

УТВЕРЖДАЮ
Декан факультета ИКТ,
ТТУ имени академика М.С.Осими,
к.т.н., доцент
Каримзода Дж.Х.



« 11 » 09 2023 г.

Заключение
расширенного заседания кафедры Автоматизированные системы
управления факультета Информационно-коммуникационных
технологий Таджикского технического университета
имени академика М.С. Осими

Диссертация на тему «Статистические закономерности распознавания однородности текстов с помощью γ -классификатора (Қонуниятҳои омории шиноҳти якҷинсагии матн бо истифода аз γ -таснифгар)» выполнена в отделе математического моделирования Института математики имени А. Джураева Национальной академии наук Таджикистана и на кафедре автоматизированных систем управления Таджикского технического университета имени академика М.С. Осими.

В период подготовки диссертации и по настоящее время диссертант Косимов Абдунаби Абдурауфович работает в Таджикском техническом университете имени академика М.С. Осими в должности и.о. доцента кафедры «Автоматизированные системы управления».

В 2011 году окончил Худжандский политехнический институт Таджикского технического университета имени академика М.С. Осими по специальности «Программное обеспечение вычислительной техники и автоматизированных систем».

Научный консультант: Усманов Зафар Джураевич, доктор физико-математических наук, академик НАНТ, профессор.

По результатам обсуждения диссертации на тему «Статистические закономерности распознавания однородности текстов с помощью γ -классификатора» принято следующее **заключение:**

Оценка выполненной соискателем работы. Диссертация Косимова Абдунаби Абдурауфовича на тему: «Статистические закономерности распознавания однородности текстов с помощью γ -классификатора» посвящена изучению проблемы распознавания однородности текстовых фрагментов.

Оценка актуальности диссертационной работы. Актуальность темы диссертационного исследования определяется тем, что она вошла в перечень первостепенных задач, записанных в Постановлении Правительства Республики

Таджикистан «Об утверждении программы применения и развития информационных технологий в таджикском языке» от 06.06.2005 № 188.

Настоящая диссертация является составной частью глобальной научной проблемы – автоматической обработки информации на естественном языке, признанной одной из актуальных проблем человечества. С надеждами на успешное разрешение последней связан вопрос о способности современной цивилизации контролировать, упорядочивать, осмысливать и использовать лавинообразный приток знаний, порождаемый её собственной деятельностью.

Одной из граней этой проблемы является проектирование автоматических систем распознавания новизны и адресности информации, охватывающих такие вопросы, как компиляция, плагиат, заимствование, идентификация авторства, сходство произведения и его перевода и т.п. В связи с развитием информационных технологий исследования в этой области знания заметно интенсифицировались по всему миру. Многочисленные научные публикации во всех высокоразвитых странах показывают особую роль данной проблематики, её непосредственное влияние на развитие науки и техники, на прогресс в сфере искусственного интеллекта, на ширококомасштабные приложения в мировой экономике.

Личное участие автора в получении результатов, изложенных в диссертации:

- 1) сформировал две электронные коллекции текстов, из которых первая предназначена для предварительного тестирования, а вторая – для оценки перспективности применения γ -классификатора;
- 2) исследовал цифровой портрет текста (ЦПТ) для распознавания автора текста;
- 3) установил статистическую эффективность применения γ -классификатора для распознавания авторов произведений;
- 4) определил минимальный размер незнакомого текста, пригодного для распознавания его автора;
- 5) исследовал эффективность применения высокочастотных элементов ЦПТ для идентификации автора текста;
- 6) установил статистическую эффективность применения γ -классификатора и исследования пригодности ЦП на основе распределения частотности различных алфавитных элементов текста для распознавания других признаков однородности, таких как тематики текста, язык, группа языков, оригинал и его перевод, стиль произведений, шифры научных работ и т.д.;
- 7) исследовал статистические закономерности распознавания однородных текстов на корпусах художественных литературных произведений;
- 8) определил эффективность применения γ -классификатора для атрибуции искусственно сгенерированных произведений авторов;
- 9) исследовал влияние порядка ЦПТ на распознавание однородности произведения с помощью γ -классификатора;
- 10) спроектировал и реализовал компьютерный программный комплекс для

распознавания (идентификации) однородности текста на основе различных ЦП текста и γ -классификатора.

Достоверность научных результатов и основных выводов подтверждается использованием фундаментальных теоретических положений, изложенных в литературных источниках, анализом обширного фактического материала, апробацией разработанных методических положений и методов оценки.

Практическая значимость диссертации и использование полученных результатов состоит в том, что на базе проведенных научных исследований разработано компьютерное программное приложение, которое с успехом применяется в государственной административной деятельности для автоматизации процесса обработки текстовой информации, в сфере криминалистики для установления авторства анонимных текстов, в области образования для обнаружения плагиата в курсовых и дипломных проектах.

Комплекс программ под названием «THR» (text homogeneity recognition) применён в следующих организациях:

1. Академия Министерства внутренних дел Республики Таджикистан.
2. Государственный комитет национальной безопасности Республики Таджикистан.
3. Институт языка и литературы имени Рудаки НАНТ.
4. Институт математики имени А.Джураева НАНТ.
5. ТГУ имени академика М.С. Осими.

Ценность научных работ соискателя, полнота изложения материалов диссертации в опубликованных работах.

Основное содержание диссертационного исследования достаточно полно отражено в 73 работах соискателя, в том числе в 34 работах, опубликованных в журналах, рекомендованных Высшей аттестационной комиссией при Президенте Республики Таджикистан:

В журналах, рекомендованных ВАК:

[1-А]. Косимов, А.А. Цифровой образ “Шахнаме” (“Книги царей”) А.Фирдоуси [Текст]. / З.Д. Усманов, А.А. Косимов // Доклады Академии наук Республики Таджикистан. – 2014. – Том 57. – № 6. – С. 471-476.

[2-А]. Косимов, А.А. Частотность букв таджикской литературы [Текст] / З.Д. Усманов, А.А. Косимов // Доклады Академии наук Республики Таджикистан. – 2015. – Том 58. – № 2. – С. 112-115.

[3-А]. Косимов, А.А. Частотность биграмм в таджикской литературе [Текст] / З.Д. Усманов, А.А. Косимов // Доклады Академии наук Республики Таджикистан. – 2016. – Том 59. – № 1-2. – С. 28-32.

[4-А]. Косимов, А.А. О распознавании авторства таджикского текста [Текст] / З.Д. Усманов, А.А. Косимов // Доклады Академии наук Республики Таджикистан. – 2016. – Том 59. – № 3-4. – С. 114-119.

[5-А]. Косимов, А.А. О множестве анаграмм в поэме А.Фирдауси “Шахнаме”

[Текст] / **А.А. Косимов** // Известия Академии наук Республики Таджикистан. Отделение физико-математических, химических, геологических и технических наук. – 2016. – № 1 (162). – С. 48-53.

[6-А]. **Косимов, А.А.** Оценка эффективности использования униграмм при идентификации текста [Текст] / **А.А. Косимов** // Доклады Академии наук Республики Таджикистан. – 2017. – Т.60. – № 3-4. – С. 132-137.

[7-А]. **Косимов, А.А.** Оценка эффективности использования биграмм при идентификации текста [Текст] / **А.А. Косимов** // Доклады Академии наук Республики Таджикистан. – 2017. – Т.60. – № 5-6. – С. 224-229.

[8-А]. **Косимов, А.А.** Оценка эффективности использования триграмм при идентификации текста [Текст] / **А.А. Косимов** // Известия Академии наук Республики Таджикистан. Отделение физико-математических, химических, геологических и технических наук. – 2017. – №1(166). – С. 51-57.

[9-А]. **Косимов, А.А.** Определение минимального объема выборки слов для идентификации текста [Текст] / **А.А. Косимов** // Вестник Таджикского национального университета, Серия естественных наук, Душанбе. – 2017. – №1/5. – С. 178-180.

[10-А]. **Косимов, А.А.** О минимальном объеме текста, необходимого для распознавания его автора [Текст] / **А.А. Косимов** // Доклады Академии наук Республики Таджикистан. – 2017. – Т.60. – № 9. – С. 398-401.

[11-А]. **Косимов, А.А.** Об идентификации текста с помощью символьных триграмм [Текст] / **А.А. Косимов, О.А. Косимов** // Вестник Технологического Университета Таджикистана, Душанбе. – 2018. – С. 37-42.

[12-А]. **Косимов, А.А.** Программный комплекс Tajik_Text_Author [Текст] / **А.А. Косимов** // Политехнический вестник, Серия: интеллект, инновации, инвестиции. – 2019. – 3(47). – С. 22-28.

[13-А]. **Косимов, А.А.** Применение специфичного цифрового портрета для идентификации авторов произведений [Текст] / **А.А. Косимов, К.С. Бахтеев** // Известия Академии наук Республики Таджикистан. Отделение физико-математических, химических, геологических и технических наук. – 2019. – №3(176). – С. 7-11.

[14-А]. **Косимов, А.А.** О распознавании автора текста на основе частотности слогов [Текст] / **Х.А. Худойбердиев, А.А. Косимов** // Доклады Академии наук Республики Таджикистан. – 2019. – Т.62. – № 11-12. – С. 641-645.

[15-А]. **Косимов, А.А.** О распознавании автора текстового фрагмента [Текст] / **А.А. Косимов, К.С. Бахтеев** // Известия Академии наук Республики Таджикистан. Отделение физико-математических, химических, геологических и технических наук. – 2019. – №4(177). – С. 18-25.

[16-А]. **Косимов, А.А.** К вопросу об автоматическом распознавании авторства и стилей произведений таджикско-персидской художественной литературы [Текст] / **З.Д. Усманов, А.А. Косимов** // Доклады Академии наук Республики Таджикистан. – 2020. – Т.63. – № 1-2. – С. 49-54.

[17-А]. **Косимов, А.А.** О распознавании автора текста на основе частотности длин предложений [Текст] / **А.А. Косимов**, К.С. Бахтеев // Доклады Академии наук Республики Таджикистан. – 2020. – Т.63. – №3-4. – С. 180-186.

[18-А]. **Косимов, А.А.** Автоматический поиск анаграмм словоформных N -грамм [Текст] / З.Д. Усманов, **А.А. Косимов** // Доклады Академии наук Республики Таджикистан. – 2020. – Т.63. – № 5-6. – С. 316-321.

[19-А]. **Косимов, А.А.** О влиянии цифрового портрета текста на распознавание автора произведения [Текст] / З.Д. Усманов, **А.А. Косимов** // Известия Академии наук Республики Таджикистан. Отделение физико-математических, химических, геологических и технических наук. – 2020. – №3(180). – С. 36-42.

[20-А]. **Косимов, А.А.** Об идентификации текста на основе частотности слогов [Текст] / Х.А. Худойбердиев, **А.А. Косимов**, Х.А. Тошхуджаев // Политехнический вестник, Серия: интеллект, инновации, инвестиции. – 2020. – 2(50). – С. 52-56.

[21-А]. **Косимов, А.А.** Об автоматическом распознавании языка произведений [Текст] / З.Д. Усманов, **А.А. Косимов** // Доклады Академии наук Республики Таджикистан. – 2020. – Т.63. – № 7-8. – С. 461-466.

[22-А]. **Косимов, А.А.** Оценка эффективности применения γ -классификатора для атрибуции искусственно сгенерированных поэм «Шахнаме» А. Фирдоуси [Текст] / М.Ё. Мухсинзода, **А.А. Косимов** // Политехнический вестник, Серия: интеллект, инновации, инвестиции. – 2020. – 4(52). – С. 35-39.

[23-А]. **Косимов, А.А.** Тестирование γ -классификатора, настроенного на распознавание языков произведений на основе латинского алфавита [Текст] / З.Д. Усманов, **А.А. Косимов** // Научный Вестник НГТУ «Системы анализа и обработки данных». – Том 82. – № 2. – 2021. – С. 83-94.

[24-А]. **Косимов, А.А.** Об автоматическом распознавании языков произведений на основе кириллического алфавита [Текст] / М.Л. Мирзохасанов, **А.А. Косимов** // Вестник Технологического университета Таджикистана, Душанбе. – 2021. – 1(44). – С. 101-107.

[25-А]. **Косимов, А.А.** Структура однородностей поэм произведения А. Фирдоуси «Шахнаме» [Текст] / **А.А. Косимов**, Н.М. Курбонов // Политехнический вестник, Серия: интеллект, инновации, инвестиции. – 2021. – 2(54). – С. 35-38.

[26-А]. **Косимов, А.А.** Об однородности оригинала и его перевода [Текст] / **А.А. Косимов** // Доклады Национальной академии наук Таджикистана. – 2021. – Т.64. – № 11-12. – С. 660-665.

[27-А]. **Косимов, А.А.** О распознавании автора текстового фрагмента на основе частотности слогов [Текст] / **А.А. Косимов** // Политехнический вестник, Серия: интеллект, инновации, инвестиции. – 2021. – 4(56). – С. 59-64.

[28-А]. **Косимов, А.А.** О распознавании автора текстового фрагмента на основе частотности буквенных биграмм [Текст] / **А.А. Косимов** // Системы анализа и обработки данных. – Том 85. – № 1. – 2022. – С. 73-82. DOI: 10.17212/2782-2001-2022-1-73-82.

[29-А]. **Косимов, А.А.** О распознавании автора текстового фрагмента на основе

частотности буквенных триграмм [Текст] / **А.А. Косимов, Н.А. Шокирова** // Вестник Технологического университета Таджикистана, Душанбе. – 2022. – 2(49). – С. 35-43.

[30-А]. **Косимов, А.А.** О влиянии порядка буквенных униграмм на распознавание автора произведения [Текст] / **А.А. Косимов** // Доклады Национальной академии наук Таджикистана. – 2022. – Т.65. – № 5-6. – С. 324-330.

[31-А]. **Косимов, А.А.** О влиянии порядка буквенных триграмм на распознавание автора произведения [Текст] / **А.А. Косимов** // Вестник Филиала МГУ имени М.В. Ломоносова в городе Душанбе, Серия естественных наук. – 2022. – № 1. – С. 14-21.

[32-А]. **Косимов, А.А.** Определение шифра специальности с помощью символьных униграмм [Текст] / **А.А. Косимов** // Вестник Филиала МГУ имени М.В. Ломоносова в городе Душанбе, Серия естественных наук. – 2023. – Том 1. – №1 (29). – С. 16-24.

[33-А]. **Косимов, А.А.** О влиянии порядка символьных триграмм на определение языка произведения [Текст] / **А.А. Косимов** // Политехнический вестник, Серия: интеллект, инновации, инвестиции. – 2023. – 1(61). – С. 34-37.

[34-А]. **Косимов, А.А.** О влиянии порядка буквенных биграмм на определение языка произведения [Текст] / **И.К. Каландарбеков, А.А. Косимов** // Вестник Филиала МГУ имени М.В. Ломоносова в городе Душанбе, Серия естественных наук. – 2023. – Том 1. – №2 (31). – С. 26-32.

Соответствие содержания диссертации избранной специальности.

Диссертация выполнена в соответствии со следующим разделом Паспорта номенклатуры специальностей научных работников по специальности 05.13.11 – “Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей” (технические науки):

- модели, методы и алгоритмы проектирования и анализа программ и программных систем, их эквивалентных преобразований, верификации и тестирования;
- модели, методы, алгоритмы, языки и программные инструменты для организации взаимодействия программ и программных систем;
- системы управления базами данных и знаний;
- программные системы символьных вычислений;
- человеко-машинные интерфейсы; модели, методы, алгоритмы и программные средства машинной графики, визуализации, обработки изображений, систем виртуальной реальности, мультимедийного общения.

Диссертация Косимова Абдунаби Абдурауфовича на тему «Статистические закономерности распознавания однородности текстов с помощью γ -классификатора» соответствует требованиям Высшей аттестационной комиссии при Президенте Республики Таджикистан, предъявляемым к диссертационным работам, и рекомендуется к защите на соискание ученой степени доктора технических наук по специальности 05.13.11 – “Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей”.

Заключение принято на заседании кафедры «Автоматизированные системы

управления» Таджикского технического университета имени академика М.С. Осими.

ПРИСУТСТВОВАЛИ: Председатель заседания – Умарализода Р.Ш., к.т.н., доцент, зав. кафедрой «Автоматизированные системы управления»; Кабилов Ш.С., к.т.н., и.о. доцент; Холов Ш.Ё., к.т.н., доцент; Джовиди Дж., к.т.н., и.о. доцент; Бадалова М.А., к.т.н., ст. преподаватель; Косимов А.А., к.т.н., и.о. доцент; Бобобеков К.М., к.т.н., ст. преподаватель; Рачабова А.С., ст. преподаватель; Давлатбекова Ш.Д., ассистент.

ПРИГЛАШЕННЫЕ: Каландарбеков И., доктор технических наук, профессор; Нарзуллоев С.А., к.т.н, доцент; Гафуров М.Х., к.т.н, доцент; Назаров А.Ш., к.э.н., доцент, зав. кафедрой «Информационная технология и защита информации»; Юнусов Н.И., к.т.н., доцент; Джалолов У.Х., к.т.н., доцент; Бандишоева Р.М., к.т.н., доцент.

Результаты голосования: «За» – 15 чел.

«Против» – 0 чел.

«Воздержались» – 0 чел.

Протокол № 1 от «11» 09 2023 г.



Умарализода Р.Ш. – к.т.н., доцент, зав. кафедрой
Автоматизированные системы управления



Рачабова А.С. – секретарь заседания