В диссертационный совет 24.2.415.04
на базе ФГАОУ ВО ТУСУР
634050, Россия, г. Томск, пр. Ленина, 40

ОТЗЫВ

на автореферат диссертации Романова Александра Сергеевича

«Методология идентификации автора текстовой информации для решения задач кибербезопасности», представленную на соискание ученой степени доктора технических наук по специальности 2.3.6 – Методы и системы защиты информации, информационная безопасность

Диссертация Романова Александра Сергеевича «Методология идентификации автора текстовой информации для решения задач кибербезопасности» посвящена созданию методологии идентификации автора текстовой информации, включая естественно-языковые тексты и исходные коды программ, для решения задач информационной безопасности. **Актуальность** представленного диссертационного исследования определяется возрастающей потребностью в создании высокоточных и производительных инструментов автоматического анализа текстовой информации и установления авторства в контексте обеспечения информационной безопасности. **Практическая значимость** работы определяется развитием технологий искусственного интеллекта в области агрегирования методов обработки естественного языка и машинного обучения и созданных новых более точных и устойчивых методов, которые позволяют анализировать не только стилистические, но и индивидуальные семантические авторские характеристики. Разработанный комплексный подход, объединяющий идентификацию авторов текстов с определением их «портретов», представляет значительный практический вклад и соответствует мировым тенденциям развития в данной области науки. Построенная методология позволяет решать критически важные задачи информационной безопасности: от противодействия киберпреступности, экстремизму и деструктивным воздействиям через распространение текстовой информации на общество и государство до подтверждения авторства учебных, научных и квалификационных работ.

Научная новизна результатов диссертационного исследования заключается в разработке комплекса взаимосвязанных методик анализа текстовых данных. Соискателем впервые предложена методология идентификации авторства, учитывающая специфику как естественных, так и искусственно сгенерированных текстов; разработана новая модель создания текста в киберсреде, интегрирующая семантические иерархические признаки, атрибуты авторов и виды деятельности; впервые созданы и верифицированы специализированные методики на основе ансамблевых архитектур нейросетей (GRU-CNN, CodeBERT, BERT, SimNN) и методов машинного обучения (SVM, fastText) с применением новых методов отбора признаков и трансферного обучения, решающие задачи бинарной и мультиклассовой классификации авторов по разным атрибутивным признакам (возрастным

группам, идеологии, гендеру и пр.), идентификации авторов с учетом возможной обфускации текстов, определения деструктивного контента, а также проверки однородности текста и выявления заимствований в условиях открытого множества авторов, что в совокупности представляет собой качественный вклад в теорию и практику информационной безопасности. Предложенные соискателем методы эффективно применены для выявления авторства текстов в различных прикладных сценариях, таких как мониторинг социальных сетей, проверка подлинности научных работ, защита интеллектуальной собственности, борьба с распространением запрещенной информации.

Результаты работы подкреплены обширными экспериментальными исследованиями, выполненных соискателем. **Обоснованность и достоверность** научных положений, выводов и рекомендаций подтверждаются использованием теоретически обоснованных методов исследования, соответствием экспериментальным данным, положительным опытом внедрения предложенных методик. Результаты исследования представлены диссертантом в высокорейтинговых отечественных и международных научных изданиях, общее количество опубликованных работ – 94, имеются зарегистрированные результаты интеллектуальной деятельности.

При ознакомлении с авторефератом возникли следующие **замечания**:

1) в автореферате на рис. 4 представлено, но не раскрыто, каким образом учитываются в методиках и гиперпараметрах моделей машинного обучения особенности и ограничения среды, в которой выполняется идентификация авторов текстов (например, динамичность состава авторов, изменчивость их «портретов» со временем, наличие преднамеренной обфускации самим автором, инструментальные ограничения на объем и разнородность доступных для анализа наборов текстов);

2) видится недостаточным методологическое обоснование представленного в автореферате выбора определенных комбинаций нейромоделей и методов машинного обучения (например, GRU+CNN+SVM) для построения ансамблей. Не ясно, насколько статистически значимым является продемонстрированный эффект от их интеграции и насколько он действительно превосходит монолитные (неансамблевые) архитектуры, особенно в свете избыточности и малой различительной силы отдельных компонентов ансамблей. Например, указание на обучение модели на трех образцах (на стр. 21) не позволяет репрезентативно оценить устойчивость полученных результатов и возможность обобщения на реальные трудоемкие сценарии анализа текста от тысяч уникальных авторов;

3) не представлен критический анализ разработанных методик в условиях, когда зафиксирована чрезмерная длительность выполнения процедуры идентификации автора (например, согласно таблице 1, в отдельных сценариях длительность анализа достигает 9…27ч для 50 авторов, при этом точность идентификации снижается до 37%), что затрудняет оперативный анализ текстов в реальном времени. В каких условиях данные временные задержки и снижение точности могут быть оправданы спецификой решаемых задач, и как они могут быть устранены?

4) ряд ключевых методов (например, трансферное обучение, семантическая кластеризация, VGG-Face, ANOVA и пр.) вводятся исключительно в контексте частных сценариев в главе 5, но не участвуют и не имеют теоретического обоснования в рамках представленной соискателем общей методологии в главе 3, что затрудняет комплексную оценку корректности, внутренней непротиворечивости, границ применимости предложенного решения;

5) имеются неточности в изложении материала автореферата, например, «state-of-the-art алгоритмы» (на стр. 13), «множество авторов ЫІ» (на стр. 15).

Заключение

Насколько можно судить по автореферату, представленная работа **является** законченным квалификационным исследованием, выполненным на актуальную тему и содержащим научные и практические результаты для решения задач информационной безопасности на базе методологии идентификации автора текстовой информации.

Диссертационное исследование Романова Александра Сергеевича на тему «Методология идентификации автора текстовой информации для решения задач кибербезопасности» **соответствует** критериям, установленным в Положении о присуждении ученых степеней (пп. 9-14), предъявляемым к докторским диссертациям. В работе изложены новые научно обоснованные технические решения, внедрение которых вносит значительный вклад в развитие страны.

Романов Александр Сергеевич **заслуживает** присуждения ученой степени доктора технических наук по специальности 2.3.6 – Методы и системы защиты информации, информационная безопасность.

Я, Калинин Максим Олегович, даю согласие на включение моих персональных данных в документах, связанных с работой диссертационного совета, и их дальнейшую обработку.

Профессор Высшей школы кибербезопасности Института компьютерных наук и кибербезопасности федерального государственного автономного образовательного учреждения высшего образования «Санкт-Петербургский политехнический университет Петра Великого», доктор технических наук, профессор, специальность 05.13.19 Методы и системы защиты информации, информационная безопасность

Калинин Максим Олегович
«17» 09 2025 г.

федеральное государственное автономное образовательное учреждение высшего образования «Санкт-Петербургский политехнический университет Петра Великого»
195251, г. Санкт-Петербург, ул. Политехническая, д. 29
тел.: +78125527632, e-mail: sci@ibks.spbstu.ru

Подпись Калинина М.О.
УДОСТОВЕРЯЮ
Ведущий специалист
по кадрам
17.09.2025