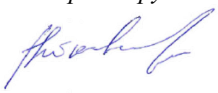
На правах рукописи



Виткова Лидия Андреевна

МОДЕЛИ, АЛГОРИТМЫ И МЕТОДИКА ПРОТИВОДЕЙСТВИЯ ВРЕДОНОСНОЙ ИНФОРМАЦИИ В СОЦИАЛЬНЫХ СЕТЯХ

Специальность: 05.13.19 – Методы и системы защиты информации, информационная безопасность

АВТОРЕФЕРАТ

диссертации на соискание ученой степени
кандидата технических наук

Санкт-Петербург – 2021

Работа выполнена в Федеральном государственном бюджетном учреждении науки «Санкт-Петербургский Федеральный исследовательский центр Российской академии наук» (СПб ФИЦ РАН) в лаборатории проблем компьютерной безопасности

Научный руководитель: **Сахаров Дмитрий Владимирович**
кандидат технических наук,
СПбГУТ им. проф. М.А. Бонч-Бруевича,
доцент кафедры Защищенных систем связи

Официальные оппоненты: **Липатников Валерий Алексеевич**
доктор технических наук, профессор
Военная академия связи имени Маршала Советского Союза С.М. Буденного Министерства обороны РФ, научно-исследовательский центр, старший научный сотрудник

Филяк Петр Юрьевич
кандидат технических наук, доцент
ФГБОУ ВО СГУ им. Питирима Сорокина
Институт точных наук и информационных технологий, кафедра информационной безопасности, доцент

Ведущая организация: **Акционерное общество «Лаборатория противодействия промышленному шпионажу»**

Защита состоится «29» июня 2021 г. в 14 часов 00 минут на заседании диссертационного совета Д 002.199.01 при Федеральном государственном бюджетном учреждении науки «Санкт-Петербургский Федеральный исследовательский центр Российской академии наук» (СПб ФИЦ РАН) по адресу: 199178, Санкт-Петербург, 14-я линия В.О., 39, комн. 401. Факс: (812)-328-44-50, тел: (812)-328-34-11.

С диссертацией можно ознакомиться в библиотеке Федерального государственного бюджетного учреждения науки «Санкт-Петербургский Федеральный исследовательский центр Российской академии наук» (СПб ФИЦ РАН) по адресу: 199178, Санкт-Петербург, В.О., 14 линия, д. 39 и на сайте http://www.spiiras.nw.ru/dissovet/

Автореферат разослан «20» мая 2021 года.

Ученый секретарь
диссертационного совета Д 002.199.01,
кандидат технических наук Абрамов Максим Викторович

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы исследования. Глубина проникновения социальных сетей в повседневную жизнь значительна, и их преимуществом является возможность участников коммуникации оперативно высказывать свое мнение большой группе людей, публиковать медиа файлы. Сегодня социальные сети (СС) являются не только средством общения, но и инструментом распространения информации. Очевидной проблемой информационной безопасности современного общества стала вредоносная информация. Стоит отметить, что террористические и преступные группировки все чаще берут на вооружение средства информационного воздействия, пишут стратегии, направленные на расширение сферы влияния и вовлечение новых адептов через СС. Именно поэтому одной из составляющих обеспечения информационной безопасности государства представляется мониторинг, анализ и активное противодействие вредоносной информации в СС.

Само понятие «вредоносная информация» рассматривается экспертами разных наук, но консенсус пока не достигнут. Изучением вопросов противодействия распространению вредоносной информации стали заниматься еще в 1990 г. Так, например, В.Н. Лопатин входил в состав парламентской комиссии Верховного Совета СССР и отвечал за вопросы информационной безопасности. В своих работах он определял «вредоносную информацию» как угрозу информационной безопасности. К этому он относил: распространение порнографии; клевету; недостоверную информацию, скрытую рекламу. В 21 веке к вредоносной информации все же чаще относят «фейковые новости». Особенно остро необходимость противодействия распространению таким новостям в СС, порождающим волны паники, возникла во время пандемии короновирусной инфекции.

Однако, в настоящее время проблема противодействия имеет крайне малое количество научно-технических решений. Известные средства обнаружения и противодействия вредоносной информации в СС не отвечают требованиям к скорости, полноте, точности и адекватности принимаемых решений. Это обусловлено несколькими причинами. Во-первых, системы разделены два не связанных модуля: (1) мониторинг; (2) противодействие. Посередине между ними находится оператор. Во-вторых, СС имеют сложную структуру и состоят из множества разнородных сообщений, что недостаточно учитывается при выборе цели противодействия, например тип сообщения, источник и другие характеристики. В-третьих, в реальном масштабе времени необходимо обрабатывать сверхбольшие объемы сообщений и в сжатые сроки выбирать цель для контрмеры, в ручном режиме оператор системы противодействия не в состоянии остановить распространение вредоносной информации.

Таким образом, основная сложность противодействия вредоносной информации в СС напрямую следует из современных тенденций развития информационной сферы, а именно, увеличения: (1) объема сообщений, содержащих вредоносную информацию; (2) скорости распространения вредоносной информации; (3) скорости тиражирования сообщений; (4) скорости появления новых источников распространения информации в СС; (5) количества способов привлечения внимания

аудитории; (6) уровня гетерогенности данных. Это обуславливает необходимость повышения эффективности противодействия вредоносной информации в социальных сетях, в том числе за счет его оперативности и обоснованности.

Таким образом, решаемая в диссертационной работе **научная задача**, заключающаяся в разработке моделей, алгоритмов и методики противодействия вредоносной информации в социальной сети, является крайне актуальной.

Степень разработанности темы.

Большое внимание вопросам противодействия вредоносной информации, анализу и оценке источников вредоносной информации в СС уделяется такими исследователями как Д.А. Губанов, И.В. Котенко, М.В. Литвиненко, Д.А. Новиков, И.Б. Саенко, А.Л. Тулупьев, Д.Ю. Турдаков, А.А. Чечулин, А.Г. Чхартишвили, A.L. Barabasi, X. Zheng и др. Множество работ посвящено информационному конфликту и противоборству. К этой группе можно отнести труды С.А. Будникова, Ю.Л. Козирацкого, В.А. Липатникова, С.И. Макаренко, С.П. Расторгуева, Д.В. Сахарова. Вопросы информатизации процессов и оценивания эффективности информационных систем раскрываются в работах М.В. Буйневича, В.П. Заболотского, А.А. Мусаева, П.Ю. Филяка, Р.М. Юсупова. Однако, несмотря на сделанный учеными существенный задел, проблема противодействия вредоносной информации в СС не может считаться разрешенной и требует проведения новых исследований.

Цели и задачи. Целью диссертационной работы является повышение эффективности противодействия вредоносной информации в СС за счет анализа источников вредоносной информации и автоматизации выбора контрмер. Для достижения данной цели в диссертационной работе поставлены и решены следующие задачи:

1) анализ существующих моделей вредоносной информации и информационного обмена;

2) анализ существующих алгоритмов оценки источников в СС, существующих систем мониторинга и методик противодействия вредоносной информации в СС;

3) разработка комплекса моделей социальной сети, источника и вредоносной информации;

4) разработка комплекса алгоритмов анализа источников вредоносной информации и ранжирования контрмер;

5) разработка методики противодействия вредоносной информации в социальных сетях;

6) разработка архитектуры и программных прототипов компонентов системы противодействия (СПД) вредоносной информации, экспериментальная и теоретическая оценка эффективности предложенных моделей, алгоритмов, методики и архитектуры.

Объектом исследования являются социальные сети, в которых возможно наличие сообщений с вредоносной информацией и их источников.

Предметом исследования являются модели, методики и алгоритмы противодействия вредоносной информации в СС.

Научная новизна результатов диссертационной работы состоит в следующем:

1. Комплекс моделей социальной сети, источника и вредоносной информации отличается от аналогов учетом структуры информационного обмена в СС, информационных объектов и вредоносной информации с использованием предложенной классификации объектов социальной сети. Разработанные модели социальной сети и источника содержат новые классы, атрибуты объектов и связи, а модель вредоносной информации, в отличие от аналогов, основана на теории множеств и состоит из взаимосвязанных объектов и признаков вредоносной информации, вместе формирующих вредоносно-информационные объекты. Также в комплекс входит авторская информационно-признаковая модель вредоносной информации.

2. Комплекс алгоритмов анализа источников вредоносной информации и ранжирования контрмер, отличается от аналогов учетом связей и зависимых атрибутов объектов в социальной сети, а также учетом таких атрибутов как потенциал источника, активность пользователей на странице источника, количество просмотров сообщения с вредоносной информацией, количество друзей и подписчиков источника. В качестве результата работы алгоритмы анализа источников формируют ранжированный список объектов воздействия. Алгоритм ранжирования контрмер отличается от аналогов учетом авторских коэффициентов и уровней сложности для каждой меры противодействия в системе и в качестве результата работы формирует ранжированный список контрмер.

3. Методика противодействия вредоносной информации в СС отличается от известных тем, что она ориентирована на автоматический и автоматизированный выбор объектов воздействия и мер противодействия вредоносной информации из списка ранжированных контрмер. Кроме того, методика отличается от аналогов использованием предложенных моделей представления социальной сети, источника, вредоносной информации, а также предложенных алгоритмов анализа источников и ранжирования контрмер.

4. Архитектура и программные прототипы компонентов СПД вредоносной информации отличаются от известных тем, что ориентированы на ранжирование и выбор доступных контрмер в системе для заданных типов вредоносной информации. Архитектура содержит оригинальные компоненты анализа и оценки источника вредоносной информации, базу данных с информацией о мерах противодействия вредоносной информации в СС, информацию об агентах реализации, через которые контрмеры будут реализованы. В силу этого архитектура позволяет формировать наборы исходных данных для исследований и разработок в области противодействия вредоносной информации, а также для исследований и разработок решений для систем поддержки принятия решения.

Теоретическая и практическая значимость работы. Теоретическая значимость диссертационной работы определяется ее вкладом в развитие теории и методов информационной безопасности, что проявляется в следующих аспектах: введены новые классы, объекты, их атрибуты и связи для анализа и оценки информационных объектов и информационного обмена в СС, расширен класс алгоритмов сортировки и ранжирования для анализа источников в СС, расширен набор критериев для мер противодействия и выделены новые функциональные связи между компонентами архитектуры СПД. Предложенный подход позволяет

формулировать научно-обоснованные требования к решению задач, связанных с анализом источника вредоносной информации в СС и с противодействием сообщению или его источнику. Кроме того, предложенные комплексы моделей и алгоритмов, методика и архитектура могут быть использованы как часть системы поддержки принятия решений оператором в интересах противодействия вредоносной информации.

Методология и методы исследования. Для решения поставленных в диссертации задач применялись как классические, так и относительно молодые методы исследования:

1) системный и сравнительный анализ применен в равной степени для получения практически всех основных научных результатов;

2) сбор, систематизация и анализ научной и технической информации о предметной области позволили создать комплекс моделей;

3) объектно-ориентированный подход и структурный синтез использовался для создания алгоритмов анализа и оценки источников;

4) с помощью методов ранжирования, экспертной оценки и методов проектирования архитектур и программных систем были созданы методика противодействия, архитектура и программные компоненты системы противодействия вредоносной информации в социальных сетях.

Положения, выносимые на защиту. Основными положениями, выносимыми на защиту, являются:

1) комплекс моделей социальной сети, источника и вредоносной информации;

2) комплекс алгоритмов анализа источников вредоносной информации и ранжирования контрмер;

3) методика противодействия вредоносной информации в социальной сети;

4) архитектура и программные компоненты системы противодействия вредоносной информации.

Степень достоверности и апробация результатов работы. Достаточная степень достоверности и обоснованность представленных в диссертационной работе научных положений обеспечиваются за счет тщательного анализа состояния исследований в данной области, подтверждается согласованностью результатов, полученных при экспериментальных исследованиях, успешной апробацией основных теоретических положений диссертации на ряде научных конференций всероссийского и международного уровня, а также публикацией основных положений, раскрывающих данные результаты, в ведущих рецензируемых научных изданиях.

Основные положения и результаты диссертационной работы были представлены на следующих научных конференциях: МНТ НТК «Актуальные проблемы инфотелекоммуникаций в науке и образовании» (АПИНО 2018, 2019, 2020) (Санкт-Петербург, Россия); 10-я Конференция по социальной информатике (SocInfo 2018) (Санкт-Петербург, Россия); 3 МНК «Интеллектуальные информационные технологии для промышленности» (IITI 2018) (Сочи, Россия); 4 МНК «Интеллектуальные информационные технологии для промышленности» (IITI 19) (Отава, Чехия); 28-Я НТК «Методы и технические средства обеспечения безопасности информации» (МиТСОБИ 2019) (Санкт-Петербург, Россия); МНТК

«Автоматизация» (RusAutoCon 2019) (Сочи, Россия); 13я МНК «Intelligent Distributed Computing (IDC 2019) (Санкт-Петербург, Россия); XI Санкт-Петербургская межрегиональная конференция «Информационная безопасность регионов России» (ИБРР 2019) (Санкт-Петербург, Россия) и др.

Реализация результатов работы. Отраженные в диссертационной работе исследования проведены в рамках следующих НИР: грант российского научного фонда РНФ № 18-71-10094 «Мониторинг и противодействие вредоносному влиянию в информационном пространстве социальных сетей»; грант РНФ № 18-11-00302 «Интеллектуальная обработка цифрового сетевого контента для эффективного обнаружения и противодействия нежелательной, сомнительной и вредоносной информации». Полученные результаты внедрены в учебный процесс СПбГУТ им. проф. М.А. Бонч-Бруевича (учебный курс «Технологии обеспечения информационной безопасности больших данных») и СПбГУПТД (учебные курсы «Комплексная защита на предприятии», «Технологии и методы программирования»), применяются в рабочем процессе репутационного агентства «Glorystory» (компания ООО «Жасмин»). Результаты исследования представлены в заявках, победивших на следующих конкурсах: 1) Конкурс инноваций и инновационных проектов в номинации «А» «Конкурс концептуальных идей, методик, рекомендаций» 2016/2017 г. от Международной академии связи; 2) Конкурс субсидий молодым ученым, молодым кандидатам наук вузов, отраслевых и академических институтов, расположенных на территории Санкт-Петербурга, в 2018 г.

Публикации. По материалам диссертационного исследования было опубликовано 20 статей, в том числе 6 в рецензируемых изданиях из перечня ВАК, 1 статья опубликована в рецензируемом международном журнале, индексируемом WoS/Scopus, 7 статей – в сборниках трудов международных конференций, индексируемых в базах WoS и/или Scopus, 6 статей опубликованы в журналах и/или в сборниках трудов конференций, включенных в РИНЦ, получены 3 свидетельства о государственной регистрации программ для ЭВМ.

Личный вклад автора в статьи, опубликованные в рецензируемых изданиях из перечня ВАК следующий: в [1] Витковой Л.А. предложена модель распространителя вредоносной информации в социальных сетях; в [2] – модели социальных сетей; в статье [3] Витковой Л.А. принадлежит алгоритм оценки сложности для выбора мер противодействия вредоносной информации в социальных сетях и алгоритм выбора коэффициентов сложности на основе экспертных оценок; в [4] Виткова Л. А. предложила общую архитектуру системы обнаружения и противодействия вредоносной информации и описала функциональную структуру компонентов противодействия. Для [6] Виткова Л.А. провела исследование подходов к анализу социальных сетей, основанных на потоковых методах и на концепции социально-сетевого анализа. Все результаты, представленные в диссертационной работе, получены лично автором в процессе выполнения научно-исследовательской деятельности.

Структура и объем диссертационной работы. Диссертационная работа включает введение, три главы, заключение, список использованных источников (162 наименования) и 6 приложений. Объем работы – 173 страницы машинописного текста; включая 35 рисунков и 15 таблиц.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во введении обоснованы важность и актуальность темы диссертационной работы, определена цель и сформулированы задачи, решение которых необходимо для ее достижения. Показаны научная новизна и практическая значимость работы. Дано краткое описание разработанных комплексов моделей и алгоритмов, методики противодействия вредоносной информации, архитектуры и программных прототипов компонентов СПД вредоносной информации в СС. Представлены основные результаты их реализации в научно-исследовательских проектах.

Первая глава диссертации посвящена исследованию задачи противодействия вредоносной информации с учетом требования к обоснованности выбора объекта воздействия и контрмеры. Определены место и роль противодействия вредоносной информации в социальных сетях в информационной безопасности государства, общества и личности. Приведены основные определения и обзор релевантных моделей, алгоритмов и методик, описаны существующие системы противодействия вредоносной информации. Определены достоинства существующих подходов и решений, выделены их основные недостатки, затрудняющие противодействие вредоносной информации. Обоснована актуальность цели исследования. Предложено использование методики, основанной на анализе источников, сортировке объектов воздействия по приоритету, ранжировании контрмер для решения поставленной в исследовании цели.

На основании проведенного исследования определено множество функциональных и нефункциональных свойств противодействия вредоносной информации в социальных сетях и требований к методике противодействия. Выделены следующие свойства противодействия вредоносной информации в СС: *оперативность* – время, необходимое для противодействия вредоносной информации в социальных сетях; *обоснованность* – совокупность учитываемых параметров для выбираемых объектов воздействия и контрмер в процессе противодействия; *ресурсопотребление* – вероятность того, что количество использованных ресурсов не будет превышать допустимое значение.

Определены входные и выходные параметры для исследования.

Дано:

$$DATASET \subseteq \{messages, sources\}, \quad (1)$$

где $messages$ – множество сообщений, содержащих вредоносную информацию, $sources$ – множество источников этих сообщений.

$$MESSAGE = < messageURL, source, activity, messageType >, \quad (2)$$

где $messageURL$ – адрес сообщения в СС, $source$ – источник сообщения, $messageType$ – тип сообщения (пост, комментарий или ответ на комментарий), $activity$ – характеристики сообщения.

$$SOURCE = < sourceID, sourceURL >, \quad (3)$$

где $sourceID$ – уникальный идентификатор источника, $sourceURL$ – адрес источника в СС.

$$ACTIVITY =< countLike, countRepost, countView, countComment >, \quad (4)$$

где $countLike$ – количество отметок «мне нравится», $countRepost$ – количество «репостов» (копий со ссылкой на источник), $countView$ – количество просмотров, а $countComment$ – количество комментариев.

Требуется найти:

$$DATASET\_MAX \subseteq \{messages\_max, sources\_max\}, \quad (5)$$

где $messages\_max$ – множество сообщений ($messages$), у которых характеристики $activity$ будут самыми высокими по сравнению с другими сообщениями в множестве $messages$, а $sources\_max$ – множество источников ($sources$), которые связаны с максимальным количеством сообщений ($messages$), входящих в множество $messages\_max$.

Сформулирована задача исследования. Она заключается в разработке: (1) комплекса моделей социальной сети, источника и вредоносной информации; (2) комплекса алгоритмов анализа источников вредоносной информации в социальных сетях и ранжирования контрмер; (3) методики противодействия вредоносной информации в социальных сетях с учетом требований к обоснованности; (4) архитектуры и программных прототипов компонентов системы противодействия вредоносной информации в социальных сетях.

Сформулирована цель исследования – повышение эффективности противодействия вредоносной информации в социальных сетях. В диссертации показатель эффективности определяется через показатель обоснованности, а также с учетом требований к оперативности и к ресурсопотреблению.

Во второй главе предложен *комплекс моделей*, описывающий социальную сеть, источник и вредоносную информацию, а также *комплекс алгоритмов* анализа источников и ранжирования контрмер, необходимые для решения поставленных задач.

На первом этапе для представления элементов социальной сети (СС) проанализированы метаданные социальных сетей и построена модель социальной сети, согласно которой СС – это совокупность взаимосвязанных классов: (1) source, (2) message, содержащих такие объекты и их атрибуты как: (1.1) id_source, (1.2) followers, (1.3) message; (2.1) url_message, (2.2) type_message, (2.3) like, (2.4) comment, (2.5) repost, (2.6) answer, (2.7) views. Характеристики объектов и их атрибутов позволяют анализировать источники вредоносной информации в социальных сетях и сравнивать их.

На втором этапе предложена модель источника, в которой новыми характеристиками являются: (1) potential, (2) last_visit, (3) registration_time, (4), social_network_name, (5) index_activity, (6) index_impact, (7) index_viewability.

На третьем этапе на базе предложенных моделей социальной сети и источника вредоносной информации разработана теоретико-множественная модель вредоносной информации в социальной сети, которая включает такие базовые

элементы как: информационный объект IO (от англ. information object), признак информационной угрозы T (от англ. threat), MIO вредоносный информационный объект (от англ. malicious information object), признак информационной угрозы, содержащийся во вредоносном информационном объекте $Token$ (от англ. token), дискретный признак информационного объекта $Feature$ (от англ. feature) и связи между объектами. Теоретико-множественная модель формально представлена следующим образом:

$$\begin{gathered} IO = \{io\}; MIO = \{io\}; MIO_i = \{io\}, \\ MIO \subset IO; \forall\, io \in MIO: io \in IO, \\ MIO_i \subseteq MIO; \forall\, io \in MIO_i: io \in MIO, \\ Token_{mio_i} \subset T;\ Token_{mio_i} = \{t\}, \\ CheckFeature(io, t) = \{True; False\}, \\ io \in MIO_i \Leftrightarrow \exists\ Token_{mio_i}: checkFeature\ (io,\ t) = True\ , \end{gathered} \tag{6}$$

где IO – множество информационных объектов, io_1 – один информационный объект, T –множество всех возможных признаков информационной угрозы, t_n – один признак, MIO –множество вредоносных информации (множество вредоносных информационных объектов), MIO_i– отдельный класс вредоносной информации, $Token_{mio_i}$– множество признаков характеризующих MIO.

Для формирования множества признаков вредоносной информации в работе предложена информационно-признаковая модель, в которую включены следующие элементы: (1) информационная угроза – задается оператором СПД; (2) вредоносная информация в СС – задается оператором СПД путем формирования набора ключевых слов; (3) информационные признаки, формирующие множество всех возможных признаков.

Разработанный комплекс моделей социальной сети, источника и вредоносной информации содержит новые классы и атрибуты объектов, новые связи между ними, а также позволяет сформировать требования к алгоритмам анализа и оценки источников и выбора контрмер.

Комплекс алгоритмов анализа источников вредоносной информации и ранжирования контрмер (рис. 1) состоит из (1) алгоритма ранжирования источников по потенциалу, (2) алгоритма оценки источников, (3) алгоритма сортировки объектов воздействия, (4) алгоритма ранжирования контрмер.

Формальная запись комплекса анализа источников и ранжирования контрмер имеет следующий вид:

$$Z = SC \rightarrow max\ , \tag{7}$$

$$\begin{cases} f_1(S) \to I_p^s = \{0,1,2\}, \\ f_2(S) \to I_i^s[0,1,2], \left(I_i = \frac{I_i}{\max I + 1}\right), \\ f_3(S) \to I_{pr}^s = I_p^s + I_i^s = [0,4], \\ f(C) \to complexity(c_x) = \frac{cw_x * \sum_{i=1}^{|KC|} w_i * \left(\sum_{j=1}^{|KC_i|} (cp_{x,i,j} * lc_{x,i,j})\right)}{100 * |KC|}, f(c) \to (0;1], \end{cases} \quad (8)$$

где: S – источник, C – контрмера, $f_1(S)$ – индекс потенциала источника (I_p^s) равен 0, 1, 2 в зависимости от количества сообщений в анализируемом наборе данных, принадлежащих источнику. Вычисляется по «алгоритму ранжирования источников по потенциалу». $f_2(S)$– индекс влиятельности источника (I_i^s), значение которого находится между 0 и 2. Вычисление индекса влиятельности происходит по «алгоритму оценки источников». $f_3(S)$ – приоритет источника (I_{pr}^s) в качестве объекта воздействия в анализируемом наборе данных. Для получения значения применяется «алгоритм сортировки объектов воздействия». $f(C)$ – ранжированные контрмеры с учетом их сложности. Ранжирование происходит согласно алгоритму ранжирования контрмер.

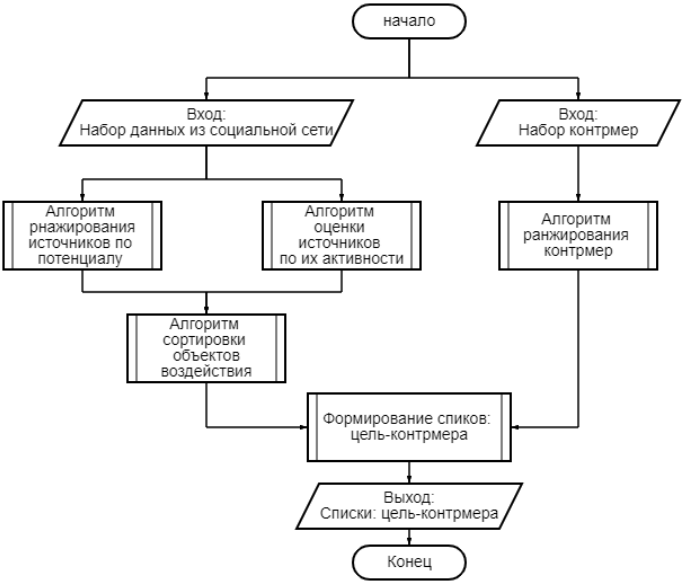


Рисунок 1 – Схема комплекса алгоритмов анализа источников и ранжирования контрмер

На выходе из комплекса алгоритмов формируются списки пар: цель-контрмера, при этом правила для выбора объектов воздействия в качестве цели ($target$) следующие:

$$\{source \in TARGET | I_{pr}^{s} \cong max\}, \quad (9)$$
$$\{message \in TARGET | I_{pr}^{s} \cong min\}, \quad (10)$$

где $TARGET$ – это множество объектов воздействия.

Разработанный комплекс алгоритмов анализа источников вредоносной информации и ранжирования контрмер отличается от существующих аналогов, учетом таких атрибутов как потенциал источника, активность пользователей на странице источника, количество просмотров сообщения с вредоносной информацией. Алгоритм ранжирования контрмер отличается от аналогов учетом коэффициентов и уровней сложности для каждой меры противодействия. При этом разработанный комплекс алгоритмов позволяет сформировать требования к методике противодействия вредоносной информации и является основой для СПД.

В третьей главе представлены методика противодействия вредоносной информации в СС, архитектура и программные прототипы системы, результаты экспериментальной и теоретической оценки методики.

Методика противодействия вредоносной информации в СС состоит из двух стадий: (1) стадия настройки и (2) стадия эксплуатации. При этом стадия настройки методики состоит из двух шагов: Шаг 1. «Настройка системы запросов», на котором оператор определяет информационные угрозы и их признаки, а система противодействия формирует и сохраняет списки угроз и их признаков. Шаг 2. «Ранжирование контрмер», на котором оператор выбирает доступные агенты реализации, система формирует и сохраняет списки доступных агентов реализации. Далее оператор выбирает доступные контрмеры, система формирует список контрмер и выбирает коэффициенты сложности контрмер на основе экспертных оценок, затем система формирует и сохраняет список ранжированных контрмер. Стадия эксплуатации методики состоит из 3х шагов и представлена на рисунке 2.

Выходными данными методики являются: (1) возможные информационные угрозы, признаки, контрмеры и их коэффициенты, доступные агенты реализации мер противодействия; (2) различные параметры объектов воздействия, согласно которым оператор распределяет свое внимание и очередность принятия решения о противодействии; (3) сформированные пары цель-контрмера для противодействия вредоносной информации в социальных сетях через доступные агенты реализации.

Разработанная методика отличается от известных использованием авторских алгоритмов анализа источников и ранжирования контрмер, благодаря чему повышается обоснованность принятия решения о противодействии цели и выбора контрмеры и сокращается время работы оператора в процессе противодействия вредоносной информации в СС.

Архитектура и программные прототипы компонентов СПД вредоносной информации представлена на рисунке 3.

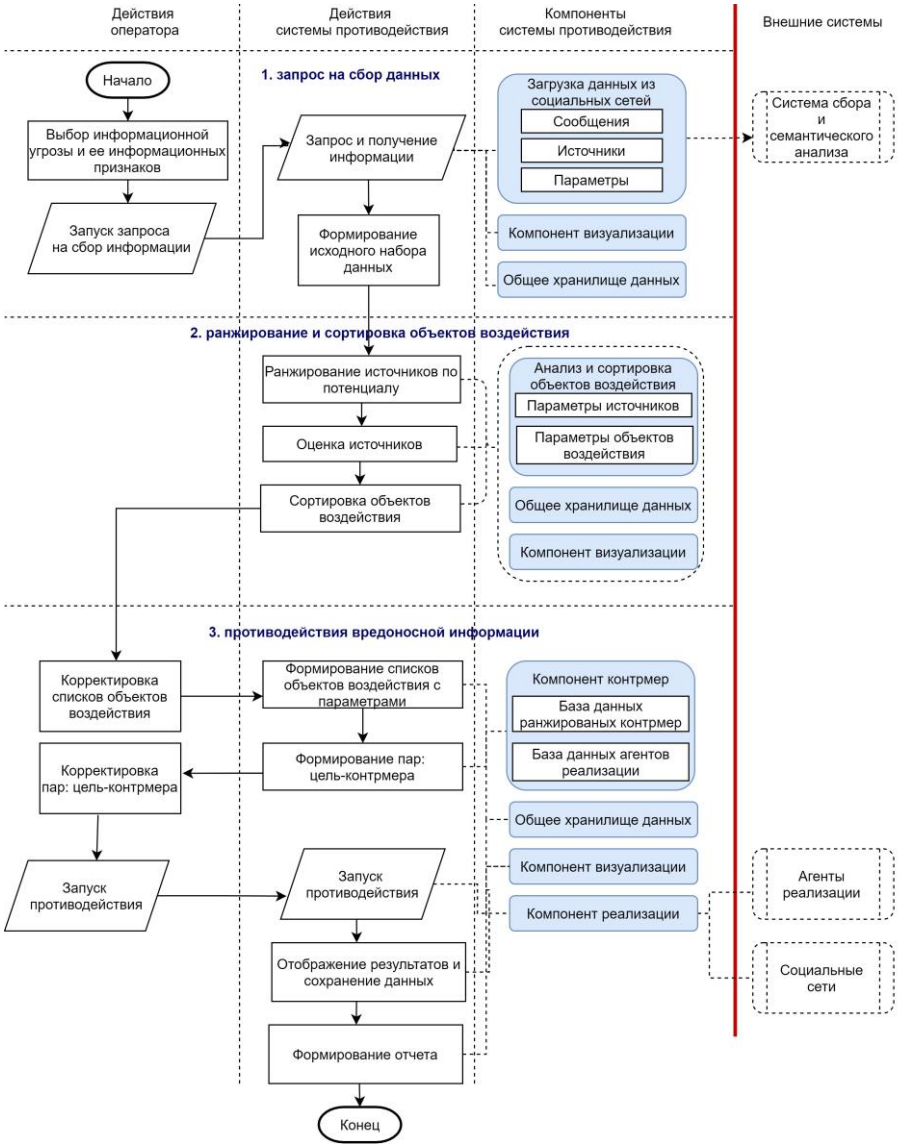


Рисунок 2 – Представление методики противодействия вредоносной информации на стадии эксплуатации

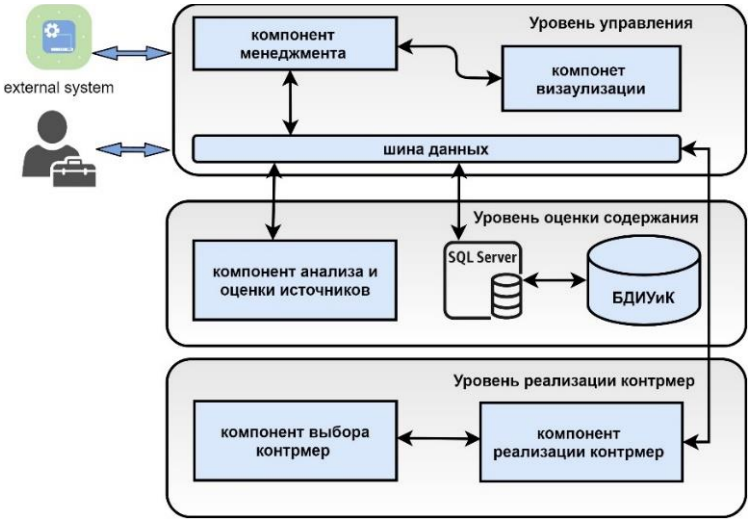


Рисунок 3 – Архитектура СПД
вредоносной информации в социальных сетях

Архитектура включает три уровня: (1) уровень управления (компонент менеджмента (управления потоками и запросами), компонент визуализации (отчетность)); (2) уровень оценки содержания (компонент анализа и оценки источников, SQL сервер и база данных); (3) уровень реализации контрмер (компонент выбора контрмер и компонент реализации контрмер).

Элементы архитектуры реализованы в качестве программных прототипов: (1) программный прототип компонента анализа и оценки источников в СС, который включает алгоритм ранжирования источников, алгоритм оценки источников и алгоритм сортировки объектов воздействия; (2) программный прототип компонента выбора контрмер, который включает алгоритм ранжирования контрмер, алгоритм экспертных оценок для формирования коэффициентов; (3) программный прототип базы данных информационных угроз и контрмер (БДИУиК), который содержит информацию о мерах противодействия вредоносной информации в СС, о типах информационных объектов, к которым контрмеры могут быть применимы, об агентах реализации, через которые могут быть реализованы меры противодействия.

Для экспериментальной оценки из социальной сети были собраны данные из социальной сети, проведены измерения времени комплекса алгоритмов анализа источников и ранжирования контрмер, получены показатели нагрузки на центральный процессор и на оперативную память. Далее для формирования исходных данных были проведены исследования и эксперименты, выяснилось, что самым затратным процессом с точки зрения оперативности является время работы оператора на стадии настройки методики и на 1-м, 4-м шагах на стадии эксплуатации методики. Для оценки показателя времени работы оператора для принятия решения

о противодействии по методике и без нее были проведены эксперименты, в которых приняли участие 10 экспертов. По итогам экспериментальной оценки оперативности рассчитана вероятность выполнения методики за заданное время, которая составляет $P_{operability}(T_m \leq T^{additional}) = 0.9942$, что соответствует предъявляемым требованиям ($P_{operability}^{acceptable} = 0{,}99$) к оперативности.

Оценка ресурсопотребления проводилась по ряду частных показателей, характерных для 2-го шага стадии эксплуатации методики противодействия вредоносной информации в СС. Учитывалась нагрузка на центральный процессор, использование оперативной памяти и время работы оператора. Показано, что оценка ресурсопотребления соответствует предъявляемым требованиям $P_{res}\left(r \leq R^{acceptable}\right) \geq P_{res}^{acceptable}$, где P_{res} – вероятность того, что ресурсы r, затрачиваемые на противодействие вредоносной информации по методике, не превышают допустимого значения $R^{acceptable} = 75\%$, $P_{res}^{acceptable}$ – это допустимое значение вероятности.

В рамках теоретической оценки проводилось сравнение показателей обоснованности для разработанной методики и аналогов, таких как решения лаборатории Касперского, Zerofox, Ithreat Cyber Group Inc и др. Показано, что разработанная методика учитывает большее количество параметров для выбираемых объектов воздействия и контрмер в ходе противодействия вредоносной информации в СС, при соблюдении требований к другим свойствам. В сравнении с аналогами количество учитываемых параметров при использовании методики больше, такое что $N_{param}^{M} > maxN_{param}^{S}$, где N_{param}^{M} – количество учитываемых параметров для методики, $maxN_{param}^{S}$ – максимальное количество учитываемых параметров для аналогов. При этом $N_{param}^{M} = 12, maxN_{param}^{S} = 8$.

Проведен сравнительный анализ разработанной методики с известными методиками по используемым функциональным возможностям, таким как: А – возможность формирования задач сбора и анализа сообщений для системы мониторинга; Б – возможность настройки доступных мер противодействия в системе; В– возможность анализа источников сообщений в полученном наборе данных; Г – возможность ранжирования и сортировки объектов воздействия в полученном наборе данных; Д – возможность ранжирования и сортировки доступных контрмер из базы контрмер для каждого набора данных; Ж – возможность выбора цели воздействия для противодействия.

Результаты приведены в таблице 1 (используются следующие обозначения и баллы: «+» – наличие параметра в работе (1 балл); «+/–» – частичное соответствии параметру (0.5 балла); «–» –отсутствие параметра (0 баллов).

Анализ результатов сравнения методики противодействия вредоносной информации в социальных сетях с аналогами позволяет сделать следующие выводы. Во-первых, ни одна из методик, кроме предложенной, одновременно не удовлетворяет всем функциональным требованиям. Во-вторых, все методики в той или иной степени позволяют ранжировать контрмеры. В-третьих, параметры сообщений, источников, контрмер учитываются только в предложенной методике и

в решении от компании Creopoint Inc. В-четвертых, отставание ближайших аналогов от предложенной методики составляет от 1,5 балла до 4-х. То есть, предложенная методика выигрывает у ближайших аналогов.

Таблица 1 – Сравнение разработанной методики с известными аналогами

Методика противодействия вредоносной информации в СС	**Учитываемые параметры**						**Оценка**
	А	**Б**	**В**	**Г**	**Д**	**Ж**	
RU2651252, Лаборатория Касперского	–	–	+	–	+/–	+/–	2
Zerofox Inc «Brand Protection»	+	+	+	–	+/–	+/–	4
Ithreat Cyber Group Inc	–	+	–	–	+/–	–	1.5
Creopoint Inc	+	+	+	+/–	+/–	+/–	4.5
ЕАИС Роскомнадзора	+	+	–	–	+/–	+/–	3
Разработанная методика	**+**	**+**	**+**	**+**	**+**	**+**	**6**

Таким образом, полученные в диссертации результаты позволяют утверждать о достижении более высокой эффективности разработанной методики по сравнению с известными, что доказывает реализацию итоговой цели исследования – повышения эффективности противодействия вредоносной информации за счет анализа источников вредоносной информации и автоматизации выбора контрмер.

В заключении представлена итоговая оценка проделанной работы, приведены основные результаты исследования и описаны перспективы дальнейшего исследования в рамках темы.

ЗАКЛЮЧЕНИЕ

Предложенные модели, алгоритмы, методика и архитектура, а также их практическая реализация в совокупности обеспечивают решение актуальной научно-технической задачи повышения эффективности противодействия распространению вредоносной информации в социальных сетях. Результаты диссертационной работы составляют следующие **итоги** исследования:

1. Предложен комплекс моделей СС, источника и вредоносной информации, отличающийся от существующих аналогов возможностью одновременного учета структуры информационного обмена в СС, источников и вредоносной информации.
2. Разработан комплекс алгоритмов анализа источников вредоносной информации и ранжирования контрмер, в котором, в отличие от существующих, учитываются связи и зависимые атрибуты объектов в СС, такие как: потенциал источника, активность пользователей на странице, количество просмотров сообщения и др. Алгоритмы ранжирования контрмер учитывают коэффициенты и уровни сложности для каждой меры противодействия.
3. Предложена методика противодействия вредоносной информации в СС, ориентированная на автоматический и автоматизированный выбор объектов воздействия и мер противодействия вредоносной информации из списка ранжированных контрмер.
4. Разработана архитектура и программные компоненты системы противодействия вредоносной информации, отличающаяся от существующих архитектур тем, что поддерживает ранжирование и выбор доступных оператору

контрмер в системе для заданной оператором вредоносной информации. Архитектура содержит оригинальные компоненты анализа и оценки источника вредоносной информации, базу данных с информацией о мерах противодействия вредоносной информации в социальных сетях.

В качестве **рекомендаций** по дальнейшей разработки темы заключаются в расширении класса алгоритмов анализа поведения источников и авторов сообщений, алгоритмов анализа распространения информации в СС, интеграции механизмов автоматического и автоматизированного противодействия в существующие архитектуры и системы.

Полученные результаты соответствуют п. 3 «Методы, модели и средства выявления, идентификации и классификации угроз нарушения информационной безопасности объектов различного вида и класса», п. 5 «Методы и средства (комплексы средств) информационного противодействия угрозам нарушения информационной безопасности в открытых компьютерных сетях, включая Интернет» паспорта специальности 05.13.19 – «Методы и системы защиты информации, информационная безопасность».

СПИСОК РАБОТ, ОПУБЛИКОВАННЫХ АВТОРОМ ПО ТЕМЕ ДИССЕРТАЦИИ

Публикации в рецензируемы научных изданиях из Перечня ВАК:

1. Виткова Л.А. Модель вредоносной информации и ее распространителя в социальных сетях / Л.А. Виткова, Д.В. Сахаров, Д.Р. Голузина // Защита информации. Инсайд. – Спб., 2020. –№3 (93). – С. 66-72. **(05.13.19)**
2. Гамидов Т.О. Разработка моделей и алгоритмов анализа данных для исследования хода инцидентов и кризисов в социальных сетях / Т.О. Гамидов, Л.А. Виткова, М.М. Ковцур // Вестник Санкт-Петербургского государственного университета технологии и дизайна. Серия 1: Естественные и технические науки. – СПб., 2020. – № 2. – С. 3-10. **(05.13.19)**
3. Виткова Л.А. Выбор мер противодействия вредоносной информации в социальных сетях / Л.А. Виткова, А.А. Чечулин, Д.В. Сахаров // Вестник Воронежского института ФСИН России. – Воронеж, 2020. – Т. 3. – С. 20-29. **(05.13.19)**
4. Виткова Л.А. Архитектура системы выявления и противодействия нежелательной информации в социальных сетях. / Л.А. Виткова, И.Б. Саенко // Вестник Санкт-Петербургского государственного университета технологии и дизайна. Серия 1: Естественные и технические науки. – СПб., 2020. – № 3. – С. 33-39. **(05.13.19)**
5. Виткова Л.А. Методика анализа аудитории канала распространения информации в социальных сетях. // Известия высших учебных заведений. Технология легкой промышленности. – СПб, 2018. – Т. 42, № 4. – С. 5-10.
6. Проноза А.А. Методика выявления канала распространения информации в социальных сетях / А.А. Проноза, Л.А. Виткова, А.А. Чечулин, И. В. Котенко, Д.В. Сахаров // Вестник Санкт-Петербургского университета. Прикладная математика. Информатика. Процессы управления. – СПб., 2018. – Т. 14, № 4. – С.362-377

Публикации в зарубежных изданиях из баз данных WOS и Scopus:

7. Kotenko I.V. The intelligent system for detection and counteraction of malicious and inappropriate information on the Internet / I.V. Kotenko, L.A. Vitkova, I.B. Saenko, O.N. Tushkanova, A.A. Branitsky// AI Communications, 2020. – Vol 33(1). – C. 1-13. – ISSN 0921-7126
8. Vitkova L.A. Selection of countermeasures against propagation of harmful information via Internet / L. A. Vitkova, A. P. Pronichev, E. V. Doynikova, I. B. Saenko // IOP Conference Series: Materials Science and Engineering, 2021 Vol 1032, – 1032 012017. – ISSN 1757-8981
9. Vitkova, L.A. The technology of intelligent analytical processing of digital network objects for detection and counteraction of inappropriate information / L.A. Vitkova, I.B. Saenko, A.A. Chechulin, I.B. Parashchuk // The 1st International Conference on Computer Technology Innovations dedicated to the 100th anniversary of the Gorky House of Scientists of Russian Academy of Science (ICCTI — 2020). Official conference proceedings, 2020. – P 13-19. – ISBN 978-5-9676-1216-9
10. Vitkova L.A. Approach to Identification and Analysis of Information Sources in Social Networks / L. A. Vitkova, M. V. Kolomeets // Proceedings of the 13th International Symposium on Intelligent Distributed Computing (IDC 2019), October 7-9, 2019, Saint-Petersburg, Russia. 2020. P. 285-293. – ISSN 1860-949X.
11. Vitkova L.A. An Approach to Creating an Intelligent System for Detecting and Countering Inappropriate Information on the Internet / L.A. Vitkova, I.B. Saenko, O.N. Tushkanova // Proceedings of the 13th International Symposium on Intelligent Distributed Computing (IDC 2019), October 7-9, 2019, Saint-Petersburg, Russia. 2020. – P. 244-254. – ISSN 1860-949X.
12. Vitkova, L.A. Hybrid Approach for Bots Detection in Social Networks Based on Topological, Textual and Statistical Features / L.A. Vitkova, Kotenko I.V., M.V. Kolomeets, O.N. Tushkanova, A.A. Chechulin // Advances in Intelligent Systems and Computing 1156 AISC, 2019, P. 412-421
13. Pronoza A.A. Visual analysis of information dissemination channels in social network for protection against inappropriate content / A.A. Pronoza, L.A. Vitkova, A.A. Chechulin, I.V. Kotenko // 3rd International Scientific Conference on Intelligent Information Technologies for Industry, IITI 2018. Sochi, Russian Federation, 17-21 September 2018. Advances in Intelligent Systems and Computing. Vol. 875, 2019. P. 95-105.
14. Kotenko I.V. Monitoring and counteraction to malicious influences in the information space of social networks / I.V. Kotenko, I.B. Saenko, A.A. Chechulin, V.A. Desnitsky, L.A. Vitkova, A.A. Pronoza // The 10th Social Informatics conference (SocInfo2018). September 25–28, 2018, Saint Petersburg, Russia. Proceedings, Part II. Lecture Notes in Computer Science, Vol.11186, Springer 2018, P.1 59-167. – ISBN 978-3-030-01158-1.

Публикации в сборниках трудов конференций, включенных в РИНЦ:

15. Виткова Л.А. Методология выявления искусственной мобилизации протестной активности в соцсетях / Л.А. Виткова, К.А. Науменко // Тезисы докладов научного семинара «Фундаментальные проблемы управления производственными процессами в условиях перехода к Индустрии 4.0» в рамках МНТК «Автоматизация», 2020. – С. 212-214
16. Виткова Л.А. Модель и алгоритмы защиты от вредоносной информации в социальных сетях // IX МНТиНМК «АПИНО». СПб: СПбГУТ. Сборник научных статей. 2020. – Т. 1. – С. 235-240.
17. Валиева К.А. Методика обнаружения вредоносной информации в информационном пространстве социальных сетей / К.А. Валиева, Л.А. Виткова, Е.В. Смирнов IX МНТиНМК «АПИНО». СПб: СПбГУТ. Сборник научных статей. 2020. – Т. 1. – С. 206-211.
18. Виткова Л.А. Противодействие распространению нежелательной информации в информационном пространстве социальных сетей / Л. А. Виткова, М.А. Справцева // IX МНТиНМК «АПИНО». СПб: СПбГУТ. Сборник научных статей. 2020. – Т. 1. – С. 258-261.
19. Виткова Л.А. О моделировании процессов выявления и противодействия террористической и экстремистской активности в интернете и социальных сетях / Л.А. Виткова, Е.В. Дойникова, А.П. Проничев // Сборник научных статей XVII Санкт-Петербургской международной конференции «Региональная информатика (РИ-2020)». СПб: СПОИСУ, 2020. – С. 117-118.
20. Виткова Л.А. Распределенный сбор и обработка данных в системах мониторинга информационного пространства социальных сетей / Л.А. Виткова, И.В. Котенко, А.В. Хинензон // VIII МНТиНМК «АПИНО». СПб: СПбГУТ. Сборник научных статей. 2019. – Т. 1. – С. 228-232

Свидетельства о государственной регистрации программ для ЭВМ:

21. Виткова Л.А. Компонент сегментации пользователей по их активности в социальных сетях / Л.А. Виткова, А.А. Чечулин, И.В. Котенко – Свидетельство о государственной регистрации программы для ЭВМ № 2019664733. Зарегистрировано в Реестре программ для ЭВМ 13.11.2019.
22. Федорченко Е.В. Компонент выбора мер противодействия нежелательной, сомнительной и вредоносной информации / Е.В. Федорченко, Л.А. Виткова, А.П. Проничев, И.Б. Саенко. – Свидетельство о государственной регистрации программы для ЭВМ № 2020665591. Зарегистрировано в Реестре программ для ЭВМ 27.11.2020.
23. Виткова Л.А. База данных для учета нежелательной информации совместно с мерами противодействия / Л.А. Виткова, Е.О. Березина, А.П. Проничев, И.Б. Саенко, И.В. Котенко – Свидетельство о государственной регистрации программы для ЭВМ № 2020622557. Зарегистрировано в Реестре программ для ЭВМ 08.12.2020.

Автореферат диссертации

ВИТКОВА
Лидия Андреевна

МОДЕЛИ, АЛГОРИТМЫ И МЕТОДИКА ПРОТИВОДЕЙСТВИЯ ВРЕДОНОСНОЙ ИНФОРМАЦИИ В СОЦИАЛЬНЫХ СЕТЯХ

Текст автореферата размещен на сайтах:
Высшей аттестационной комиссии Министерства науки и высшего образования Российской Федерации
https://vak.minobrnauki.gov.ru/
Федеральное государственное бюджетное учреждение науки «Санкт-Петербургский Федеральный исследовательский центр Российской академии наук»
http://www.spiiras.nw.ru/dissovet/

Подписано в печать 27.04.2021
Формат 60x84 1/16. Бумага офсетная. Печать офсетная.
Усл.печ.л. 1,0. Тираж 100 экз.
Заказ № ___