диссертация является законченной научно-квалификационной работой, в которой решена актуальная научная проблема повышения производительности существующих алгоритмов нахождения множества всех формальных понятий и родственных с ней задач, путем декомпозиции исходного контекста с сохранением всех искомым формальных понятий.

Полученные результаты соответствуют п. 5 «Разработка и исследование моделей и алгоритмов анализа данных, обнаружения закономерностей в данных и их извлечениях, разработка и исследование методов и алгоритмов анализа текста, устной речи и изображений» паспорта специальности 05.13.17 – Теоретические основы информатики.

Считаю, что диссертационная работа Монгуш Ч.М. «Разработка метода и средств фрагментации и дефрагментации формальных контекстов»

соответствует требованиям п.9 «Положения о присуждении учёных степеней» постановления Правительства Российской Федерации от 24.09.2013 г. № 842, а её автор Монгуш Чодураа Михайловна заслуживает присуждения ей ученой степени кандидата физико-математических наук.

Главный научный сотрудник,
заведующий лабораторией
автоматизированных систем массового
обслуживания и обработки сигналов
ИПУ РАН, доктор технических наук



Фархадов Маис Паша Оглы

Почтовый адрес, телефон: Российская Федерация, 117997, г. Москва, ул.
Профсоюзная, д. 65, Федеральное государственное бюджетное учреждение
науки Институт проблем управления им. В.А. Трапезникова Российской
академии наук, тел. +7 (495) 334-89-10

E-mail: mais@ipu.ru



Подпись

Фархадова М.П.

ЗАВЕРЯЮ

ИНЖЕНЕР

Н.П.

17.12.19