



УТВЕРЖДАЮ

Проректор по научной работе

ФГБОУ ВО «ПГТУ»

док-р физ.-мат. наук, доцент

Д.В. Иванов

03

2024 г.

## ЗАКЛЮЧЕНИЕ

федерального государственного бюджетного образовательного учреждения  
высшего образования  
«Поволжский государственный технологический университет»

**Диссертация** «Метод и алгоритм обнаружения спама на основе выделения признаков электронных писем с использованием контентной фильтрации» выполнена на кафедре информационной безопасности Федерального государственного бюджетного образовательного учреждения высшего образования (ФГБОУ ВО) «Поволжский государственный технологический университет».

**В период подготовки диссертации соискатель** Корелов Сергей Викторович проходил обучение в аспирантуре ФГБОУ ВО «Поволжский государственный технологический университет».

**В 2001 г. окончил** Академию Федерального агентства правительственной связи и информации при Президенте Российской Федерации по специальности 220200 «Автоматизированные системы обработки информации и управления».

**В 2024 году окончил** аспирантуру ФГБОУ ВО «Поволжский государственный технологический университет» по направлению 10.06.01 «Информационная безопасность», направленность «Методы и системы защиты информации, информационная безопасность».

**Справки** со сведениями о сдаче кандидатских экзаменов выданы:

- 26.05.2021 г. ФГБОУ ВО «Национальный исследовательский Нижегородский государственный университет им. Н.И. Лобачевского», кандидатский экзамен «Английский язык» сдан 11.06.2003 г.;

- 26.05.2021 г. ФГБОУ ВО «Национальный исследовательский Нижегородский государственный университет им. Н.И. Лобачевского», кандидатский экзамен «История и философия науки (технические науки)» сдан 19.05.2021 г.;

- 30.10.2023 г. ФГБОУ ВО «Поволжский государственный технологический университет», кандидатский экзамен «05.13.19 Методы и системы защиты информации, информационная безопасность» сдан 26.03.2021 г.

**Научный руководитель** – Сидоркина Ирина Геннадьевна, д.т.н., профессор, зав. кафедры информационной безопасности ФГБОУ ВО «Поволжский государственный технологический университет».

**По итогам обсуждения принято следующее заключение:**

1. Диссертация Корелова Сергея Викторовича является законченной научно-квалификационной работой по научной специальности 2.3.6. Методы и системы защиты информации, информационная безопасность, соответствующей п. 9 Положения о присуждении ученых степеней, утвержденного Постановлением Правительства Российской Федерации № 842 от 24.09.2013 г. (с последующими изменениями), в которой содержатся научно обоснованные результаты решения задачи обнаружения спама, имеющие важное практическое значение для повышения эффективности обнаружения спама в сфере разработки, совершенствования и применения методов и средств защиты информации.

**2. Соискателем лично получены все основные результаты, выносимые на защиту:**

1. Разработана модель электронного почтового сообщения для классификации электронных писем на основе метода «генетических карт», отличающаяся от известных моделей методом выделения значимых

последовательностей символов текста (признаков электронных писем на основе их содержания, термов), позволяющим усилить смысловое содержание термов.

2. Разработан метод классификации электронных писем для обнаружения спама и идентификации легальных электронных писем, отличающийся от известных использованием разработанной модели электронных писем, применение которого позволяет повысить эффективность обнаружения спама и достоверность идентификации легальных электронных писем с учетом меняющихся информационных потребностей конкретного пользователя (персонализации), а также снизить количество неклассифицированных писем.

3. Разработан алгоритм классификации электронных писем, отличающийся от известных наличием дополнительной процедуры определения «схожести» термов на основе расстояния Левенштейна, обеспечивающей вычисление мер принадлежности классифицируемого электронного письма к классам спама и легальных, применение которого позволяет повысить достоверность идентификации легальных электронных писем с учетом меняющихся информационных потребностей конкретного пользователя (персонализации), а также уменьшить количество неклассифицируемых писем.

4. Разработана архитектура подсистемы классификации электронных писем для обнаружения спама и идентификации легальных электронных писем, отличающаяся от известных блоком выделения термов и блоком нечеткой классификации, реализующая предложенные в работе метод и алгоритм, применение которых позволяет повысить достоверность идентификации легальных электронных писем с учетом меняющихся информационных потребностей конкретного пользователя (персонализации).

В перечисленных в автореферате работах соискателем лично получены следующие результаты:

- в работах [2-5, 8-16, 18] проведены анализ, исследование и обоснование возможности применения метода «генетических карт» для представления электронных писем; разработана модель электронного почтового сообщения

для классификации электронных писем на основе указанного метода и проведены анализ, исследование и обоснование ее элементов и ключевых параметров;

- в работах [16, 18, 19] проведено исследование применения различных весов термов в задаче обнаружения спама с использованием разработанной модели электронных писем, а также разработаны процесс (метод) классификации электронных писем и алгоритм его реализации в задаче обнаружения спама;

- в работе [17] представлена архитектура прототипа подсистемы классификации электронных писем на основе разработанной модели электронных писем.

Опубликованные работы полностью отражают основное содержание диссертационной работы. Основные положения и результаты, выносимые на защиту, отражены в публикациях автора: по главе 1 – [1, 3, 6, 7]; по главе 2 – [2-5, 8-15]; по главе 3 – [16, 18, 19]; по главе 4 – [17-19]. Одна работа написана автором единолично, другие совместно с научным руководителем или другими членами научного коллектива.

**3. Достоверность полученных результатов и выводов** основана на том, что предложенные в диссертационной работе решения подтверждаются:

- корректной постановкой задач;
- применением известных технологий и методов, успешно используемых в других прикладных областях;
- апробацией разработанных модели, метода, алгоритма и программного модуля.

Выводы и положения диссертации научно обоснованы и подтверждены положительными оценками на научных конференциях и результатами экспериментальных исследований автора.

#### **4. Научная новизна работы заключается в следующем:**

1. Разработана модель электронного почтового сообщения для классификации электронных писем на основе метода «генетических карт», отличающаяся от известных моделей методом выделения значимых последовательностей символов текста (признаков электронных писем на основе их содержания, термов), позволяющим усилить смысловое содержание термов.

2. Разработан метод классификации электронных писем для обнаружения спама и идентификации легальных электронных писем, основанный на положениях задачи классификации текстовых документов, отличающийся использованием разработанной модели электронных писем, применение которого позволяет повысить эффективность обнаружения спама и достоверность идентификации легальных электронных писем, а также снизить количество неклассифицированных писем.

3. Разработан алгоритм классификации электронных писем на основе методов машинного обучения, отличающийся наличием дополнительной процедуры определения «схожести» термов на основе расстояния Левенштейна, обеспечивающей вычисление мер принадлежности классифицируемого электронного письма к классам спама и легальных для повышения достоверности идентификации электронных писем, позволяющий осуществить программную реализацию разработанных модели и метода.

4. Разработана архитектура подсистемы классификации электронных писем для обнаружения спама и идентификации легальных электронных писем на основе разработанного алгоритма, отличающаяся от известных блоком выделения термов и блоком нечеткой классификации, реализующая предложенные в работе метод и алгоритм, применение которых позволяет повысить достоверность идентификации легальных электронных писем с учетом меняющихся информационных потребностей конкретного пользователя (персонализации).

**5. Практическая значимость полученных результатов заключается в следующем:**

Практическая значимость полученных результатов заключается в разработке программных модулей исследовательского прототипа подсистемы классификации электронных писем для обнаружения спама и идентификации легальных электронных писем. Применение разработанных модели и метода позволяет повысить эффективность обнаружения спама и достоверность идентификации легальных электронных писем с учетом меняющихся информационных потребностей конкретного пользователя (персонализации) с точностью классификации до 0,995 и полнотой классификации до 0,993, а также снизить количество ошибочно классифицированных и неклассифицированных писем.

**6. Ценность научной работы заключается в том, что в результате выполненных исследований:**

- предложена новая модель электронного почтового сообщения для классификации электронных писем с использованием метода выделения значимых последовательностей символов текста (признаков электронных писем на основе их содержания), позволяющего усилить смысловое содержание термов;

- предложен метод классификации электронных писем для обнаружения спама и идентификации легальных электронных писем с использованием разработанной модели;

- предложен алгоритм классификации электронных писем с использованием методов машинного обучения и дополнительной процедуры определения «схожести» термов на основе расстояния Левенштейна, обеспечивающей вычисление мер принадлежности классифицируемого электронного письма к классам спама и легальных;

- предложена архитектура подсистемы классификации электронных писем для обнаружения спама и идентификации легальных электронных писем на основе разработанного алгоритма.

## **7. Обоснование выбранной специальности и отрасли науки диссертации**

Диссертация «Метод и алгоритм обнаружения спама на основе выделения признаков электронных писем с использованием контентной фильтрации» соответствуют следующим пунктам паспорта научной специальности 2.3.6. «Методы и системы защиты информации, информационная безопасность»:

п. 3 «Методы, модели и средства выявления, идентификации, классификации и анализа угроз нарушения информационной безопасности объектов различного вида и класса» – разработаны модель электронного почтового сообщения, учитывающая содержание электронных писем конкретного пользователя (персонализацию), метод классификации электронных писем для обнаружения спама и идентификации легальных электронных писем, алгоритм классификации электронных писем, архитектура подсистемы классификации электронных писем для обнаружения спама и идентификации легальных электронных писем;

п. 5 «Методы, модели и средства (комплексы средств) противодействия угрозам нарушения информационной безопасности в открытых компьютерных сетях, включая Интернет» – разработаны модель электронного почтового сообщения, учитывающая содержание электронных писем конкретного пользователя (персонализацию), метод классификации электронных писем для обнаружения спама и идентификации легальных электронных писем, алгоритм классификации электронных писем, архитектура подсистемы классификации электронных писем для обнаружения спама и идентификации легальных электронных писем;

п. 15 «Принципы и решения (технические, математические, организационные и др.) по созданию новых и совершенствованию существующих средств защиты информации и обеспечения информационной безопасности» – разработаны модель электронного почтового сообщения, учитывающая содержание электронных писем конкретного пользователя (персонализацию), метод классификации электронных писем для обнаружения спама и идентификации легальных электронных писем, алгоритм классификации электронных писем, архитектура подсистемы классификации электронных писем для обнаружения спама и идентификации легальных электронных писем.

**Отрасль науки** – технические науки, поскольку приведенные результаты исследований в области обнаружения спам дают существенный технический эффект при их использовании и внедрении.

## **8. Полнота изложения материалов диссертации**

Основные результаты диссертации опубликованы в 19 печатных работ, в том числе в 4 статьях в научных изданиях из Перечня рецензируемых научных изданий, рекомендованных ВАК, в 15 статьях в других изданиях.

### *Статьи в научных изданиях из Перечня рецензируемых научных изданий, рекомендованных ВАК*

1. Корелов С. В., Ротков Л. Ю., Рябов А. А. Вероятностный метод идентификации спама // Доклады Томского государственного университета систем управления и радиоэлектроники. 2010. № 1 (21), часть 1. С. 150-152.

2. Корелов С. В., Петров А. М., Ротков Л. Ю., Горбунов А. А. Модель электронных писем в задаче обнаружения спама // Вестник Поволжского государственного технологического университета. Сер.: Радиотехнические и инфокоммуникационные системы. 2020. № 2 (46). С. 44-54. DOI:10.25686/2306-2819.2020.2.44.



3. Корелов С. В., Петров А. М., Ротков Л. Ю., Горбунов А. А. Предобработка текстов электронных писем в задаче обнаружения спама // Труды учебных заведений связи. 2020. Т. 6. № 4. С. 80-90. DOI:10.31854/1813-324X-2020-6-4-80-90.

4. Корелов С. В., Петров А. М., Сидоркина И. Г., Ротков Л. Ю., Горбунов А. А. Выбор размера кодовой таблицы в модели электронных писем // Вестник Поволжского государственного технологического университета. Сер.: Радиотехнические и инфокоммуникационные системы. 2021. № 3 (51). С. 49-62. DOI:10.25686/2306-2819.2021.3.49.

#### *Другие публикации по теме диссертации*

5. Корелов С. В., Крюков А. К., Ротков Л. Ю. Применение метода построения генетической карты текста для идентификации спама // Труды XII научной конференции по радиофизике, посвященной 90-летию со дня рождения М. М. Кобрин (Нижний Новгород, 7 мая 2008 г.) /Под ред. А. В. Якимова, С. М. Грача. Нижний Новгород: Изд-во ТАЛАМ, 2008. С. 277-278.

6. Агаджанов В. В., Корелов С. В., Ротков Л. Ю. Обнаружение спама при помощи аппарата wavelet-преобразований // Труды XII научной конференции по радиофизике, посвященной 90-летию со дня рождения М. М. Кобрин (Нижний Новгород, 7 мая 2008 г.) /Под ред. А. В. Якимова, С. М. Грача. Нижний Новгород: Изд-во ТАЛАМ, 2008. С. 276-277.

7. Корелов С. В., Грачева О. К. Идентификация спама на классах сообщений // Труды XIV научной конференции по радиофизике, посвященной 80-й годовщине со дня рождения Ю. Н. Бабанова (Нижний Новгород, 7 мая 2010 г.) /Под ред. С. М. Грача, А. В. Якимова. Нижний Новгород: ННГУ, 2010. С. 288-289.

8. Корелов С. В., Ротков Л. Ю. Метод генетических карт в задаче идентификации спама // Информационно-измерительные и управляющие системы. 2011. № 3. Т. 9. С. 72-75.

9. Корелов С. В. Обнаружение текстового спама методом генетических карт // Труды XV научной конференции по радиофизике, посвященной 110-й годовщине со дня рождения А. А. Андропова (Нижний Новгород, 1-13 мая 2011 г.) /Под ред. С. М. Грача, А. В. Якимова. Нижний Новгород: ННГУ, 2011. С. 265-267.

10. Корелов С. В., Ротков Л. Ю. Идентификация текстового спама методом генетических карт // Вестник Нижегородского университета им. Н. И. Лобачевского. 2012. № 4 (1). С. 101-104.

11. Корелов С. В., Петров А. М., Ротков Л. Ю., Горбунов А. А. Определение длины выборки в модели электронных писем // Вестник Пермского национального исследовательского политехнического университета. Электротехника, информационные технологии, системы управления. 2020. № 4 (36). С. 31-47. DOI:10.15593/2224-9397/2020.4.02.

12. Корелов С. В., Петров А. М., Ротков Л. Ю., Горбунов А. А. К вопросу об определении численного значения параметра в модели электронных писем // Труды XXIV научной конференции по радиофизике, посвященной 75-летию радиофизического факультета (Нижний Новгород, 13 – 31 мая 2020 г.). Нижний Новгород: ННГУ, 2020. С. 471-474.

13. Корелов С. В., Петров А. М., Ротков Л. Ю., Горбунов А. А. К вопросу об определении численного значения параметра модели электронных писем // Материалы всероссийской научно-технической конференции «Автоматизированные системы управления и информационные технологии» (г. Пермь, 9–11 июня 2020 г.). 2020. Т.2. С. 519-525.

14. Корелов С. В., Петров А. М., Ротков Л. Ю., Горбунов А. А. Комбинирование значений параметра модели электронных писем // Материалы XII Международной интернет-конференции молодых ученых, аспирантов, студентов «Инновационные технологии: теория, инструменты, практика» (г. Пермь, 16 ноября – 31 декабря 2020 г.). 2020. С. 448-455.

15. Корелов С. В., Петров А. М., Сидоркина И. Г., Горбунов А. А. Анализ результатов реализации подхода к выделению термов в модели электронных

писем на случайность // Труды XXV научной конференции по радиофизике, (Нижний Новгород, 14 – 26 мая 2021 г.). Нижний Новгород: ННГУ, 2021. С. 498-502.

16. Корелов С. В., Петров А. М., Сидоркина И. Г., Ротков Л. Ю. Применение весов термов в задаче обнаружения спама с использованием модели электронных писем // Труды XXVI научной конференции по радиофизике, посвященной 120-летию М. Т. Греховой, (Нижний Новгород, 12 – 27 мая 2022 г.). Нижний Новгород: ННГУ, 2022. С. 522-526.

17. Корелов С. В., Петров А. М., Сидоркина И. Г., Ротков Л. Ю. Подсистема классификации электронных писем на основе модели электронных писем // Информационные технологии обеспечения комплексной безопасности в цифровом обществе: сборник материалов VI Всероссийской молодежной научно-практической конференции с международным участием (г. Уфа, 19-20 мая 2023 года) / отв.ред. Д. С. Юнусова – Уфа: РИЦ УУНиТ, 2023. – С. 31-35.

18. Корелов С. В., Петров А. М., Сидоркина И. Г., Ротков Л. Ю. Применение весов термов в задаче обнаружения спама // Труды XXVII научной конференции по радиофизике (Нижний Новгород, 15 – 25 мая 2023 г.). Нижний Новгород: ННГУ, 2023. С. 516-521.

19. Корелов С. В., Петров А. М., Сидоркина И. Г., Ротков Л. Ю. Модель процесса классификации электронных писем и алгоритм его реализации в задаче обнаружения спама // Труды Международного научно-технического конгресса «Интеллектуальные системы и информационные технологии – 2023» («ИС & ИТ-2023», «IS&IT'23»). Научное издание в 2-х т. Т. 2. – Таганрог: Изд-во Ступина С. А., 2023. С. 3-9.

**Диссертация** Корелова Сергея Викторовича соответствует п. 14 Положения о порядке присуждении ученых степеней:

- отсутствуют недостоверные сведения об опубликованных соискателем ученой степени работах, в которых изложены основные научные результаты диссертации;

- соискатель ссылается на авторов и источники заимствования.

Диссертация «Метод и алгоритм обнаружения спама на основе выделения признаков электронных писем с использованием контентной фильтрации» Корелова Сергея Викторовича рекомендуется к защите на соискание ученой степени кандидата технических наук по научной специальности 2.3.6. Методы и системы защиты информации, информационная безопасность.

**Заключение принято на заседании** кафедры информационной безопасности ФГБОУ ВО «Поволжский государственный технологический университет».

**Присутствовало на заседании** 21 человек, в том числе 3 доктора наук.

**Результаты голосования:** «за» – 21 человек, «против» – нет, «воздержалось» – нет.

Протокол № 11 от « 15 » 03 2024 г.

Декан ФИиВТ,  
доцент кафедры ИБ

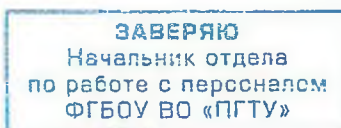


А.А. Кречетов

Секретарь расширенного  
заседания кафедры ИБ



Е.С. Кубашева



Исаева С.А.

15 03 2024