

ТЕХНОЛОГИИ ПОИСКА ДИСКРЕДИТИРУЮЩЕЙ ИНФОРМАЦИИ В СОЦИАЛЬНЫХ МЕДИА

В.А. Минаев, К.М. Бондарь, А.С. Коломина

В статье рассматриваются вопросы выявления, анализа и предотвращения угроз дискредитации организаций и конкретных индивидов с использованием социальных медиа. Обращается внимание на применение математических методов и моделей при описании процессов распространения дискредитирующей информации. Дается определение, обсуждаются основные формы и приводятся примеры дискредитации. Показана роль социальных медиа в системе реализации дискредитирующих воздействий и технологические особенности медийного воздействия. Обсуждаются большие данные (Big Data), представляющие структурированные и неструктурированные массивы данных большого объема, а также современные инструменты их обработки для анализа, прогнозирования и принятия управленческих решений; нейронные сети для поиска смысла информации, включая выявление манипулятивных воздействий на человека; визуализация взаимосвязей между ключевыми фигурами, событиями и организационными структурами, обеспечивающая понимание эффективных стратегий для управления ситуацией. Делается вывод о том, что в условиях роста информационных вызовов и усложнения цифровых рассмотренные подходы становятся незаменимыми для защиты репутации, обеспечения доверия к медиа информации.

Ключевые слова: технологии поиска, дискредитирующая информация, социальные медиа, моделирование.

Введение

Киберпространство, являясь ключевым элементом современной информационной среды, продолжает стремительно развиваться. С его расширением открываются не только новые возможности для коммуникации, сотрудничества и обмена данными, но и возрастают угрозы, связанные с деструктивным использованием цифровых технологий [1]. Одной из значимых проблем становится распространение дискредитирующей информации, подрывающей доверие к организациям, социальным институтам, отдельным лицам.

Отсутствие четкой системы выявления, анализа и предотвращения угроз дискредитации существенно увеличивает вероятность информационных атак на вышеуказанные объекты. В этом контексте создание математических моделей и разработка методов анализа признаков дискредитации сотрудников, основанных на данных из открытых источников, становится не просто актуальными, но и крайне важными задачами.

Необходимость научного исследования указанной темы особенно заметна на фоне растущего числа информационных манипуляций, нацеленных на подрыв репутации объектов из государственных и коммерческих структур. Такие действия приводят к серьезным последствиям не только локального, но и глобального характера, сказываясь на политической, экономической и социальной стабильности как отдельных групп социума, так и общества в целом [2].

Дискредитирующие воздействия в Сети

В современном мире воздействие на общественное сознание и умы отдельных индивидов становится важнейшим фактором, определяющим работу организаций и их сотрудников. Дискредитация, как социальное явление, проявляется в различных формах — от банальной клеветы до манипуляций общественным мнением, нанося значительный ущерб отдельным личностям и целым организациям.

Дискредитация определяется как *преднамеренный и целенаправленный подрыв доверия к организации или отдельному индивиду путем распространения негативной, искаженной или ложной информации о них.*

Таким образом, основные формы дискредитации проявляются в: клевете – осознанном распространении недостоверных сведений об определенном объекте; искажении фактов – намеренной манипуляцией информацией для создания негативного восприятия того или иного объекта; фейковых новостях – вымышленных событиях и фактах, распространяемых в СМИ; ассоциации с негативным контекстом – связывании объекта с нежелательными идеями или персонажами [3].

Необходимо отметить, что термин «социальный» в контексте настоящей статьи обозначает принадлежность к обществу и взаимодействие между людьми, в то время как «медиа» подразумевает средство передачи информации в обществе. Объединяя эти понятия, термин «социальные медиа» можно определить как платформы, через которые люди делятся информацией, выражают мнения и обсуждают различные события или явления, происходящие в их жизнедеятельности.

Современный мир превратился в «глобальную деревню» благодаря стремительному развитию социальных медиа. Это означает, что все люди во всем мире могут в любой момент общаться друг с другом.

Социальные медиа (СМ) выступают как интерактивные платформы, созданные для упрощения коммуникации между частными пользователями, организациями и корпорациями. СМ постепенно вытесняют традиционные методы маркетинга и становятся неотъемлемым инструментом продвижения товаров и услуг, активно способствуют продвижению концепции «узнаваемости бренда».

СМ обеспечивают возможность обмена идеями, фотографиями и видеоматериалами, что делает их важным компонентом современных маркетинговых стратегий в широком плане, но одновременно открывают нишу для противодействия таким стратегиям

со стороны конкурента, противника, да и просто завистника, к чужому успеху, достижению, путем дискредитации последних.

Ключевыми составляющими дискредитирующих действий выступают три системно взаимосвязанных элемента: социальное взаимодействие, контент и коммуникационные средства. Эти три элемента формируют инфокоммуникационную основу в виде СМ для реализации дискредитирующих воздействий (рис. 1 [4]).

Очевидно, что для успешного функционирования всех представленных компонентов в «системе» дискредитации, если хотите так величать в целом это негативное явление, они должны быть интегрированы и доступны в нужное время. Раскроем кратко их содержание.

Социальное взаимодействие предполагает установление и развитие связей между пользователями социальных медиа через платформы.

Контент — это данные, доступные в социальных медиа, способные привлечь внимание аудитории.

Коммуникационные средства выполняют функцию обмена информацией и ускорения взаимодействия.

В эпоху глобальной цифровизации СМ стали важнейшим инструментом общения, предоставляя доступ к информации, влияя на общество, индивида как на локальном, так и на глобальном уровнях.

СМ значительно изменили восприятие людьми окружающего мира, способы коммуникации и обмена мыслями. Их преимущества многогранны [5], включая следующие:

1. **Анонимность** – можно скрывать свою личность, используя псевдонимы.
2. **Разнообразие** доступной информации.
3. **Публичное представление** жизнедеятельности.
4. **Оперативность** распространения информации.

5. *Формирование* новых социальных связей.

6. *Субъективность* воздействия, вытесняющая объективность восприятия данных.

7. *Изменение форматов* представления информации, включая трансформацию текстов, изображений и видеоматериалов.

8. *Запаздывание цензурирования*, не успевающего за скоростью информационных изменений.

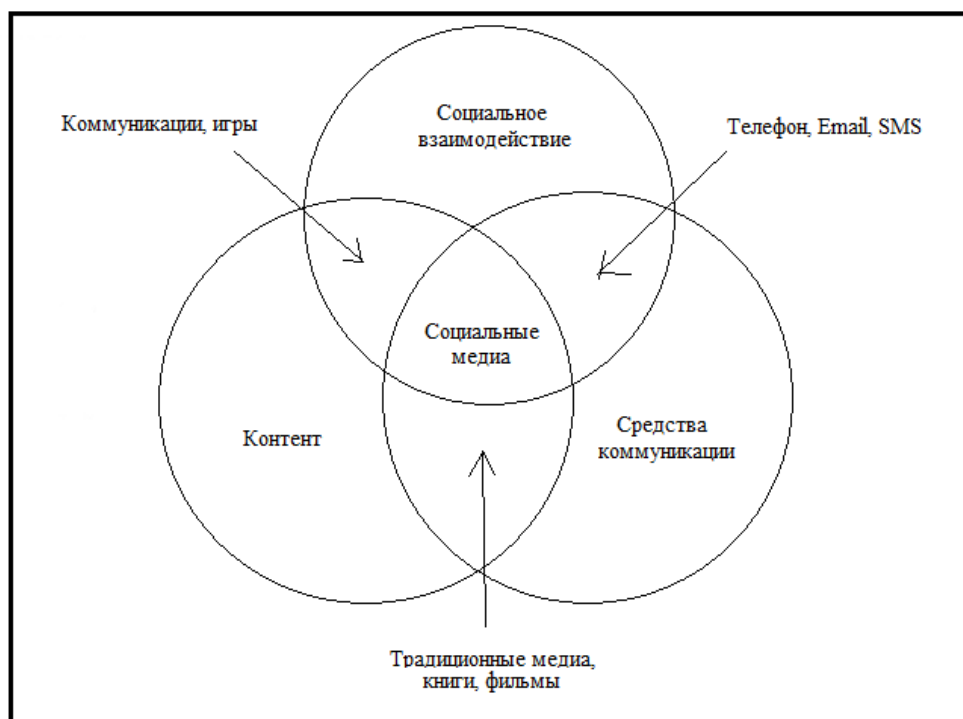


Рис. 1. Социальные медиа в системе реализации дискредитирующих воздействий [4]

СМ предоставляет по большому счету семь основных элементов:

1. *Личность* – личная информация, предоставляемая пользователем.

2. *Беседы* – общение между пользователями в социальных сетях.

3. *Общий доступ* – количество контента, которым делятся пользователи.

4. *Присутствие*, социальные медиа могут предоставлять информацию о пользователях независимо от того, находятся ли они в Сети или нет.

5. *Отношения* – уровень взаимодействия между пользователями.

6. *Репутация* показывает, насколько пользователи осведомлены о своем статусе и статусе других пользователей в социальных сетях.

7. *Группы* – структурный элемент, который описывает способность пользователей создавать сообщества и подгруппы в них.

Данные элементы представляют собой ценную информацию, которая может быть использована для дискредитации при наличии недоброжелательных намерений. Более того, применение методов открытой разведки позволяет выявлять даже ту информацию, которую пользователь не стремился распространять.

Роль медийных платформ в системе дискредитации

Медийные платформы представляют собой сложные структуры, объединяющие частных лиц и организации посредством реализации своих социально-экономических, эмоциональных и финансовых отношений и связей.

СМ выступают каналом коммуникации, связывающим людей, которые демонстрируют множество моделей поведения. В рамках исследований авторов изучаются вопросы математического

моделирования распространения дискредитирующей информации.

Сегодня СМ являются основой распространения дискредитирующей информации в мировом киберпространстве. Примеры подобного рода можно встретить в разных странах. Так, в США дискредитация политиков и бизнесменов стала частью повседневной реальности, особенно в период избирательных кампаний.

СМ используются для распространения недостоверных сведений о кандидатах, серьезно влияя на имидж и их общественное восприятие. Например, в ходе президентской кампании 2016 года против Хиллари Клинтон проведена масштабная кампания, включавшая публикацию ложной информации о ее здоровье и использовании личного сервера электронной почты [6]. Эти действия заметно ухудшили ее репутацию в глазах избирателей.

Дональд Трамп также подвергся атакам, включая обвинения в связях с Россией, что, в конечном итоге, повлияло на выборную кампанию и на общественное мнение в период его президентства [7].

В Европе дискредитация становится все более распространенным явлением, особенно в деловой сфере. В Германии компании сталкиваются с атаками на своих руководителей, когда конкуренты используют компрометирующую информацию для снижения доверия. Это негативно сказывается на моральном состоянии сотрудников, финансовом положении компаний и их репутации. Яркий пример – скандал с Volkswagen в 2015 году, приведший к значительным финансовым потерям и серьезному удару по репутации компании на международной арене [8].

На корпоративном уровне дискредитация ухудшает атмосферу в коллективе. А именно, сотрудники, ставшие мишенью информационных атак, часто чувствуют изоляцию и беспомощность, что снижает их мотивацию к продуктивному труду.

В Японии такие случаи уже приводили к массовым увольнениям персонала и утрате доверия к компаниям. Например, скандал вокруг Kobe Steel в 2017 году, связанный с подделкой сертификатов качества, подорвал

репутацию фирмы, привел к падению стоимости акций [9].

Дискредитация негативно влияет на отношения организаций с клиентами. Она снижает доверие с их стороны, что влечет за собой финансовые потери.

Очевидно, что для минимизации такого рода рисков необходимо активно работать над развитием положительного имиджа компании, укреплять внутренние коммуникации и развивать культуру доверия в ней. При этом обучение сотрудников медиаграмотности становится важным фактором в защите от информационных манипуляций в их отношении и сохранения репутации компании.

Возможности киберпространства предоставляют злоумышленникам мощные инструменты для манипуляции общественным мнением. Социальные сети и мессенджеры, информационные платформы для обмена контентом, различные форумы создают условия для стремительного распространения искаженной или ложной информации. Специальные алгоритмы, управляющие контентом, усиливают влияние провокационных или эмоциональных негативно насыщенных публикаций, позволяя дискредитирующим материалам быстро становиться вирусными.

Особую роль в этом процессе играют технологии ботов и фейковых аккаунтов, активно используемых для создания искусственной популярности дискредитирующего контента. Эти технологии помогают создавать иллюзию широкой общественной поддержки, усиливая влияние манипулятивной информации на аудиторию.

Распространение ложной информации нередко остается безнаказанным, поскольку многие не проверяют ее достоверность. Важную роль в этом играют популярные блогеры и медийные личности, которые формируют удобные нарративы для ими же формируемых целевых общественных групп. Используя возможности социальных сетей, видеоплатформ и подкастов, они создают специализированный контент, насыщенный эмоциями, который, благодаря алгоритмам цифровых платформ, стремительно набирает популярность. В результате формируются

новые медиа-системы, способствующие усилению и ускорению распространения дискредитирующей информации.

Технологические особенности медийного воздействия

Создание современных медиа-систем – процесс сложный, обусловленный взаимодействием целого ряда факторов. Важнейшим из них выступают алгоритмы, управляющие распределением контента в СМ. Такие алгоритмы ориентированы на повышение вовлеченности пользователей в реальную или выдуманную ситуацию и активно продвигают информацию, вызывающую яркие эмоции – гнев, страх или восхищение. Именно поэтому эмоционально насыщенные материалы приобретают намного больший охват по сравнению с нейтральными сообщениями.

Доступность цифровых технологий открыла широкие возможности для участия в создании контента. Сегодня им занимаются не только профессиональные журналисты, но и блогеры, активисты и даже анонимные авторы. Этот процесс значительно снижает порог входа в информационное пространство, позволяя обойти устоявшиеся традиционные стандарты журналистики, включающих проверку достоверности данных.

Влияние новостных систем усиливается за счет взаимодействия различных информационных платформ. Материалы, появившиеся в одной из них, быстро распространяются на другие, создавая взаимоусиливающие циклы. Кроме того, алгоритмы помогают формировать узкие сообщества, где пользователи окружены информацией, соответствующей их убеждениям. Это создает иллюзию ее достоверности, подкрепленную поддержкой якобы единомышленников.

Конечно же, нельзя забывать и про коммерческую составляющую, связанную с тем, что лидеры мнений, представляющие те или иные платформы, заинтересованы в повышении охватов и вовлеченности пользователей, поскольку это напрямую влияет на их доходы от рекламы и спонсорских контрактов. При этом эмоционально насыщенные, скандальные темы становятся основным инструментом для

привлечения внимания и удержания аудитории. Их контент часто более персонализирован, чем материалы традиционных медиа. Они делятся своими переживаниями, демонстрируют эмоции и активно взаимодействуют с подписчиками, формируя ощущение искренности.

Стоит отметить, что подчас скептическое отношение в определенных кругах населения к крупным медиакомпаниям также способствует популярности альтернативных источников информации. Многие воспринимают блогеров как непредвзятых комментаторов, которые несут правду, игнорируемую традиционными СМИ.

Медиа-системы, значительно изменив информационное пространство, в то же время стали катализатором распространения дискредитирующей информации. За счёт эмоциональной насыщенности и вирусности такие материалы усиливают поляризацию в обществе, углубляя разрыв между разными социальными группами, подрывая доверие к официальным источникам информации, усложняя борьбу с манипулятивными данными.

Современные лидеры мнений и создаваемые ими новостные системы оказывая колоссальное воздействие на восприятие информации и формирование общественного мнения, активно участвуют в формировании и управлении восприятием событий. Поэтому понимание механизмов их деятельности и факторов, способствующих их популярности, является необходимым шагом для разработки моделей и методов противодействия манипуляциям и предотвращения распространения дискредитирующей информации в киберпространстве. Ясно, что только технологически подкреплённый подход способен противостоять нарастающей в нем информационной нестабильности в условиях нарастающей неопределённости и угроз.

Для противодействия возникшим вызовам требуется применение современных технологий, таких как большие данные, нейронные сети, инструменты визуализации данных.

Большие данные (Big Data) представляют собой структурированные и неструктурированные массивы данных

большого объёма, а также современные автоматизированные инструменты их обработки для последующего применения полученных результатов для анализа, прогнозирования и принятия управленческих решений

Нейронные сети предоставляют возможность поиска смысла текстов, аудио- и видео информации, включая выявление манипулятивных воздействий на человека, например, DeepFake, или других форм информационных искажений.

Визуализация данных отображает взаимосвязи между ключевыми фигурами, событиями и организационными структурами, делая анализ более понятным и наглядным, обеспечивая эффективные стратегии для управления ситуацией.

Заключение

В заключение, говоря детально о технологиях комплексного анализа содержимого Сети при поиске информации дискредитирующего характера, рассмотрим наиболее эффективные на сегодняшний день инструменты:

Для анализа текстов с целью обнаружения манипулятивных технологий, включая DeepFake и другие формы искажений, предлагается, на наш взгляд, наиболее точная нейронная сеть BERT (Bidirectional Encoder Representations from Transformers) и ее производные [10, 11]. Она активно используется для задач классификации текста, анализа тональности, обнаружения фейковых новостей и манипулятивных текстов.

При аудиальном анализе применяются модели типа Wav2Vec 2.0 и OpenAI Whisper для расшифровки речи и выявления манипулятивной интонации. Для анализа видео используются EfficientNet, XceptionNet и MesoNet, которые показывают высокую результативность в обнаружении визуальных манипуляций. Обучение названных моделей производится на специализированных наборах данных: для текста – LIAR (корпус фейковых новостей) и PAN datasets, для аудио — VoxCeleb и Librispeech, а для видео — DeepFake Detection Challenge (DFDC) и FaceForensics++.

Для автоматизированного извлечения данных в социальных сетях среди наиболее эффективных инструментов иностранного происхождения можно выделить PhantomBuster и Octoparse, а также Netlytic, используемые для анализа текстов и связей между пользователями. Для интеграции с Twitter используется специализированная библиотека Tweepy, с Facebook – Facebook Graph API, для общего веб-скрейпинга – Scrapy. Для анализа сложных динамических страниц и создания парсеров, способных работать с закрытыми группами, эффективен инструмент Selenium. Библиотека BeautifulSoup применяется для обработки HTML-структур, что делает её удобной для работы с контентом, собранным из нестандартных источников.

При работе с социальными сетями и мессенджерами в России эффективно применяются отечественные разработки, такие как поисковая система SEUS – программный комплекс для мониторинга, поиска и анализа информации. Применяется для развития безопасного информационного пространства и защиты российского общества от деструктивного воздействия.

Интеграция рассмотренных подходов и технологий позволит не только эффективно выявлять и предотвращать угрозы дискредитирующих воздействий в обществе, но и создавать условия для противостояния их распространению в социальных медиа. В условиях роста информационных вызовов и усложнения цифровых систем такие подходы становятся незаменимыми инструментами для защиты репутации, обеспечения доверия к медиа информации.

Таким образом, цифровизация средств коммуникации произвела революцию в процессах наблюдения за социальными и политическими событиями, а также реагирования на них. Быстро растущее влияние социальных медиа стало использоваться политиками, служащими различных структур, обычными гражданами для распространения деструктивной информации, включая процессы дискредитации. Эта статья посвящена исследованию влияния социальных медиа на отношения между людьми и организациями через эффект дискредитации. Социальные

медиа гражданам не только позволяют напрямую общаться, но и, к сожалению, активизируют участие масс в дискредитации в форме обратной связи через комментарии на сайтах социальных медиа. Деструктивная информация при этом играет важную роль в формировании имиджа той или иной компании, того или иного гражданина. Для исследования этого явления весьма конструктивен метод математического моделирования, который позволяет не только проводить латентно-структурный анализ явления дискредитации, но и идентифицировать статистически ненаблюдаемые параметры и переменные, свойственные ему. Результаты показывают, что исследования системы дискредитации могут стать шагом вперед для нейтрализации этого крайне негативного феномена, эффективно используя платформу социальных медиа.

Список литературы

1. Минаев В.А., Бондарь К.М., Рабчевский А.Н., Федорович В.Ю. Противодействие экстремистской идеологии в социальных медиа: математические модели и методы: Монография / Под ред. В.А. Минаева, К.М. Бондаря. Хабаровск: РИО ДВЮИ МВД России, 2023. 232 с.
2. Бондаренко Е. Н., Анализ речевой стратегии дискредитации в лингвистической экспертизе (на примере интернет-комментария) // Филологические науки. Вопросы теории и практики. 2014. № 10 (40). Ч. 1. С. 27-29.
3. Кихтан В. В., Мамиева Б. Ю. К вопросу о манипулировании в современных СМИ // Вестник Волжского университета имени В. Н. Татищева. 2018. № 2. URL: <https://cyberleninka.ru/article/n/k-voprosu-o-manipulirovanii-v-sovremennyh-smi> (дата обращения: 5.12.2024).
4. Fatema S., Yanbin L. and Fugui D. Social media influence on politicians' and citizens' relationship through the moderating effect of political slogans // *Frontiers in Communication*, 2022. С. 2-21.
5. Васильев Э. А., Дьяченко Н. Н. Субъекты дискредитации сотрудников органов внутренних дел // Юридическая наука и практика: Вестник Нижегородской академии МВД России. 2019. № 4 (48). С. 56-60.
6. Президентская кампания Хиллари Клинтон (2016). URL: [https://ru.wikipedia.org/wiki/Президентская_к_ампания_Хиллари_Клинтон_\(2016\)](https://ru.wikipedia.org/wiki/Президентская_к_ампания_Хиллари_Клинтон_(2016)) (дата обращения 5.12.2024).
7. Federal Prosecution of Donald Trump (classified documents case). URL: [https://translated.turbopages.org/proxy_u/en-ru.ru.0a3a3912-675aaa2f-5458b0cc-74722d776562/https/en.wikipedia.org/wiki/Federal_prosecution_of_Donald_Trump_\(classified_documents_case\)](https://translated.turbopages.org/proxy_u/en-ru.ru.0a3a3912-675aaa2f-5458b0cc-74722d776562/https/en.wikipedia.org/wiki/Federal_prosecution_of_Donald_Trump_(classified_documents_case)) (дата обращения 5.12.2024).
8. Volkswagen Emissions Scandal. URL: https://translated.turbopages.org/proxy_u/en-ru.ru.3f92f82e-675aaa84-1539d6c8-74722d776562/https/en.wikipedia.org/wiki/Volkswagen_emissions_scandal (дата обращения 5.12.2024).
9. Kobe Steel. URL: https://en.m.wikipedia.org/wiki/Kobe_Steel (дата обращения 15.12.2024).
10. Минаев В. А., Поликарпов Е. С., Симонов А. В. Применение глубинных нейронных сетей для выявления деструктивного контента в социальных медиа // *Информация и безопасность*. 2021. Т. 24. № 3. С. 361-372.
11. Минаев В.А., Симонов А.В. «Просеивание» телеграмм-каналов при поиске контента экстремистского характера // *Информация и безопасность: научный журнал*. 2023. Том 26. № 1. С. 25-30.

Московский университет МВД РФ им. В.Я. Кикотя
Moscow University of the Internal Affairs Ministry of Russia named after V. Ya. Kikot

Дальневосточный институт МВД РФ им. И.Ф. Шилова
Far Eastern Institute of the Internal Affairs Ministry of Russia named after I.F. Shilov

Поступила в редакцию 6.12.2024

Информация об авторах

Минаев Владимир Александрович – д-р техн. наук, профессор, профессор кафедры специальных информационных технологий, Московский университет МВД РФ им. В.Я. Кикотя, Москва, e-mail: m1va@yandex.ru

Бондарь Константин Михайлович – канд. техн. наук, доцент кафедры информационно-технического обеспечения Дальневосточного института МВД РФ им. И.Ф. Шилова, Хабаровск, e-mail: bondar_km@mail.ru

Коломина Анна Сергеевна – адъюнкт кафедры специальных информационных технологий, Московский университет МВД РФ им. В.Я. Кикотя, Москва, e-mail: ann.kolomina@mail.ru

TECHNOLOGIES FOR DEFAMATORY INFORMATION SEARCHING IN SOCIAL MEDIA

V.A. Minaev, K.M. Bondar, A.S. Kolomina

In article identifying, analyzing and preventing threats to discredit organizations and specific individuals using social media are considered. Attention is drawn to the use of mathematical methods and models in describing the processes of defamatory information propagation. The definition is given, the main forms are discussed and examples of discrediting are given. The role of social media in the system of discrediting effects and technological features of media exposure are shown. Big Data is discussed, representing structured and unstructured arrays of large amounts of data, as well as modern processing tools for analysis, forecasting and management decision-making; neural networks for finding the meaning of information, including the identification of manipulative effects on humans; visualization of the relationships between key figures, events and organizational structures, providing an understanding of effective strategies for managing the situation. It is concluded that in the context of the growth of information challenges and the complexity of digital, the considered approaches become indispensable for protecting reputation and ensuring trust in media information. It is concluded that in the context of increasing information challenges and the complexity of digital applications, the considered methods and models are becoming indispensable tools for protecting reputation and ensuring trust in media information.

Keywords: search technologies, defamatory information, social media, modeling.

Submitted 6.12.2024

Information about the authors

Vladimir A. Minaev – Dr. Sc. (Technical), Professor, Professor of the Special Information Technologies Department of the Moscow University of the Internal Affairs Ministry of Russia named after V. Ya. Kikot, e-mail: m1va@yandex.ru

Konstantin M. Bondar – Cand. Sc. (Technical), Associate Professor of the Information and Technical Support Department Far Eastern Institute named after of the I.F. Shilov of the Internal Affairs Ministry of the Russian Federation, Khabarovsk, e-mail: bondar_km@mail.ru

Anna S. Kolomina – Postgraduate Student of the Special Information Technologies Department of the Moscow University of the Internal Affairs Ministry of Russia named after V. Ya. Kikot, Moscow, e-mail: ann.kolomina@mail.ru