

УДК 004

АВТОМАТИЗИРОВАННАЯ ИНФОРМАЦИОННАЯ СИСТЕМА ОБНАРУЖЕНИЯ КЛЮЧЕВЫХ СЛОВ В ЗАДАННЫХ ЧАТАХ МЕССЕНДЖЕРА TELEGRAM

Пачкин Е.Д., студент гр. ПИБ-232, II курс

Научный руководитель: Прокопенко Е.В., к.ф.-м.н., доцент

Кузбасский государственный технический университет имени Т.Ф. Горбачёва
г. Кемерово

В современном мире нарастают угрозы экстремизма и терроризма в цифровом пространстве, особенно в мессенджерах, таких как Telegram. В последние годы наблюдается рост использования этих платформ экстремистскими группами для распространения идеологии, вербовки новых последователей и координации действий.

Несмотря на то, что Telegram предоставляет пользователям возможность анонимного общения, это также создает серьезные трудности для правоохранительных органов и служб безопасности, которые пытаются отслеживать и предотвращать террористическую деятельность. Шифрование сообщений, закрытые группы и большое количество контента затрудняют мониторинг и анализ информации, что может привести к серьезным последствиям для общества. По запросу компетентных государственных структур создается далее описываемый продукт.

В частности, время расходуется на проверку открытых пользовательских чатов и анализ отправляемых в них сообщений, их контекста отправки. Преимущества разработки информационной системы по проверке сообщений в чатах мессенджера Telegram имеет ряд преимуществ:

1. Снижение трудозатрат человека и влияния человеческого фактора.
2. Автоматизированный анализ существующих чатов.
3. Кратное увеличение охвата анализируемой информации.
4. Обработанная и структурированная информация по соответствующим лицам.
5. Редкое администрирование в удобное для управляющего лица время с возможностью отслеживания статуса исполнения его действий и примерным сроком их окончания

Telegram – один из самых популярных мессенджеров, где ежедневно обмениваются миллионы сообщений. В открытых группах и каналах участники обсуждают новости, учебу, работу и множество других тем, среди которых могут возникать и запрещенный законодательством РФ. Объем информации в таких чатах огромен, и отследить всё вручную практически невозможно. Здесь на помощь приходит **автоматизированная**

информационная система мониторинга, которая способна анализировать переписку без постоянного участия человека.

Например, существует крупный чат в Telegram с десятками тысяч сообщений в день. Один модератор или аналитик физически не сможет прочитать каждое сообщение и вовремя отреагировать на обнаруженный запрещенный контент. Ручной труд не только отнимает массу времени, но и подвержен влиянию человеческого фактора: человек может устать или пропустить что-то важное. Автоматизированная система решает эту проблему, беря на себя рутинный мониторинг и первичную обработку сообщений. Это снижает трудозатраты и минимизирует влияние человеческого фактора, обеспечивая более стабильный и быстрый контроль над ситуацией.

Цель работы: создание автоматизированной информационной системы по анализу ранее отправленных и новых отправляемых сообщений в заданных чатах мессенджера Telegram с приоритетом на снижение трудозатрат лица, принимающего решения по управлению системой, для снижения влияния угроз экстремизма и терроризма в цифровом пространстве.

Заказчик: ГУ МВД России по Кемеровской области – Кузбассу.

Главное преимущество такой системы – **снижение трудозатрат человека**. Вместо команды людей, круглосуточно перечитывающих чаты, работу выполняет программа. Это означает, что люди освобождаются от монотонной рутины. Система самостоятельно собирает новые сообщения из чатов, фильтрует их по заданным критериям и выделяет ту информацию, которая действительно требует внимания.

Минимизация **человеческого фактора** означает уменьшение вероятности ошибки или субъективной оценки. Автоматизированный мониторинг действует по четко заданным алгоритмам: он не отвлекается и не устает. При установленном маркере отслеживания определенных групп ключевых слов, система сработает каждый раз, когда встретит их в найденном сообщении. Человек-модератор мог бы пропустить такое сообщение в потоке текста или из-за усталости, а машина – нет. В результате повышается надежность контроля: снижается риск, что такое сообщение останется незамеченным.

Кроме того, уменьшение ручного труда напрямую экономит время и ресурсы. То, на что у человека ушли бы часы, программа обработает за считанные минуты. Таким образом, достигается **повышение эффективности**: меньше людей тратят время на мониторинг, при этом качество отслеживания только возрастает.

Автоматизированная система мониторинга способна **анализировать существующие чаты** Telegram в реальном времени. Система разработана на языке программирования Python состоит из нескольких компонентов:

- **Интеграция с Telegram:** Специальный пользовательский аккаунт подключается к нужным чатам через Telegram API, для этого используется библиотека эмуляции пользовательских действий Telethon

(рис. 1). Он добавляется как невидимый наблюдатель, откуда будет получать все новые сообщения, либо проверять ранее отправленные. Для удобного управления пользовательским ботом используется привычный Telegram bot, написанный посредством использования библиотеки aiogram (рис. 2).

- **Хранение данных:** Все собранные сообщения сохраняются в базе данных. Это позволяет не только отслеживать новые сообщения, но и накопить историю переписки для дальнейшего анализа. Для этого выбрано средство управления базами данных PostgreSQL, подключающееся через ORM-технологию посредством библиотеки sqlalchemy и asynccpg.
- **Модуль анализа текста:** Этот компонент обрабатывает поступающие сообщения. Он может искать заданные ключевые слова или фразы. Для корректности одновременно используется как механизм стемминга слов, так и лемматизация в сообщениях и ключевых заданных фраз посредством встроенных возможностей библиотеки ruromgru3 и ускорение анализа текста через усовершенствованный механизм работы библиотеки flashtext .
- **Система оповещений:** Когда в чат приходит сообщение, представляющее интерес, система автоматически уведомляет ответственных лиц. Это выполняется путем мгновенное уведомления в приватный канал администратора со всей подробной информацией и пользователю.

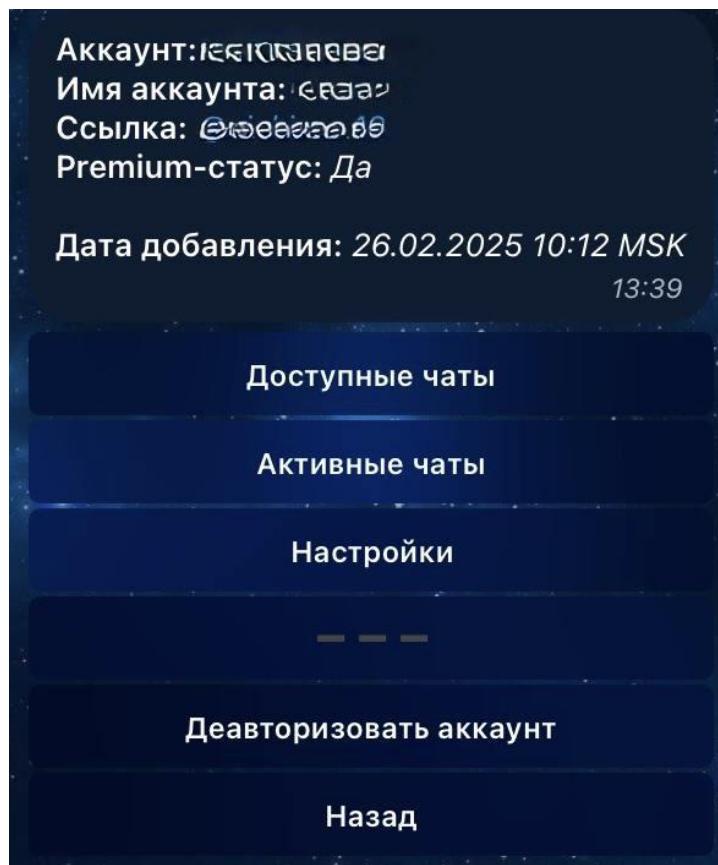


Рисунок 1 – Пользовательский аккаунт

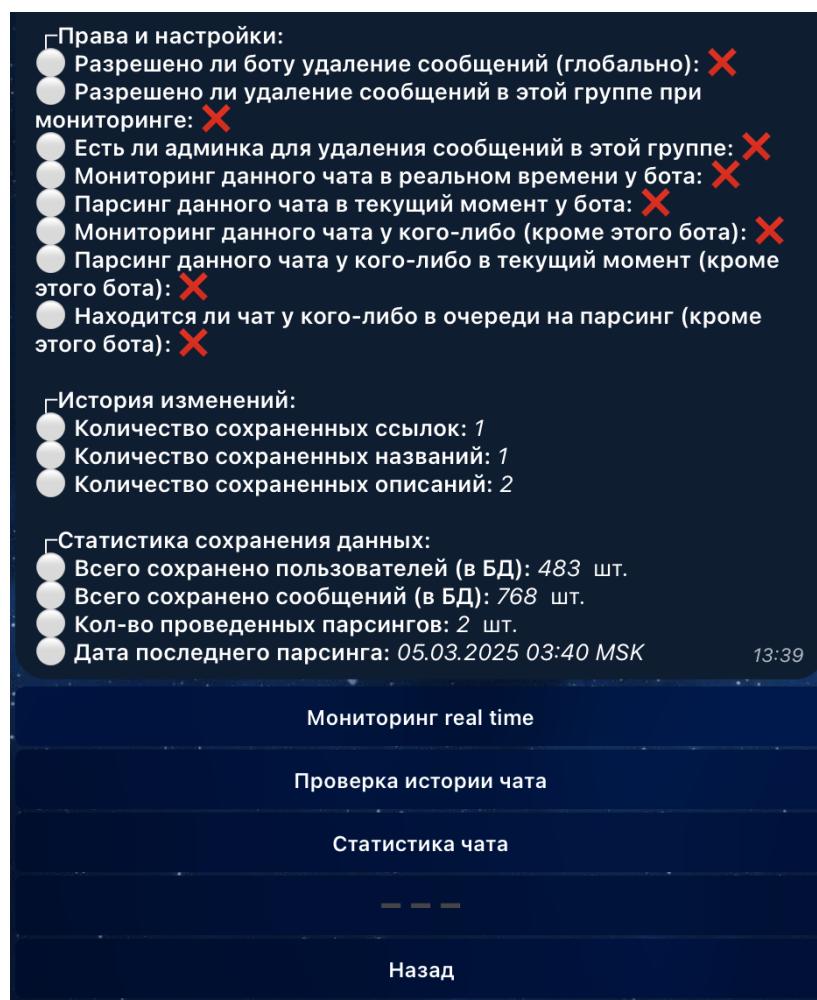


Рисунок 2 – Параметры группы

Благодаря этому **автоматизированному анализу** все существующие чаты, подключенные к системе, непрерывно просматриваются и оцениваются. Система может собирать статистику, в том числе, сколько сообщений написал каждый участник, какие ключевые слова он употреблял, выявлять тенденции.

Одно из ключевых преимуществ автоматизированной системы — **кратное увеличение охвата анализируемой информации**. В то время как человек способен одновременно качественно следить лишь за ограниченным числом чатов, программа легко масштабируется. Она может мониторить десятки или даже сотни групп одновременно, не теряя качества анализа. Представьте, что нужно отслеживать информацию сразу в 50 тематических чатах — вручную это практически невыполнимая задача, а для компьютера разница между одним и пятьюдесятью чатами не принципиальна, нужно лишь больше вычислительных ресурсов. В результате автоматизация дает **кратный рост охвата**: охватываются все необходимые чаты, сколько бы их ни было.

Автоматизированная система мониторинга Telegram спроектирована так, чтобы требовать **минимального администрирования**. Это значит, что после первичной настройки она не нуждается в постоянном ручном

управлении. Управляющее лицо, администратор системы, может время от времени проверять ее работу или корректировать настройки, но нет необходимости следить за ней весь день или каждый день. Все рутинные действия: сбор сообщений, обновление базы, фильтрация, уведомления, – выполняются при наступлении соответствующих событий.

На данный момент система находится в опытной эксплуатации.

Список литературы:

1. Шелудько, В. М. Язык программирования высокого уровня Python. Функции, структуры данных, дополнительные модули: учебное пособие / В. М. Шелудько. – Ростов-на-Дону, Таганрог: Издательство Южного федерального университета, 2017. – 107 с. – ISBN 978-5-9275-2648-2. – Текст: электронный // Электронно-библиотечная система IPR BOOKS: сайт. – URL: <http://www.iprbookshop.ru/87530.html> (дата обращения: 22.02.2024). – Режим доступа: для авторизир. Пользователей.
2. Федоров, Д. Ю. Программирование на языке высокого уровня Python : учебное пособие для прикладного бакалавриата / Д. Ю. Федоров. – 2-е изд., перераб. И доп. – Москва : Издательство Юрайт, 2019. – 161 с. – (Бакалавр. Прикладной курс). – ISBN 978-5-534-10971-9. – Текст: электронный // ЭБС Юрайт [сайт]. – URL: <https://urait.ru/bcode/437489> (дата обращения: 15.02.2025).
3. Светозарова, Г.И. Практикум по программированию на алгоритмических языках / Г.И. Светозарова, Е.В. Сигитов, и др.. – М.: Наука, 2018. – 320 с.
4. Бэрри, Пол Изучаем программирование на Python / Пол Бэрри. – М.: Эксмо, 2023. – 897 с.
5. Савельева, Н. В. Основы программирования. Курс лекций / Н.В. Савельева. – М.: Интернет-университет информационных технологий, 2021. – 264 с.
6. Алекс, Дэвис Асинхронное программирование. Руководство / Дэвис Алекс. - М.: ДМК Пресс, 2018. - 311 с.
7. Лебедев, С. Н. Алгоритмы и структуры данных на Python / С. Н. Лебедев. – Екатеринбург: Уральский университет, 2021. – 450 с. – ISBN 978-5-7996-1234-7. – Текст: электронный // ЭБС УРФУ: сайт. – URL: <https://www.urfu.ru/> (дата обращения: 28.01.2025).
8. Демиденко, А. Telegram Bot. Руководство по созданию бота в мессенджере Телеграм. / А. Демиденко. – Москва : Автор, 2023. – 29 с. – ISBN 978-5-04-542280-2.
9. Бейдер Д. Чистый Python. Тонкости программирования для профи / Д. Бейдер. – Спб : Питер, 2022. – 288 с.