Business Email Compromise Detection: Решение трека от Group-IB

By DR Team:

Вахрушев Дмитрий, Мухаметзянов Ренас

Постановка задачи кратко

Построить алгоритм выявления компрометации электронной почты.

Общий подход

- Рассматриваем задачу как задачу детекции аномалий
- А точнее как Novelty detection
- Эмбеддинг текста + модель для Novelty detection

Этапы работы

1. Предобработка данных. (Токенизация, лемматизация, удаление стоп-слов)
2. Векторное представлений документов, feature engineering
3. Подбор моделей.
4. Оценка качества моделей
5. Разработка интерфейса

Архитектура решения

Предобработка

1. Парсинг email
2. Удаление знаков пунктуации, специальных символов (вроде \n)
3. Удаление стоп-слов
4. Лемматизация

Векторное представление

Начали с Doc2Vec для векторного представления письма.

Позже перешли на LaBSE (Language-agnostic BERT Sentence Embedding)

Также использовали признаки: есть ли в письме html разметка, есть ли в письме ссылки

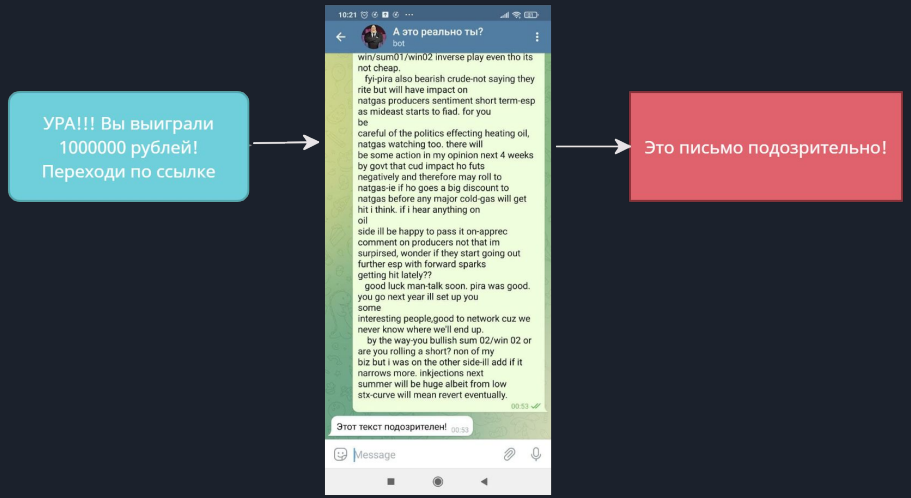
Модели и их качество

1. OneClassSVM
2. Local Outlier Factor

–

Интерфейс для взаимодействия - Telegram бот

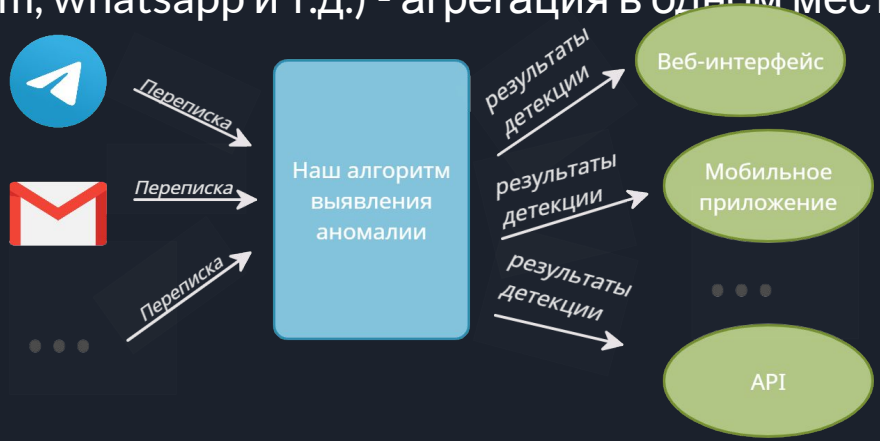
- http://t.me/compomise_detection_bot
- Удобный и простой канал взаимодействия



SCAN ME

Бизнес сторона решения и развитие продукта

- Социальная инженерия - самый эффективный вид взлома . Технических способов защиты от неё на данный момент немного
- Защита деловой переписки на различных платформах (email, telegram, whatsapp и т.д.) - агрегация в одном месте



Бизнес сторона решения и развитие продукта

- Модель монетизации: подписка с размером оплаты в зависимости от количества сотрудников в организации (при этом использование продукта условно-бесплатно до какого-то количества пользователей)
- Дополнительные сервисы по аналитике переписки

Спасибо за внимание!