

МИНИСТЕРСТВО ВНУТРЕННИХ ДЕЛ
РЕСПУБЛИКИ КАЗАХСТАН

АЛМАТИНСКАЯ АКАДЕМИЯ
ИМЕНИ МАКАНА ЕСБУЛАТОВА

**ИССЛЕДОВАНИЕ СОВРЕМЕННЫХ МЕТОДОВ
ДЕАНОНИМИЗАЦИИ ПОЛЬЗОВАТЕЛЕЙ
В СОЦИАЛЬНЫХ СРЕДАХ**
Методические рекомендации

Алматы
2025

Обсуждено и одобрено на заседании научно-методическом совете Алматинской академии МВД Республики Казахстан им. М. Есбулатова (протокол №5 от «15» мая 2025 года)

Рецензенты:

Байшоланова К.С. – Казахский национальный университет им. аль-Фараби, профессор, д.э.н.

Калдыбаев А.Б. – начальник управления противодействия киберпреступлений ДП г. Алматы полковник полиции.

Кадырова Р.Т., Уралова Ф.С., Суюнбай Ж.: Исследование современных методов деанонимизации пользователей в социальных средах: Методические рекомендации. – Алматы: ООНИиРИР Алматинской академии МВД Республики Казахстан им. М. Есбулатова, 2025. – 40 с.

Методические рекомендации посвящены вопросам деанонимизации пользователей в социальных средах и раскрывают актуальные подходы, применяемые для идентификации лиц, скрывающих свою личность в интернет-пространстве. Особое внимание уделяется таким технологиям, как OSINT, стилистический анализ текстов, анализ цифровых следов, а также методы машинного обучения для поведенческой аналитики. Работа рассчитана на широкий круг исследователей, как для ученых, так и для практических работников, а также для курсантов, слушателей, магистрантов и докторантов высших учебных заведений правоохранительных органов.

© Алматинская академия МВД
Республики Казахстан
им. М. Есбулатова, 2025

Введение

В последние годы наблюдается резкий рост анонимной активности в цифровом пространстве, особенно на платформах, таких как Telegram. Участники подобных коммуникаций всё чаще скрывают свою личность, используя фейковые аккаунты, ботов и ссылки, ведущие на подозрительные ресурсы. Это создает риски для безопасности пользователей и препятствует эффективной идентификации нарушителей. В таких условиях особое значение приобретает разработка и внедрение современных методов деанонимизации – инструментов, позволяющих установить личность анонимных пользователей на основе анализа их поведения и цифровых следов.

Применение технологий машинного обучения, обработки естественного языка (NLP), а также OSINT-методов (разведка по открытым источникам) [21] открывает новые возможности для автоматической обработки текстов, выявления скрытых закономерностей и корреляции между действиями пользователей. В совокупности с использованием контейнерной архитектуры (Docker) данные подходы позволяют создавать гибкие системы, работающие в реальном времени и обладающие высокой точностью анализа.

Настоящие методические рекомендации направлены на формирование у обучающихся теоретических знаний и практических навыков по применению цифровых технологий для анализа поведения пользователей в социальных медиа и установления их возможной идентичности. Документ охватывает как концептуальные аспекты анонимности в сети, так и реализацию прикладных решений: от сбора данных до построения аналитических моделей и визуализации результатов.

Рекомендации предназначены для курсантов, магистрантов, слушателей, а также специалистов в области информационной безопасности, цифровой криминалистики и анализа интернет-данных.)

1 Теоретические основы деанонимизации пользователей в цифровых платформах

Деанонимизация – это процесс установления личности пользователя, который намеренно или по умолчанию скрывает свою идентичность в цифровом пространстве. В условиях стремительного развития социальных сетей и мессенджеров данное понятие приобретает особую актуальность, особенно в свете увеличения числа инцидентов, связанных с анонимной активностью. Анонимность в сети может проявляться в разных формах:

- Техническая анонимность – скрытие IP-адреса и других сетевых параметров с помощью VPN, прокси и Тор.
- Псевдонимность – использование вымышленных имен, никнеймов или аватаров вместо реальных данных.
- Поведенческая анонимность – стремление не оставлять уникальных паттернов поведения, по которым можно было бы идентифицировать пользователя.
- Контентная анонимность – размещение информации, не раскрывающей личность (например, обезличенные сообщения, изображения без EXIF-данных и пр.).

Процедуры деанонимизации могут быть направлены как на раскрытие личности конкретного субъекта, так и на выявление связи между аккаунтами одного и того же пользователя. В основе таких подходов лежит анализ цифровых следов – совокупности информации, которую оставляет пользователь при взаимодействии с интернет-ресурсами. Ключевыми направлениями деанонимизации являются:

- анализ поведенческих и временных паттернов,
- стилистическая экспертиза текстов (stylometry),
- сопоставление данных из разных источников (cross-platform correlation),
- визуализация социальных связей.

Современные методы деанонимизации нередко интегрируют возможности машинного обучения и OSINT [37] [38],

что позволяет автоматизировать сбор, сопоставление и интерпретацию информации.

Таким образом, деанонимизация представляет собой междисциплинарную область, сочетающую знания в сфере кибербезопасности, лингвистики, поведенческого анализа и цифровой криминалистики.

1.1 Понятие и типы анонимности в цифровой среде

Анонимность в цифровой среде представляет собой состояние, при котором личность пользователя не может быть напрямую или достоверно установлена другими участниками коммуникации или техническими средствами. Это явление является неотъемлемой частью современного интернета и может рассматриваться как с позитивной стороны – в контексте защиты частной жизни и свободы выражения мнений, так и с негативной – в аспекте нарушения прав, безопасности и осуществления противоправных действий.

Цифровая анонимность формируется на пересечении технологических, поведенческих и юридических факторов. Пользователи могут скрывать свою личность намеренно, например, используя псевдонимы, специализированные инструменты обхода идентификации (VPN, Tor, прокси-серверы), или непреднамеренно – просто не предоставляя конкретных персональных данных при взаимодействии с цифровыми сервисами. Независимо от мотива, анонимность представляет собой ситуацию, в которой уровень доступности сведений о субъекте строго ограничен.

Выделяют несколько основных типов анонимности, различающихся по способу и глубине сокрытия идентичности. Техническая анонимность основывается на использовании инструментов, маскирующих сетевые параметры пользователя. Такие решения позволяют скрыть IP-адрес, геолокацию, сведения об устройстве и другие цифровые идентификаторы.

Это достигается через маршрутизацию трафика через промежуточные серверы, шифрование соединений и отказ от

куки-файлов. Однако даже при высоком уровне технической защиты сохраняется возможность косвенного установления личности за счёт сопоставления поведенческих данных.

Псевдонимность является наиболее распространённой формой анонимности в социальных платформах. Пользователь может действовать под вымышленным именем или никнеймом, не раскрывая своей реальной личности. При этом сохраняется постоянство аккаунта и истории взаимодействий, что позволяет формировать устойчивый цифровой образ. Псевдонимность не исключает возможность деанонимизации, особенно в случаях, когда аккаунт связан с другими учетными записями или содержит элементы личной информации.

Поведенческая анонимность касается особенностей действий пользователя в цифровом пространстве. Каждый человек обладает уникальным стилем взаимодействия с информационными системами: ритмом написания сообщений, характерными ошибками, предпочтениями во времени активности. Эти особенности могут быть использованы как цифровой отпечаток личности и поддаются анализу с помощью методов машинного обучения. Даже при отсутствии имени, IP-адреса или иных прямых признаков, поведенческий след может указать на конкретного человека или группу.

Контентная анонимность проявляется в стиле оформления сообщений, выборе слов, тематики и контекста коммуникации. Стилометрический анализ позволяет определить автора текста или соотнести несколько сообщений между собой, опираясь на лингвистические особенности. Такой подход особенно актуален для раскрытия скрытых аккаунтов в социальных сетях, где автор старается маскировать своё участие.

Наконец, сетевая анонимность рассматривается в рамках структурных связей между аккаунтами. Даже если конкретные пользователи не раскрывают личных данных, их коммуникации, подписки, участие в группах и пересечения с другими учетными записями могут быть проанализированы для построения социальной сети. Графовые методы и алгоритмы выявления сообществ позволяют установить принадлежность

анонимного пользователя к определённой группе или даже конкретной личности.

Таким образом, анонимность в цифровой среде – это многогранное явление, охватывающее не только технические аспекты скрытия данных, но и поведенческие, лингвистические и сетевые характеристики. Понимание типов и уровней анонимности необходимо для разработки эффективных методов деанонимизации и оценки рисков в области информационной безопасности. Особенно важно учитывать, что анонимность не является абсолютной: при наличии достаточного объёма данных и соответствующих аналитических инструментов её можно нарушить, не прибегая к прямому доступу к персональной информации [39].

1.2 Социальные платформы как среда для анонимной активности

Современные социальные платформы предоставляют пользователям уникальные возможности для общения, самоидентификации, распространения контента и формирования цифровых сообществ. Вместе с тем, именно эти же платформы становятся пространством для анонимной активности, в том числе связанной с нарушением прав, распространением дезинформации и иными потенциально опасными действиями. Степень анонимности, предоставляемая пользователю, варьируется в зависимости от архитектуры платформы, политики конфиденциальности и уровня модерации [2][3].

Одной из ключевых причин популярности анонимности в социальных сетях является возможность свободного выражения мыслей без страха преследования, стигматизации или цензуры. Тем не менее, высокая степень анонимности часто используется и в противоправных целях: скрытие личности при распространении вредоносной информации, координации незаконной деятельности или манипуляциях в общественном пространстве [9].

Для анализа возможностей анонимной активности целесообразно рассмотреть различные типы платформ (рис. 1) по критерию идентификации пользователя, доступности личных данных и уровня контроля со стороны администрации.

Платформа	Регистрация с номером/почтой	Возможность использования псевдонима	Отображение username	Уровень модерации	Условие деанонимизации
Telegram	Да (телефон)	Да	Да	Низкий	Через OSINT, поведенческий анализ
Reddit	Электронная почта	Да	Да	Средний	Через историю постов и IP
Twitter (X)	Да (почта/телефон)	Да	Да	Средний/высокий	Анализ сети связей, лексики
4chan	Нет	Да (анонимно)	Нет	Очень низкий	Почти невозможна
Facebook	Да (настоящее имя по правилам)	Нет	Да	Высокий	Высокая вероятность

Рис.1. Условия анонимности на популярных платформах

Как видно из рисунка 1, платформы Telegram и Reddit предоставляют пользователям возможность сохранять высокий уровень анонимности за счёт минимальных требований к идентификации личности и отсутствия строгой верификации учетных записей. Это делает такие платформы привлекательными для лиц, желающих скрыть свою идентичность. При этом отсутствие строгой модерации и возможность использования псевдонимов осложняет процессы выявления источника сообщений или реального владельца аккаунта [5].

Telegram является одним из наиболее показательных примеров платформы с условной анонимностью. Несмотря на обязательную регистрацию по номеру телефона, сам номер может быть скрыт от других пользователей, а идентификация часто ограничивается никнеймом. При этом с помощью внешних инструментов – таких как Telegram API, OSINT-анализа,

парсинга открытых групп – возможно построение поведенческого профиля пользователя и выявление признаков повторяющейся активности [40].

На другом полюсе находятся платформы вроде Facebook, где политика «реальных имён» формирует низкий уровень анонимности. Хотя на практике не все пользователи следуют этим правилам, модерация и алгоритмы обнаружения фейковых аккаунтов значительно затрудняют анонимную активность.

Фактор	Описание	Влияние
Наличие обязательной верификации	Требование подтверждения личности или телефона	Снижает
Возможность скрыть метаданные	Прямое или косвенное отражение ID, IP, геолокации	Повышает
Публичность профиля	Доступность истории сообщений и активности	Снижает
Интеграция с другими сервисами	Возможность сопоставления данных с другими платформами	Снижает
Политика модерации	Активное удаление подозрительного или вредоносного контента	Снижает
Возможность использования ботов	Применение автоматических аккаунтов без привязки к личности	Повышает

Рис.2. Факторы, влияющие на степень анонимности

Анализ показывает, что уровень анонимности зависит не только от структуры платформы, но и от поведения самого пользователя. Даже в условиях ограниченной идентификации цифровые следы – в виде времени активности, структуры сообщений, используемой лексики – могут быть использованы для установления личности при наличии соответствующих инструментов анализа [6] [14].

Таким образом, социальные платформы одновременно служат и средством для обеспечения конфиденциальности, и пространством, где злоупотребление анонимностью может приводить к реальным угрозам. Понимание механизмов, влияющих на анонимность, является ключом к разработке методов деанонимизации, особенно в рамках работы по информационной безопасности, цифровой криминалистике и мониторингу открытых источников данных.

2 Методы и инструменты идентификации пользователей

Для решения задач деанонимизации используются современные методы обработки текста, машинного обучения и анализа открытых источников (OSINT). Эти подходы позволяют выявлять пользователей, скрывающих свою личность, на основе их поведения, языка и цифровых следов.

Технологии обработки естественного языка (NLP) [4] применяются для анализа структуры текста, стиля письма, повторяющихся выражений. Это помогает установить связь между разными аккаунтами одного автора или выявить аномалии в сообщениях.

Модели машинного обучения, такие как Random Forest и BERT, используются для автоматической классификации сообщений и оценки их анонимного характера. Эти методы позволяют обнаружить подозрительные сообщения и провести поведенческий анализ [7].

Сбор информации из открытых источников (OSINT) включает анализ username, временных меток, активности и связей между пользователями. Такие данные позволяют строить гипотезы об истинной личности анонимного участника.

Все компоненты системы реализованы с помощью Docker-контейнеров, что обеспечивает гибкость, повторяемость и удобную интеграцию для последующего анализа.

2.1 OSINT и анализ открытых данных

В рамках современного информационного пространства разведка по открытym источникам (OSINT – Open Source Intelligence) представляет собой систематизированный подход к сбору и анализу данных, доступных публично и легально. Использование OSINT является одним из ключевых методов в задачах деанонимизации, так как позволяет идентифицировать пользователя по совокупности его цифровых следов, не прибегая к взлому или несанкционированному доступу к закрытым системам.

OSINT-анализ основывается на принципе мультиканального сопоставления информации: даже если пользователь активно маскирует личность в одном сервисе, его активность в других источниках (форумы, соцсети, базы данных, публичные API) может невольно раскрывать идентифицирующие признаки. Этапы OSINT-анализа [13]:

- Определение целей исследования – формулируется задача: установить личность, найти связь между аккаунтами, проверить подлинность данных и т.д.
- Первичный сбор информации – осуществляется с использованием Telegram API, поисковых операторов, агрегаторов утечек, специализированных OSINT-платформ (например, Spiderfoot, Maltego, Recon-ng).
- Анализ и фильтрация данных – извлечённые данные очищаются от дубликатов, классифицируются по типам (username, email, IP, phone, timestamp и пр.).
- Построение взаимосвязей – с помощью графового анализа выявляются связи между аккаунтами, сервисами и действиями пользователя.
- Интерпретация и документирование результатов – создаётся отчёт, включающий визуализации, ссылки и интерпретации.

Источник	Тип данных	Применение
Telegram API	Username, ID, сообщения, дата	Идентификация активности, временные паттерны
Google Search / Dorks	Имя, email, сайты, утечки	Сопоставление аккаунтов, поиск следов по никнейму или email
Maltego	Связанные сущности, профили	Визуализация связей между учетными записями
ExifTool	Метаданные изображений	Извлечение координат, устройств, времени съемки
HaveIBeenPwned, LeakCheck	Базы утечек	Поиск совпадений по email, логинам
GitHub, Reddit, Twitter	Контент, стиль, активность	Кросс-платформенный анализ поведения и стиля

Рис.3. Примеры источников OSINT-данных в задачах деанонимизации [1].

Пример применения OSINT-анализа – работа с псевдонимом, найденным в Telegram. Имея username, исследователь может провести автоматизированный поиск в Google с помощью операторов («username» site:twitter.com), проверить наличие в базах утечек, проанализировать время и стиль публикаций в разных сервисах. Если пользователь оставлял один и тот же ник на нескольких платформах, создается возможность установить единый цифровой профиль.

Кроме того, в случае наличия мультимедийного контента - фотографий, голосовых сообщений – применяются методы анализа метаданных и лингвистической экспертизы. Например, изображение может содержать EXIF-информацию с точными координатами, моделью устройства и временем создания.

Таким образом, OSINT [22] [23] позволяет не только дополнить классические методы деанонимизации, но и самостоятельно служить мощным инструментом выявления и отсле-

живания цифровых личностей. Его эффективность обусловлена не только объемом доступных данных, но и способностью аналитика интерпретировать связи и выявлять нетривиальные зависимости между, казалось бы, разрозненными источниками информации.

2.2 Технологии машинного обучения и NLP в деанонимизации

Современные методы деанонимизации всё чаще опираются на инструменты машинного обучения (ML) и обработки естественного языка (NLP) [8]. Эти технологии позволяют эффективно анализировать большие массивы текстовых и поведенческих данных, выделять уникальные характеристики пользователей и восстанавливать связи между анонимными цифровыми профилями. Применение ML и NLP способствует созданию интеллектуальных систем, способных не только фиксировать факт анонимной активности, но и выдвигать гипотезы о возможной личности пользователя или группе, к которой он принадлежит [12].

1. Роль машинного обучения в задачах деанонимизации

Машинное обучение представляет собой область искусственного интеллекта, в рамках которой алгоритмы обучаются на исторических данных для того, чтобы делать предсказания или принимать решения без явного программирования. В контексте деанонимизации ML используется для:

- классификации сообщений и поведения пользователей по степени подозрительности;
- определения вероятности принадлежности разных аккаунтов одному субъекту;
- выявления аномалий в цифровой активности;
- распознавания шаблонов поведения или речи, уникальных для конкретных пользователей.

Один из распространённых алгоритмов в задачах классификации – Random Forest. Это ансамблевый метод, основан-

ный на построении множества решающих деревьев, где итоговый результат определяется голосованием. Он показывает высокую точность при работе с табличными данными, что делает его полезным для анализа текстов в векторном представлении (например, после применения TF-IDF). В рамках данной работы модель Random Forest применялась для классификации текстов, полученных из Telegram, по признакам, характерным для анонимной активности: наличие ссылок, стилистика, ключевые слова, временные интервалы публикации [24] [25].

Более сложные и контекстуально чувствительные модели – например, нейросетевой трансформер BERT (Bidirectional Encoder Representations from Transformers) – используются для более глубокого понимания текста. BERT способен учитывать контекст каждого слова в предложении, что делает его особенно полезным при анализе коротких сообщений, типичных для социальных сетей и мессенджеров. Такие модели позволяют выполнять не только классификацию сообщений, но и их кластеризацию, извлечение сущностей (имён, названий, локаций), а также семантический анализ содержания.

2. Применение NLP для анализа стиля и идентификации авторства

Обработка естественного языка (NLP) – это область, объединяющая лингвистику, информатику и искусственный интеллект. Она включает методы анализа текстов, позволяющие «читать» сообщения машиной и извлекать из них значимую информацию. В задачах деанонимизации особое значение имеет стилометрия – анализ стиля письма, направленный на выявление автора текста [15] [26].

Каждый человек, независимо от намерений, использует определённый словарный запас, структуру предложений, пунктуацию, синтаксические обороты. Эти параметры могут быть количественно измерены и использоваться в качестве признаков при построении модели авторства. Например, анализ частотности использования определённых частей речи,

длины слов, использования сокращений и эмодзи, а также ритма написания (если доступны временные метки) позволяет создать уникальный лингвистический профиль пользователя.

NLP также применяется для извлечения именованных сущностей (NER – Named Entity Recognition), то есть распознавания имён, организаций, географических объектов в тексте. Эта информация помогает установить, в каком контексте общается пользователь, о чём пишет, и не указывает ли текст на личные данные. Использование таких методов позволяет не только проанализировать сообщения, но и выделить потенциальные признаки самоидентификации, намеренно или случайно оставленные пользователем [27] [28].

3. Объединение ML и NLP в единую аналитическую систему

Интеграция моделей машинного обучения и NLP позволяет создать мощную аналитическую платформу, которая способна выполнять полный цикл анализа: от обработки необработанных текстов до визуализации результатов и формирования отчётов. В рамках данной работы был реализован следующий процесс [29] [30]:

- Сбор данных: извлечение сообщений из Telegram через API, включая текст, дату, username, user_id.
- Предобработка: очистка текста от лишних символов, токенизация, нормализация, лемматизация.
- Фичеизация: преобразование текстов в векторное представление (TF-IDF, embeddings), извлечение поведенческих признаков (время отправки, частота, длина).
- Классификация: применение моделей Random Forest и BERT для определения категории активности.
- Авторская стилистика: дополнительный NLP-анализ для выявления уникальных языковых характеристик.
- Визуализация: отображение результатов анализа в веб-интерфейсе (например, в виде таблиц, графиков, тепловых карт).

Такой подход позволяет систематически выявлять анонимных пользователей, отслеживать цифровые шаблоны поведения и строить профили, полезные для задач кибербезопасности, расследований или мониторинга нарушений.

Машинное обучение и обработка естественного языка играют центральную роль в современных подходах к деанонимизации. Их применение позволяет перейти от ручного анализа к автоматизированной интерпретации цифровых данных, сохраняя при этом гибкость и масштабируемость. Комплексный анализ текстов, поведенческих признаков и сетевой активности обеспечивает более полное понимание цифрового следа пользователя и повышает точность процедур идентификации в условиях анонимности [16].

2.3 Использование Docker для развертывания аналитической системы

Для обеспечения стабильной и переносимой работы комплексной аналитической системы по деанонимизации пользователей используется технология контейнеризации Docker. Она позволяет упаковать программные модули в изолированные окружения (контейнеры), в которых уже установлены все зависимости и библиотеки. Такой подход существенно упрощает развертывание, обновление и масштабирование системы [10] [11].

Разработка системы в рамках данного исследования предполагала построение архитектуры, включающей модули сбора, анализа и визуализации данных. Все компоненты были объединены в единую инфраструктуру с использованием Docker Compose – инструмента, позволяющего одновременно запускать несколько связанных контейнеров с помощью одного конфигурационного файла.

На изображении ниже представлено подтверждение успешного запуска аналитической системы:

```
C:\Users\User\Downloads\projecttest>docker-compose up -d
[+] Running 7/7
  ✓ Network projecttest_monitoring-network  Created
  ✓ Container mongodb                      Started
  ✓ Container enrichment_script_container Started
  ✓ Container telegram-collector           Started
  ✓ Container threat-analyzer             Started
  ✓ Container telegram-notifier           Started
  ✓ Container dashboard                   Started
```

Рис.4. Успешный запуск контейнеров с помощью команды docker-compose up – d

Как видно из изображения, система состоит из следующих компонентов:

- mongodb – контейнер с базой данных, в которой хранятся сообщения, метаданные и результаты анализа.
- enrichment_script_container – вспомогательный модуль для обогащения данных (например, определение временных меток, языковых характеристик).
- telegram-collector – модуль сбора данных из Telegram через API.
- threat-analyzer – компонент анализа сообщений, включающий модели машинного обучения.
- telegram-notifier – сервис, отправляющий уведомления об обнаруженной активности.
- dashboard – веб-интерфейс для визуализации и мониторинга системы.

Все контейнеры были успешно созданы и запущены, что свидетельствует о корректной конфигурации и совместимости между модулями. Использование Docker обеспечивает:

- гибкость: система может быть запущена на любой платформе с поддержкой Docker;
- масштабируемость: при необходимости можно дублировать или заменять отдельные модули;
- воспроизводимость: одинаковая среда в тестировании, разработке и эксплуатации;

- безопасность: изоляция процессов друг от друга и от основной операционной системы.

Для построения масштабируемой и надёжной инфраструктуры аналитической платформы по деанонимизации пользователей была использована технология контейнеризации Docker. Применение контейнеров позволило разделить функциональные модули системы на изолированные сервисы, каждый из которых выполняет строго определённую задачу. Такой подход обеспечивает повторяемость, простоту развертывания и лёгкость обновления компонентов без необходимости переустановки всей системы [17].

Инфраструктура развёртывается с использованием Docker Compose, который автоматически запускает все модули, указанные в конфигурационном файле. При запуске контейнеров система формирует собственную виртуальную сеть, через которую осуществляется обмен данными между сервисами. Каждый модуль может быть протестирован или масштабирован независимо от других компонентов.

На схеме ниже представлена архитектура аналитической системы:

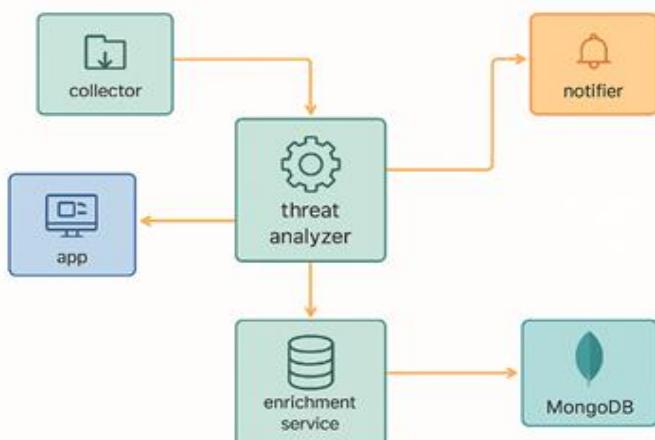


Рис.5. Архитектура контейнерной аналитической системы

Описание компонентов:

- collector – модуль сбора сообщений из Telegram с помощью API, сохраняющий данные для дальнейшего анализа;
- threat analyzer – основной аналитический модуль, использующий модели машинного обучения и NLP для анализа содержания сообщений и выявления признаков анонимной активности;
- enrichment service – вспомогательный сервис, обогащающий данные (например, определяющий временные интервалы активности, категории сообщений и др.);
- MongoDB – база данных, где хранятся тексты сообщений, метаданные и результаты анализа;
- notifier – сервис, отвечающий за отправку уведомлений при выявлении подозрительных сообщений или цифровых шаблонов активности;
- app (dashboard) – веб-интерфейс для визуализации данных, управления фильтрами и экспортом отчётов.

Таким образом, Docker выступает как базовая технология для инфраструктурного уровня системы деанонимизации. Его использование обеспечивает устойчивость, переносимость и ускоряет внедрение в реальные исследовательские или оперативные среды.

3 Практическая реализация системы

В рамках исследования была разработана и развернута аналитическая система, предназначенная для автоматизированного сбора, обработки и анализа сообщений из открытых Telegram-каналов и чатов с целью выявления анонимной активности и потенциально подозрительных пользователей. Система реализована на основе модульной архитектуры с использованием технологии контейнеризации Docker, что обеспечивает её гибкость и устойчивость.

Сбор сообщений осуществляется через модуль telegram-collector, взаимодействующий с Telegram API. Для каждого сообщения извлекаются текст, дата и время, идентификатор

пользователя, `username` (если доступен), а также метаданные, отражающие источник и структуру сообщения. Полученные данные сохраняются в базе MongoDB, работающей в отдельном контейнере.

Следующий этап обработки включает предобработку текста (очистку, токенизацию, нормализацию) и передачу в модуль `threat-analyzer`, где применяются модели машинного обучения – Random Forest и BERT [18][19]. Эти алгоритмы осуществляют классификацию сообщений по признакам, характерным для анонимной или подозрительной активности, а также анализируют лингвистические и поведенческие особенности.

Дополнительно реализован модуль `telegram-notifier`, который в случае обнаружения сообщений, соответствующих заданным критериям, отправляет уведомления администраторам или ответственным лицам. Это повышает оперативность реагирования на потенциальные угрозы [47].

Завершающим компонентом является `dashboard` – веб-интерфейс, разработанный для визуализации результатов анализа. Он отображает сообщения, категории, уровень риска, а также данные об отправителях. Это позволяет исследователям отслеживать динамику активности, фильтровать данные и экспортировать отчёты.

Система прошла тестирование в локальной и облачной среде, продемонстрировала устойчивую работу в реальном времени и может быть адаптирована для использования в задачах цифровой криминалистики, мониторинга социальных медиа и кибераналитики.

3.1 Сбор и хранение данных через Telegram API

Одним из ключевых этапов реализации аналитической системы является автоматизированный сбор данных из Telegram, где активно распространяется анонимный контент. Telegram предоставляет гибкий API-интерфейс, который позволяет получать доступ к публичным сообщениям, информации об авторах,

а также к другим метаданным – без необходимости вмешательства в защищённые каналы передачи данных.

Для доступа к Telegram API [41] необходимо зарегистрировать собственное приложение на официальной платформе my.telegram.org. В результате регистрации пользователь получает уникальные параметры – `api_id` и `api_hash`, которые используются для аутентификации при подключении к серверу Telegram.

Пример конфигурации, используемой в рамках настоящей системы, представлен на рисунке ниже:

App configuration

App api_id:	21730420	🔒
App api_hash:	ea9d67e635316e8394b995de0824e730	🔒
App title:	deanon of fraudsters	
Short name:	deanon	
	alphanumeric, 5–32 characters	

Рис. 6. Конфигурация приложения Telegram API (пример)

После регистрации приложения устанавливается клиентская библиотека, такая как Telethon или Pyrogram. В данной реализации был использован Telethon, как наиболее стабильная и функционально расширенная библиотека на Python.

Модуль `telegram-collector`, запущенный в контейнере Docker, подключается к Telegram через API и подписывается на нужные каналы, группы или чаты. Система извлекает следующие поля:

- `message_id` – уникальный идентификатор сообщения;
- `chat_id` и `chat_title` – данные об источнике;
- `sender_id`, `username`, `phone_number` – идентификаторы автора (если доступны);
- `message_text` – полный текст сообщения;
- `date_time` – временная метка;

– media_type, entities, reply_to, forwarded_from – дополнительные характеристики контекста.

Пример кода для извлечения сообщений с помощью Telethon:

```
from telethon.sync import TelegramClient
client = TelegramClient('session_name', api_id, api_hash)
async def collect_messages():
    async for message in client.iter_messages('target_channel'):
        print(message.sender_id, message.text) [42]
```

Данный процесс может быть легко масштабирован на множество каналов и чатов за счёт списков или фильтров.

Собранные данные сохраняются в базе MongoDB, которая обеспечивает хранение в формате JSON-документов. Каждое сообщение сохраняется как отдельный документ, содержащий вложенные структуры: текст, автор, дата, категории (если применимы), статус анализа и т.д.

Преимущество MongoDB в данном случае заключается в её способности гибко обрабатывать полуструктурные данные, а также в высокой скорости обработки запросов к коллекции сообщений [43].

Пример структуры документа в MongoDB:

```
{
    «message_id»: 1432,
    «text»: «Пожалуйста, перейдите по ссылке для восстановления доступа...»,
    «username»: «@support_fake_bot»,
    «sender_id»: 678901234,
    «date»: «2025-05-15T11:32:00»,
    «source»: «@example_channel»,
    «risk_score»: null,
    «is_phishing»: null,
    «analyzed»: false
}
```

Поскольку Telegram может содержать личные данные, важно ограничивать сбор только открытых сообщений, до-

ступных без подписки или авторизации. Система не обрабатывает личные чаты, приватные группы и сообщения с ограниченным доступом. Все собранные данные используются исключительно в исследовательских целях и хранятся в защищённой среде Docker-контейнера [20].

Таким образом, Telegram API предоставляет все необходимые инструменты для масштабируемого и автоматического сбора сообщений. Полученные данные формируют основу для дальнейшей лингвистической, поведенческой и классификационной обработки в рамках системы анализа цифровой активности и деанонимизации.

3.2 Предобработка текстов и извлечение признаков

Предобработка текстов – это обязательный этап в построении любой системы анализа текстовой информации, особенно в задачах, связанных с классификацией сообщений и идентификацией поведенческих характеристик пользователя. Целью данного этапа является приведение текстов к формату, пригодному для анализа, а также извлечение ключевых признаков (фичей), необходимых для работы алгоритмов машинного обучения [44].

Первоначально собранные сообщения из Telegram содержат большое количество шумовых данных: лишние пробелы, спецсимволы, эмодзи, ссылки, упоминания пользователей, а также различное форматирование. Эти элементы не несут полезной смысловой нагрузки и могут искажать результаты анализа. Поэтому применяются стандартные шаги предварительной обработки:

- Очистка текста: удаление HTML-тегов, ссылок, специальных символов, повторяющихся знаков препинания.
- Нормализация: приведение текста к нижнему регистру.
- Токенизация: разбиение текста на отдельные слова (токены).

– Стоп-слова: удаление часто встречающихся, но неинформативных слов (например, «это», «как», «что»).

– Лемматизация: приведение слов к их базовой форме (например, «платил» → «платить»).

После предварительной обработки текстов производится извлечение признаков – числовых или категориальных характеристик, которые будут использоваться в качестве входных данных для моделей машинного обучения. Наиболее распространённые подходы:

– TF-IDF (Term Frequency – Inverse Document Frequency): отражает важность термина в конкретном документе по отношению к корпусу текстов.

– N-граммы: последовательности из N слов, фиксирующие устойчивые выражения или шаблоны.

– Метафичи: длина сообщения, количество ссылок, наличие ключевых слов, частотность использования символов, а также время и день публикации.

Предобработка и извлечение признаков выполняются в модуле анализа threat-analyzer автоматически. Результаты передаются в модели классификации (Random Forest, BERT) [35] [36] для дальнейшего анализа содержания и поведенческого профиля.

Таким образом, качественная предобработка и фичеизация текста являются фундаментом надёжной аналитической системы в задачах деанонимизации и поведенческой оценки активности пользователей.

3.3 Классификация сообщений и поведенческий анализ

После этапа предобработки текстов и извлечения признаков следующим шагом является классификация сообщений, а также анализ поведенческих характеристик пользователей. Эти процессы направлены на выявление сообщений, содержащих признаки подозрительной активности, и на построение профилей анонимных участников по их цифровому поведению.

Классификация текстов осуществляется с помощью моделей машинного обучения. В данной системе используются две модели:

Random Forest – алгоритм ансамблевого обучения, работающий на основе решающих деревьев. Он обучается на размеченных данных (текстах с метками категорий) и позволяет оценить вероятность принадлежности нового сообщения к определённому классу (например, подозрительное, фейковое, безвредное).

BERT – глубокая нейросетевая модель, учитывающая контекст слов в предложении. Она особенно эффективна для анализа коротких, контекстно-насыщенных сообщений, типичных для социальных платформ. Модель способна понимать смысл фразы и выявлять тональность, намерение и тематическую направленность текста.

В процессе классификации каждое новое сообщение анализируется на предмет совпадения с ранее обученными категориями, и при превышении порогового значения вероятности ему присваивается соответствующая метка. Результаты сохраняются в базе данных и визуализируются в интерфейсе системы [31] [32].

Параллельно с содержательным анализом производится поведенческий анализ пользователя. Он включает в себя:

- временные характеристики активности (время суток, частота сообщений);
- структуру сообщений (длина, наличие ссылок, эмодзи, повторяемость);
- шаблоны взаимодействия с другими участниками (ответы, пересылки);
- аномалии в ритме и стиле публикаций.

На основе этих данных строится поведенческий профиль, позволяющий выявить повторяющиеся схемы и подозрительную активность, даже если сообщения написаны с разных аккаунтов.

Таким образом, совмещение классификации и поведенческого анализа усиливает возможности системы деанонимизации и позволяет выявлять не только конкретные сообщения, но и анонимных участников, ведущих координированную или вредоносную деятельность [46].

3.4 Механизмы деанонимизации пользователей

Деанонимизация пользователей – это процесс установления реальной или частичной идентичности субъектов, скрывающихся за псевдонимами, никнеймами или техническими средствами анонимности в цифровом пространстве. В контексте социальных платформ, таких как Telegram, деанонимизация особенно актуальна из-за высокой степени допуска к анонимной коммуникации.

Механизмы деанонимизации, реализованные в рамках данной системы, основываются на сочетании лингвистического, поведенческого и метаданных-анализа. сбору данными [33] [34].

1. Сопоставление метаданных

Для каждого пользователя собираются технические параметры сообщений: время публикации, ID отправителя, username (если указан), язык сообщения, информация о пересылке или ответе. Сравнение таких параметров в пределах разных чатов позволяет выявлять повторяющиеся шаблоны поведения и подозрительную активность под разными учетными записями.

2. Лингвистическая деанонимизация (стилометрия)

Каждый пользователь характеризуется уникальным стилем письма: выбором слов, пунктуацией, структурой предложений, использованием эмодзи и сокращений. Анализируя эти параметры с помощью инструментов NLP, система формирует языковой профиль, который может быть сопоставлен с другими сообщениями. Таким образом выявляется, принадлежат ли разные тексты одному автору [45].

3. Поведенческая корреляция

Отслеживается активность пользователя по времени суток, дням недели, длине сообщений и тематике. Система выделяет поведенческие паттерны, сравнивает их между собой и строит поведенческие кластеры, что позволяет объединять аккаунты одного пользователя.

4. Кросс-платформенный OSINT-анализ

Username или никнейм проверяются в открытых источниках (поисковики, другие соцсети, утечки данных). Если имя используется на других платформах, где раскрыты личные данные, это даёт дополнительную основу для деанонимизации.

Таким образом, комплексный подход к деанонимизации, основанный на анализе текстов, поведения и открытых данных, позволяет идентифицировать участников, даже если они не раскрывают свою личность напрямую.

3.5 Визуализация результатов через веб-интерфейс

Визуализация результатов играет ключевую роль в аналитических системах, так как обеспечивает удобную интерпретацию больших объёмов данных и позволяет пользователю быстро принимать решения на основе полученной информации. В рамках данного проекта был реализован полнофункциональный веб-интерфейс (Dashboard), разработанный с использованием фреймворка Flask (или FastAPI) и JavaScript-библиотек для графического отображения информации (например, Chart.js, D3.js).

Интерфейс доступен локально по адресу localhost:5000 и представляет собой административную панель для мониторинга и анализа сообщений, полученных из Telegram. Система отображает как агрегированные, так и детализированные данные в виде графиков, диаграмм и таблиц.

На первом экране пользователь получает сводную статистику: общее количество обработанных сообщений, распределение по уровням угроз (высокий, средний, низкий), а также динамику поступления инцидентов.

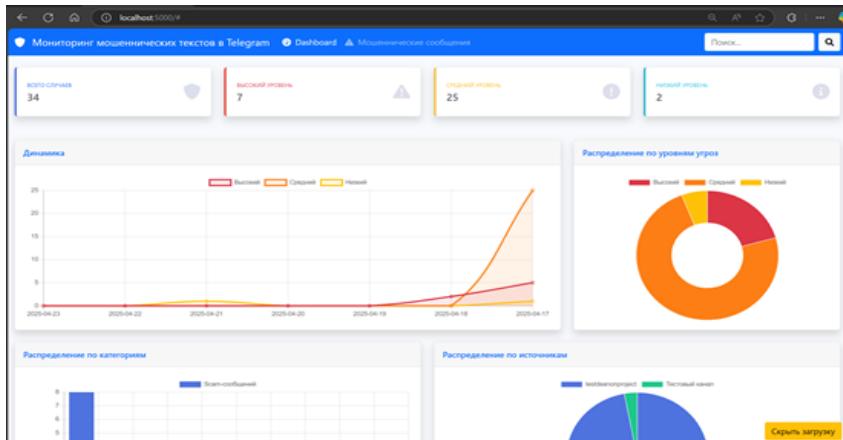


Рис.7. Главная панель аналитической системы

В верхней части отображаются ключевые метрики:

- Всего случаев – общее количество проанализированных сообщений;
- Высокий / Средний / Низкий уровень угрозы – категоризация по степени риска, определённая моделью;
- График динамики – временная кривая роста выявленных сообщений;
- Кольцевая диаграмма – наглядное распределение по уровням угроз.

Такой подход позволяет отслеживать тенденции и пики активности в режиме реального времени.

Следующие визуальные блоки на панели – это диаграммы по категориям и источникам. Система классифицирует сообщения по заранее определённым типам: «инвестиционное мошенничество», «фишинг», «фейковые лотереи», «социальная инженерия» и др.

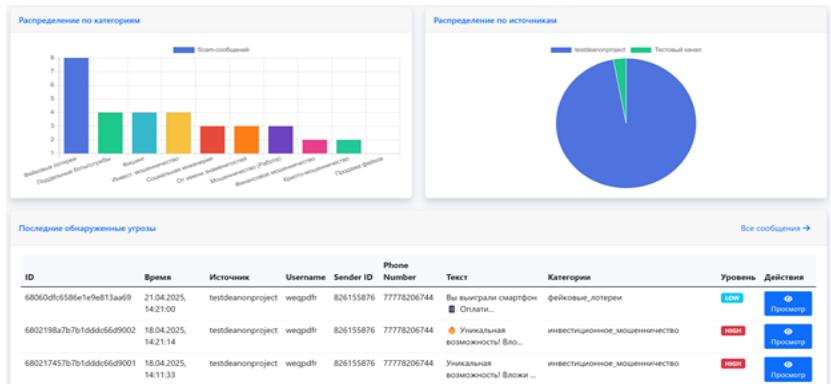


Рис.8. Распределение по категориям и Telegram-источникам

Гистограмма по категориям показывает, какие типы сообщений преобладают в анализируемом потоке.

Круговая диаграмма по источникам демонстрирует относительный вклад различных каналов и чатов (например, testdeanonproject, тестовый канал).

Эти визуализации полезны для выявления целевых источников повышенного риска и фокусировки мониторинга.

Ниже представлены таблицы, содержащие информацию о каждом выявленном сообщении. Пользователь может видеть:

- Время получения;
- Название источника (чат, канал);
- Username и Sender ID;
- Номер телефона (если доступен);
- Категорию и уровень риска;
- Функцию просмотра подробностей.

Мониторинг мошеннических текстов в Telegram

Dashboard

Поиск...

Фильтр по уровню

Список инцидентов

Phone

ID	Время	Источник	Username	Sender ID	Number	Текст	Категория	Уровень	Действия
68010df6516e1e9e813aa9	21.04.2023, 14:21:00	testdeanonproject	wegoff	826155876	77778206744	Вы выиграли смартфон 📱 Отплата доставку (1500)...	фейковые_лотереи	LOW	
6802198a7b7b1dd56669002	18.04.2023, 14:21:14	testdeanonproject	wegoff	826155876	77778206744	⚠️ Уникальная возможность! Вложи 1000\$! и получи...	инвестиционное_мошенничество	MEDIUM	
680217457b7b1dd56669001	18.04.2023, 14:11:33	testdeanonproject	wegoff	826155876	77778206744	Уникальная возможность! Вложи 1000\$! и получи пар...	инвестиционное_мошенничество	MEDIUM	
6801446c698c8d4e5914ca8c	17.04.2023, 23:11:57	testdeanonproject	qututut	642295472	Нет данных	“Икон Макс” [**предлагает] (https://t.me/DivCenter...)	мошенничество_от_ищечи_знакомств	HIGH	
680130724986d4379ae0fc	17.04.2023, 21:46:42	testdeanonproject	daratello	543452636	77076828477	Удомка без опыта 5000\$ в день. Пиши “РАБОТА” 🚀...	мошенничество_с_работой	HIGH	
68013068c4986d4379ae0fb	17.04.2023, 21:46:37	testdeanonproject	daratello	543452636	77076828477	Вы выиграли смартфон 📱 Отплата доставку (1500)...	фейковые_лотереи	HIGH	
68013068c4986d4379ae0fa	17.04.2023, 21:46:32	testdeanonproject	daratello	543452636	77076828477	Радача крепты! Отправь 0.1 ETH — получи 1 ETH. ⚡...	крипто_мошенничество	HIGH	
68012602dc4986d4379ae0f9	17.04.2023, 21:32:45	testdeanonproject	suhanbek	667503074	Нет данных	“Мама, привет. У меня проблемы, телефон забыл.	социальная_инженерия	HIGH	

Рис.9. Таблица обнаруженных сообщений с фильтрацией по уровням риска

Также доступна функция поиска и фильтрации по дате, категории, источнику или уровню угрозы. Это обеспечивает гибкость в работе с данными и позволяет быстро находить интересующие случаи.

При нажатии на кнопку «Просмотр», открывается модальное окно с полной детализацией случая. Помимо основных данных о сообщении, отображается дополнительная информация о предполагаемом пользователе из базы данных (если его удалось идентифицировать по Sender ID или username).

Основная информация

ID: 6800f7f9570842e34accd6c2

Время: 17.04.2025, 17:45:45

Источник: testdeanonproject

Уровень угрозы: **HIGH**

Категории: инвестиционное_мошенничество

Вероятность: 98.0%

Информация о пользователе (из БД)

Имя: Петров Иван Сидорович

Таб. номер: EMP001

Отдел: Отдел Разработки

Должность: Ведущий разработчик

Телефон: +7 (495) 123-45-67 доб. 101

Текст сообщения

Уникальная возможность! Вложи 10000Т и получи гарантированный доход без риска!

Обработанный текст

уникальная возможность вложи Т получи гарантированный доход без риска

Рис.10. Полная информация об инциденте и пользователе

Интерфейс отображает:

- ID сообщения и время;
- Источник, уровень угрозы, категория, вероятность классификации;
- Информация о пользователе (ФИО, отдел, должность, телефон);
- Исходный текст сообщения и его предобработанный вариант.

Такая детализация особенно важна для аналитиков и сотрудников служб безопасности, так как позволяет сопоставить сообщения с базой данных пользователей или расследовать конкретный инцидент.

Заключение

Современное цифровое пространство предоставляет широкие возможности для взаимодействия, однако вместе с этим возрастают и уровень анонимной активности, представляющей потенциальную угрозу информационной безопасности. Платформы с минимальной модерацией и высокой степенью конфиденциальности, такие как Telegram, создают благоприятную среду для сокрытия личности участников и затрудняют реагирование на инциденты.

В данных методических рекомендациях рассмотрены теоретические и прикладные аспекты деанонимизации пользователей в социальных средах. Было проанализировано понятие анонимности, приведены её основные формы и уязвимости, описаны современные методы анализа цифровых следов, включая OSINT-инструменты, технологии машинного обучения и обработку естественного языка.

Реализация подобной системы с использованием контейнерных технологий (Docker) позволяет обеспечить масштабируемость, надежность и гибкость архитектуры. Построенная система прошла апробацию на примере анализа сообщений в Telegram и показала эффективность в задачах выявления анонимных участников и построения их поведенческих профилей.

Представленные материалы имеют важное значение для подготовки будущих специалистов в области цифровой криминастики, информационной безопасности и анализа интернет-данных. Методические рекомендации могут быть использованы как в учебных целях, так и в рамках научных исследований и прикладных проектов, связанных с анализом активности в социальных сетях.

Таким образом, разработанные подходы и предложенные решения вносят вклад в формирование практических навыков по выявлению анонимных угроз и повышению цифровой безопасности в современной информационной среде.

Список использованных источников

1. Vorobyev D., & Kotenko I. (2021). OSINT techniques for social media analysis: A survey. *Security and Communication Networks*. <https://www.hindawi.com/journals/scn/2021/6693845/>
2. Wulczyn E., Thain N., & Dixon L. (2017). Ex machina: Personal attacks seen at scale. *Proceedings of WWW*.
3. Zhong H., Li H., Squicciarini A.C., & Memon N. (2016). Content-driven detection of cyberbullying on the Instagram social network. *IJIS*.
4. Chen Y., Zhou Y., Zhu S., & Xu H. (2012). Detecting offensive language in social media to protect adolescent online safety. *Privacy, Security and Trust Conference*.
5. Zhou, Zhiwei, et al. (2021). Hate speech detection on social media: A review. *Multimedia Tools and Applications*.
6. Zampieri M., et al. (2019). Predicting the type and target of offensive posts in social media. *Proceedings of NAACL*.
7. Davidson T., Warmsley D., Macy M., & Weber I. (2017). Automated hate speech detection and the problem of offensive language. *ICWSM*.
8. Basile V., et al. (2019). SemEval-2019 Task 5: Multilingual detection of hate speech against immigrants and women in Twitter. *SemEval*.
9. Boyd, Danah M., & Ellison, Nicole B. (2007). Social Network Sites: Definition, History, and Scholarship. *Journal of Computer-Mediated Communication*.
10. Suler J. (2004). The online disinhibition effect. *CyberPsychology & Behavior*.
11. Brenner, Susan W. (2007). The privacy privilege: The cybersecurity implications of anonymous communication. *Journal of Law, Technology & Policy*.
12. Balduzzi M., Platzer C., Holz T., Kirda E., & Kruegel C. (2010). Abusing social networks for automated user profiling. *RAID*.

13. Vorobyev D., & Kotenko I. (2021). OSINT techniques for social media analysis: A survey. *Security and Communication Networks*.
14. Дьяконов В.А., Ефимова И. В. (2020). Распознавание агрессии и токсичности в социальных сетях с использованием нейросетей. Информационные технологии и телекоммуникации.
15. Карпова О.Н., Петров И.В. (2021). Анализ текстовой агрессии в социальных медиа. Вестник НГУ. Серия: Информационные технологии.
16. Кузнецова Н.В. (2022). Методы и инструменты анализа пользовательской агрессии в интернете. Научно-технические ведомости СПбГПУ.
17. Cheng L., Shu K., Wu S., Silva Y.N., Hall D.L., & Liu H. (2020). Unsupervised Cyberbullying Detection via Time-Informed Gaussian Mixture Model. arXiv preprint arXiv:2008.02642. <https://arxiv.org/abs/2008.02642>
18. Dadvar M., & Eckert K. (2018). Cyberbullying Detection in Social Networks Using Deep Learning Based Models; A Reproducibility Study. arXiv preprint arXiv:1812.08046. <https://arxiv.org/abs/1812.08046>
19. Bajjatiya P., Gupta S., Gupta M., & Varma V. (2017). Deep Learning for Hate Speech Detection in Tweets. arXiv preprint arXiv:1706.00188. <https://arxiv.org/abs/1706.00188>
20. Malik J.S., Qiao H., Pang G., & van den Hengel A. (2022). Deep Learning for Hate Speech Detection: A Comparative Study. arXiv preprint arXiv:2202.09517. <https://arxiv.org/abs/2202.09517>
21. Oakley Browne T., Abedin M., & Chowdhury M.J.M. (2023). A systematic review on research utilising artificial intelligence for open source intelligence (OSINT) applications. *Artificial Intelligence Review*. <https://arxiv.org/abs/2202.09517>
22. Evangelista J.R.G., et al. (2021). Systematic literature review to investigate the application of open source intelligence (OSINT) with artificial intelligence. *Journal of Applied Security Research*, 16(3), 345-369. https://link.springer.com/chapter/10.1007/978-3-031-45237-6_5

23. Azeez N.A., et al. (2021). Cyberbullying detection in social networks: Artificial intelligence approach. *Journal of Cyber Security and Mobility*, 10(4), 745-774. https://link.springer.com/chapter/10.1007/978-3-031-45237-6_5
24. Gomez C.E., Sztainberg M.O., & Trana R.E. (2022). Curating cyberbullying datasets: A human-AI collaborative approach. *International Journal of Bullying Prevention*, 4(1), 35-46. https://link.springer.com/chapter/10.1007/978-3-031-45237-6_5
25. Deliri S., & Albanese M. (2015). Security and privacy issues in social networks. In *Data management in pervasive systems* (pp. 195–209). Springer. https://link.springer.com/chapter/10.1007/978-3-031-45237-6_5
26. Zhou Z., et al. (2021). Hate speech detection on social media: A review. *Multimedia Tools and Applications*. <https://link.springer.com/article/10.1007/s11042-021-11034-0>
27. Basile V., et al. (2019). SemEval-2019 Task 5: Multilingual detection of hate speech against immigrants and women in Twitter. *Proceedings of SemEval*. <https://aclanthology.org/S19-2007/>
28. Davidson T., Warmsley D., Macy M., & Weber I. (2017). Automated hate speech detection and the problem of offensive language. *ICWSM*. <https://ojs.aaai.org/index.php/ICWSM/article/view/14955>
29. Brenner S.W. (2007). The privacy privilege: The cybersecurity implications of anonymous communication. *Journal of Law, Technology & Policy*. <https://illinoisjltc.com/journal/wp-content/uploads/2013/10/Brenner.pdf>
30. Balduzzi M., Platzer C., Holz T., Kirda E., & Kruegel C. (2010). Abusing social networks for automated user profiling. *RAID*. https://link.springer.com/chapter/10.1007/978-3-642-15512-3_4
31. Aniello Castiglione, Roberto De Prisco, Alfredo De Santis, Ugo Fiore, and Francesco Palmieri. 2014. A botnet-based command and control approach relying on swarm intelligence. *Journal of Network and Computer Applications* 38 (2014), 22-33.
32. Klingberg S. (2022). Digital Policing of Open Source Intelligence and Social Media Using Artificial Intelligence. In

Artificial Intelligence and National Security (pp. 101-111). Springer. https://link.springer.com/chapter/10.1007/978-3-031-06709-9_6

33. Cheng L., et al. (2020). Unsupervised Cyberbullying Detection via Time-Informed Gaussian Mixture Model. arXiv preprint arXiv:2008.02642. <https://arxiv.org/abs/2008.02642>

34. Dadvar M., & Eckert K. (2018). Cyberbullying Detection in Social Networks Using Deep Learning Based Models; A Reproducibility Study. arXiv preprint arXiv:1812.08046. https://link.springer.com/chapter/10.1007/978-3-031-06709-9_6

35. Bajjatiya P., et al. (2017). Deep Learning for Hate Speech Detection in Tweets. arXiv preprint arXiv:1706.00188. <https://arxiv.org/abs/2008.02642>

36. Malik J.S., et al. (2022). Deep Learning for Hate Speech Detection: A Comparative Study. arXiv preprint arXiv:2202.09517. <https://arxiv.org/abs/1812.08046>

37. Oakley Browne T., et al. (2023). A systematic review on research utilising artificial intelligence for open source intelligence (OSINT) applications. Artificial Intelligence Review. <https://arxiv.org/abs/1706.00188>

38. Evangelista J.R. G., et al. (2021). Systematic literature review to investigate the application of open source intelligence (OSINT) with artificial intelligence. Journal of Applied Security Research, 16(3), 345-369. <https://arxiv.org/abs/2202.09517>

39. Azeez N.A., et al. (2021). Cyberbullying detection in social networks: Artificial intelligence approach. Journal of Cyber Security and Mobility, 10(4), 74-774. <https://link.springer.com/article/10.1007/s10207-024-00868-2>

40. Gomez C.E., et al. (2022). Curating cyberbullying datasets: A human-AI collaborative approach. International Journal of Bullying Prevention, 4(1), 35-46. https://link.springer.com/chapter/10.1007/978-3-031-45237-6_5

41. Официальный сайт Telegram API: <https://core.telegram.org/api>

42. Официальная документация Pyrogram: <https://docs.pyrogram.org/>

43. Vernit Garg and Laxmi Ahuja. 2019. Password Guessing Using Deep Learning. In 2019 2nd International Conference on Power Energy, Environment and Intelligent Control (PEEIC). IEEE, 38-40.
44. Dongqi Han, Zhiliang Wang, Ying Zhong, Wenqi Chen, Jiahai Yang, Shuqiang Lu, Xingang Shi, and Xia Yin. 2020. Practical traffic-space adversarial attacks on learning-based nidss. arXiv preprint arXiv:2005.07519 (2020).
45. Yisroel Mirsky and Wenke Lee. 2021. The creation and detection of deepfakes: A survey. ACM Computing Surveys (CSUR) 54, 1 (2021), 1-41.
46. Gavai, Gaurang, et al. «Detecting insider threat from enterprise social and online activity data» Proceedings of the 7th ACM CCS international workshop on managing insider security threats. 2015.
47. Zhou, Qingyu, et al. «Neural document summarization by jointly learning to score and select sentences» arXiv preprint arXiv:1807.02305 (2018).

Содержание

Введение	3
1 Теоретические основы деанонимизации пользователей в цифровых платформах	4
1.1 Понятие и типы анонимности в цифровой среде	5
1.2 Социальные платформы как среда для анонимной активности	7
2 Методы и инструменты идентификации пользователей	10
2.1 OSINT и анализ открытых данных	11
2.2 Технологии машинного обучения и NLP в деанонимизации	13
2.3 Использование Docker для развертывания аналитической системы	16
3 Практическая реализация системы	19
3.1 Сбор и хранение данных через Telegram API	20
3.2 Предобработка текстов и извлечение признаков	23
3.3 Классификация сообщений и поведенческий анализ	24
3.4 Механизмы деанонимизации пользователей	26
3.5 Визуализация результатов через веб-интерфейс	27
Заключение	32
Список использованных источников	33

Верстка:
Туренова Б.Ю.

Отдел организации научно-исследовательской и редакционно-издательской работы Алматинской академии МВД Республики Казахстан имени М. Есбулатова 050060 Алматы, ул. Утепова, 29

Подписано в печать 02 июня 2025 г.
Формат 60x84 1/16 Бум. тип. №1. Печать на ризографе. Уч.-изд. п.л. 1,3.
Тираж 50 экз. Уч.-изд. п.л. 2.
Тираж 50 экз.