Telegram-бот для автоматизированного анализа файлов и ссылок на наличие вредоносного контента и фишинговых угроз

Актуальность

Telegram широко используется для обмена файлами и ссылками, что делает его удобной средой для распространения вредоносного ПО и фишинговых ресурсов. Большинство существующих Telegram-ботов для проверки файлов имеют ограничение 20 МБ, из-за чего невозможно анализировать архивы, установщики и другие потенциально опасные файлы.

Цель проекта

Создание Telegram-бота, обеспечивающего автоматизированную проверку файлов и ссылок на наличие вредоносного и фишингового контента до взаимодействия пользователя с объектом.

Задачи

- Проанализировать угрозы, распространяемые через файлы и ссылки в Telegram.
- Изучить существующие решения для проверки файлов и ссылок.
- Разработать Telegram-бота для автоматизированного анализа файлов и ссылок.
- Реализовать механизм проверки файлов размером более 20 МБ.
- Интегрировать внешние сервисы анализа и провести тестирование.

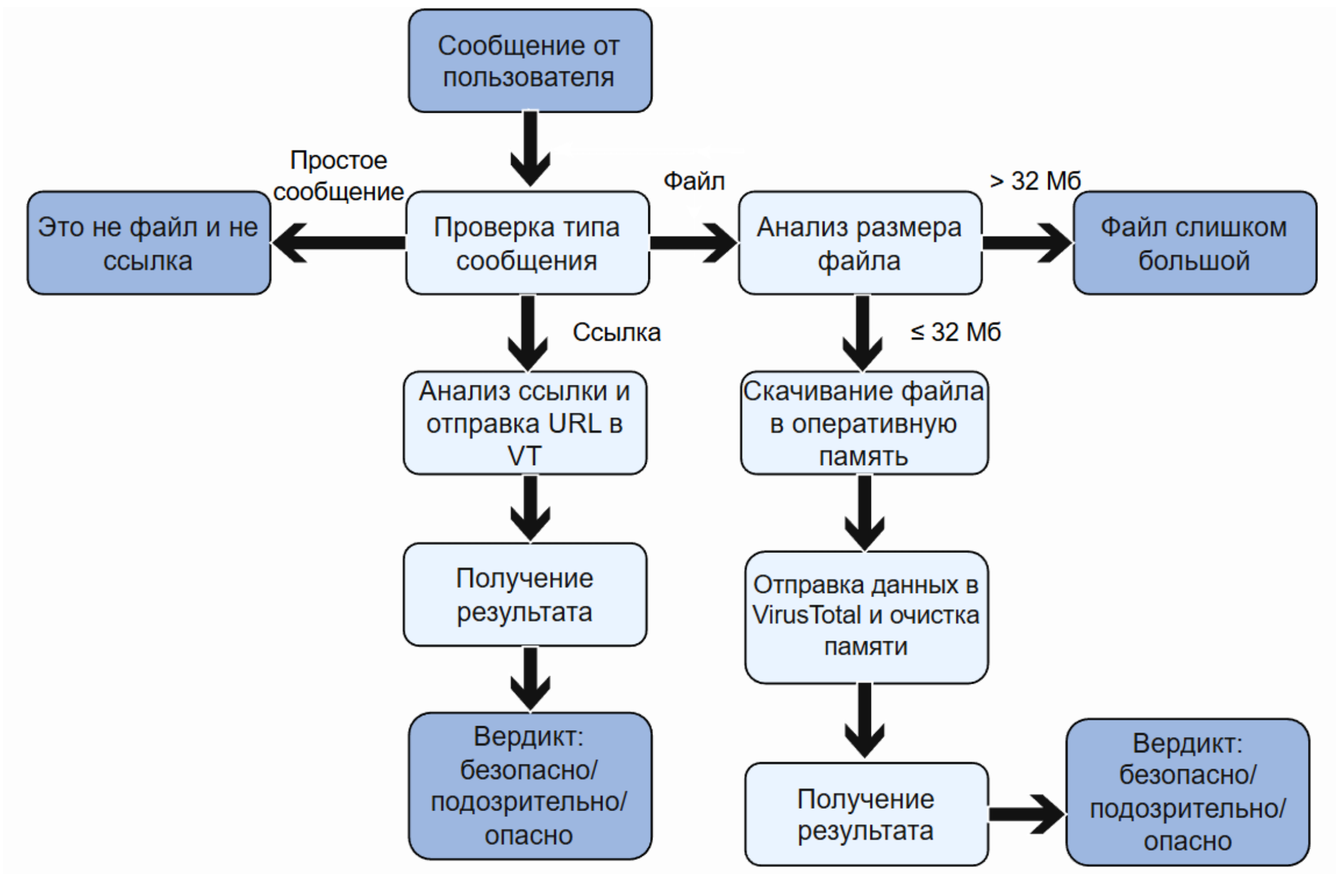
Используемые инструменты

- *Python 3* - основной язык разработки
- *Pyrogram* - реализация логики бота, обработка команд и работа с протоколом MTProto.
- *VirusTotal API* - сервис по выявлению вредоносного и фишингового контента.
- *Aiohttp* - обеспечение асинхронного взаимодействия с внешними сервисами.

Сравнение с аналогами

Критерий сравнения	Dr.Web Bot (Официальный)	@VirusTotal_AV_bot (Популярный аналог)	Разрабатываемый проект
Максимальный размер файла	20 МБ	320 МБ	32 МБ (Лимит Free API VT)
Прозрачность и доверие	Высокое (Известная компания)	Низкое. Риск утечки данных владельцу бота.	Высокое (Open Source). Код открыт, токены у владельца.
Приватность в группах	Читает все сообщения	Читает все сообщения. Риск пассивного сбора переписки.	Privacy Mode ON. Читает только сообщения с командой /scan.
Риск утечки данных	Низкий	Экстремально высокий. Риск скрытого сохранения файлов третьим лицом.	Нулевой. Файлы обрабатываются строго в оперативной памяти.
Формат отчета	Краткий вердикт	Спам-список (перегрузка информацией)	Краткий вердикт
Точность анализа	Средняя (1 движок)	Нестабильная (Пропуски угроз)	Высокая (Актуальный API v3)

Алгоритм работы



Интерфейс бота

Тестирование продукта

№	Тип объекта	Описание сценария	Ожидаемый результат	Фактический результат
1	URL	Легитимная ссылка на общедоступный ресурс	Вердикт: «Безопасно»	Успешно (Безопасно)
2	URL	Свежая фишинговая ссылка из базы OpenPhish	Вердикт: «Опасно» или «Подозрительно»	Успешно (Опасно)
3	Файл	Тестовый файл EICAR (имитатор вируса)	Вердикт: «Опасно»	Успешно (Опасно)
4	Файл	Безопасный текстовый документ (docx)	Вердикт: «Безопасно»	Успешно (Безопасно)
5	Файл	Файл размером 45 МБ (превышение лимита)	Отказ в проверке, сообщение об ошибке	Успешно (Сообщение о лимите 32 МБ)
6	Файл	Файл размером 28мб (обычные боты не смогли бы обработать файл)	Проверка пройдет успешно и файлу будет вынесен вердикт	Проверка прошла успешно, вынесен вердикт
7	Система	Стресс-тест API: отправка 5 файлов подряд за 10 секунд (превышение квоты Free API)	Бот не падает, на 5-м файле выдает уведомление о лимите	Успешно (Сообщение: « ⏳ Лимит API превышен»)

Ссылка на бота (в Telegram)

Ссылка на репозиторий (GitHub)

База знаний (MITRE ATT&CK)

T1566 - Phishing. Распространение вредоносных ссылок и файлов через Telegram с использованием доверия участников чата.

T1204 - User Execution. Активация угрозы происходит только после действий пользователя, что позволяет предотвратить атаку на этапе доставки контента.