Инженерный вестник Дона, №8 (2023)
ivdon.ru/ru/magazine/archive/n8y2023/8625

Извлечение текстовой информации из изображений модифицированного текста

Г.И. Мисюков

Финансовый университет при Правительстве Российской Федерации

Аннотация: В данной статье описывается разработка метода по извлечению текстовой информации при передаче изображений с модифицированными символами, которые помогают обходить существующие средства защиты информации (далее - СЗИ) и выводить чувствительную к распространению информацию во вне организации. В основе разрабатываемого модуля используется язык программирования python с библиотеками, которые расширяют его функционал. В статье описаны процесс подготовки данных и создания алгоритма, вариант размещения внутри предприятия и интеграция с СЗИ, а также предложены дальнейшие шаги развития модуля.
Ключевые слова: Информационная безопасность, утечка информации, анализ текста, анализ изображений, анализ модифицированной информации, стеганография

Введение

Количество утечек информации с каждым годом стремительно растет, за период 2022 год общее число утечек в мире возросло более, чем в 3,5 раза [1, 2]. Доля умышленных нарушений также не отстает и растет с каждым годом, в 2022 году более 2/3 утечек информации, случившихся по вине сотрудников, носили умышленный характер.

Наиболее популярными каналами утечки являются следующие [3]:

- почта;
- веб-почта;
- мессенджеры;
- съемные носители;
- личные устройства.

Каждый из этих каналов уже умеют контролировать и отслеживать передачу не является особо сложным, однако, и сами злоумышленники не стоят на месте придумывая новые способы фишинга пользователей [4] при попытке украсть информацию извне или стенографировать информацию [5] при краже изнутри предприятия.

К современным методам стеганографии можно отнести посимвольную замену, увеличивание межсимвольных интервалов и сокрытие изображений в аудиодорожках.

После получения чувствительной к распространению информации следующим шагом для злоумышленника является ее передача. Для скрытой передачи чувствительной информации существует множество методов и средств, самый простой и популярный способ – посимвольная замена. Посимвольная замена преимущественно используется при текстовой передаче и легко обнаруживается при наличии механизма считывания вводимого через клавиатуру текста или, например, механизма извлечения информации из буфера обмена. При создании скриншота отправляемого текста злоумышленником – шанс на обнаружение подобной передачи сводится к 0, поскольку ни одна из существующих систем не поддерживает транскрипцию многосимвольной замены текста, либо же просто не распознает подобные символы.

Стоит отметить, что некоторая модификация текста может быть обработана, однако слабым местом будет являться то, что многие системы защиты работают на транслитерации по ГОСТ Р 7.0.34-2014 «Правила упрощенной транслитерации русского письма латинским алфавитом». Помимо транслитерации по ГОСТ, число опечаток или намеренных ошибок должно быть не более 20% от длины слова, а также вводятся дополнительные требования к минимальной длине слова. В связи с этим и появляется потребность в создании системы, способной считывать передаваемый текст, как в виде текстовой информации как таковой, так и при считывании текста с картинок.

Таким образом можно утверждать, что на данный момент нет технологии, позволяющей однозначно определять извлекаемый текст, если он был изменен, поскольку символы попросту не будут сопоставляться с

существующей базой. На выходе мы получим бессмысленный текст, на который не среагирует ни одна политика безопасности.

Дополнительным обоснованием актуальности разработки системы распознавания модифицированного текста является статистика по вовремя не отслеженным инцидентам информационной безопасности. Возвращаясь к статистике, можно увидеть, что порядка 60% связанных с утечкой информации по описанным ранее каналам имели в себе вложения, 70% из которых составляли изображения, а порядка 15% изображений имели внутри себя измененные символы [6, 7], что не позволило вовремя отследить и предотвратить дальнейшее распространение информации.

Описание технологии извлечения нестандартных символов текста из изображений текста

Перед описанием создаваемой технологии сперва следует исследовать и определить слабые места текущих. СЗИ, направленным на контроль передачи и непосредственно предотвращающим утечки, являются DLP-системы, среди отечественных можно выделить «InfoWatch Traffic Monitor», «Solar Dozor», «СёрчИнформ КИБ» и «Гарда Предприятие», все они способны извлекать текст из изображений и анализировать его по существующим политикам, на основе одного из двух OCR-движков (при необходимости можно использовать и другие, но в таком случае надо согласовывать с вендором) – это:

- ABBY FindReader;
- Tesseract.

Оба этих движка позволяют извлекать текст, однако специфика их работы направленна на осмысленный текст без ошибок и изменений, что негативно влияет на процесс обработки при измененных символах.

Как видно на практике (рисунок 1, 2), при отправке изображения с измененными символами Система никак не реагирует на имеющийся в нем текст и, как следствие, не извлекает его для дальнейшей текстовой отправки.

ķØНФuΔęнцUâλЬНыę дãнНыΣ

Рисунок 1. Модифицированный текст

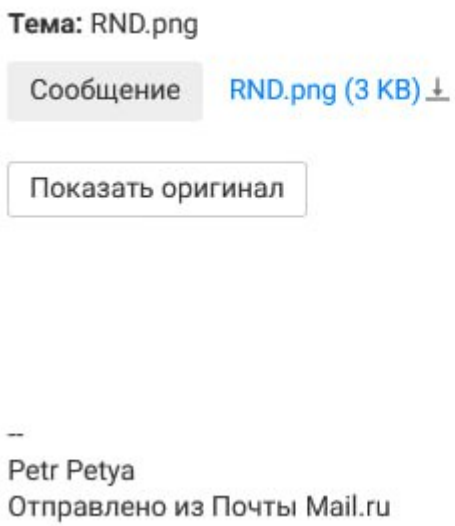

Рисунок 2. Поле извлеченного текста

Для генерации текстов с измененными символами создадим модуль, преобразующий вводимый текст на текст с альтернативными символами, обращаясь к каждому отдельному символу по его индексу и заменяя на похожий символ из соответствующего списка, этот модуль необходим для полноты проверки конечного результата с максимально большим набором предложений.

После чего приступим к созданию модуля, который позволит извлекать модифицированный текст на базе Python с использованием библиотек OpenCV, PIL, Matplotlib, Pandas, Numpy и Typing.

На основе открытых источников создана база модифицированных символов.

Процесс работы алгоритма состоит в следующем:

1. С использованием библиотеки CV2, происходит определение отдельных символов в перехваченном изображении (рисунок 3).

Рисунок 3. Определение символов

2. Извлекаем каждый отдельный символ, используя библиотеку PIL и границы вырезаемого объекта на основе данных из прошлого шага (рисунок 4).

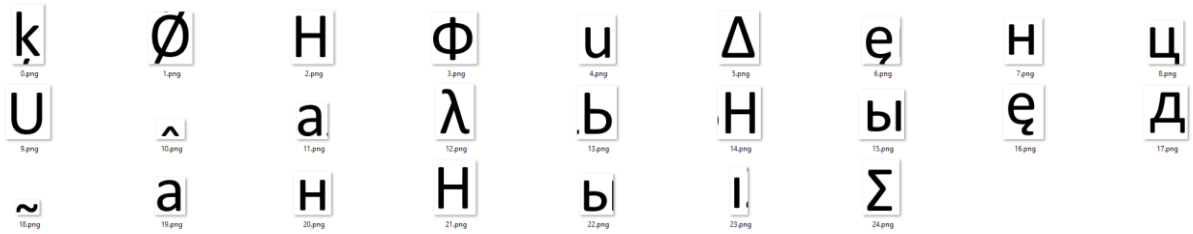

Рисунок 4. Извлеченные символы

3. Избавляемся от символов недостаточного размера, обращаясь к данным о его размере в пикселях используя библиотек PIL.
4. Производим анализ каждого символа на сопоставление с базой данных символов (поиск наиболее вероятностно схожего), посредством использования функции matchTemplate библиотеки CV2.
5. Читаем символ первой буквы названия изображения наиболее вероятностно схожего символа и добавляем символа в переменную типа «строка».
6. Выводим полученное из названия изображение текста в файл.
7. Отправляем файла через постфикс на специальный почтовый ящик в виде bcc-копии.
8. Анализируем перехваченное DLP-системой письмо.

Архитектурная интеграция

Архитектурная интеграция может быть реализована в двух режимах:

1 – Агентское решение [8], анализирующее ввод с клавиатуры и изучающее буфер обмена на предмет наличия в нем текста, чувствительного к распространению (в таком случае будет использован разработанный на этапе тестирования описанной ранее системы модуль для посимвольной автозамены текста).

2 – Шлюзовый перехват писем [9] и прочих отправляемых изображений и удерживание отправки до момента извлечения текста и вынесения вердикта на основе политик защиты информации и упоминаемых в извлеченном тексте слов.

Преимущество первого варианта – блокировка передачи сообщения, но только текстового, поскольку время обработки изображения повлияет на общее увеличение «времени отправки» и вызовет подозрения, изображения же будут передаваться на сервер обработки, после чего администратору СЗИ придет сообщение о нарушении политики и создании инцидента.

Второй же вариант позволяет реализовать блокировку на отправку почтовых писем посредством анализа копии почтового трафика и удержанием основного письма до момента вынесении вердикта о легитимности текста в теле письма и текста, извлеченного из изображений, потенциальным сигналом о СЗИ, в таком случае будет потенциальный получатель письма, который впоследствии просигнализирует о том, что письмо не было получено.

Блокировать передачу в мессенджерах, к сожалению, не выйдет, однако мы все еще можем отлавливать передаваемый текст и вложения, после чего их анализировать и детектировать человека, ответственного за распространение чувствительной информации.

Архитектурная схема с компонентами проектируемой СЗИ при передаче информации вовне в шлюзовом варианте с анализом почтового трафика, представлена на рисунке 5.

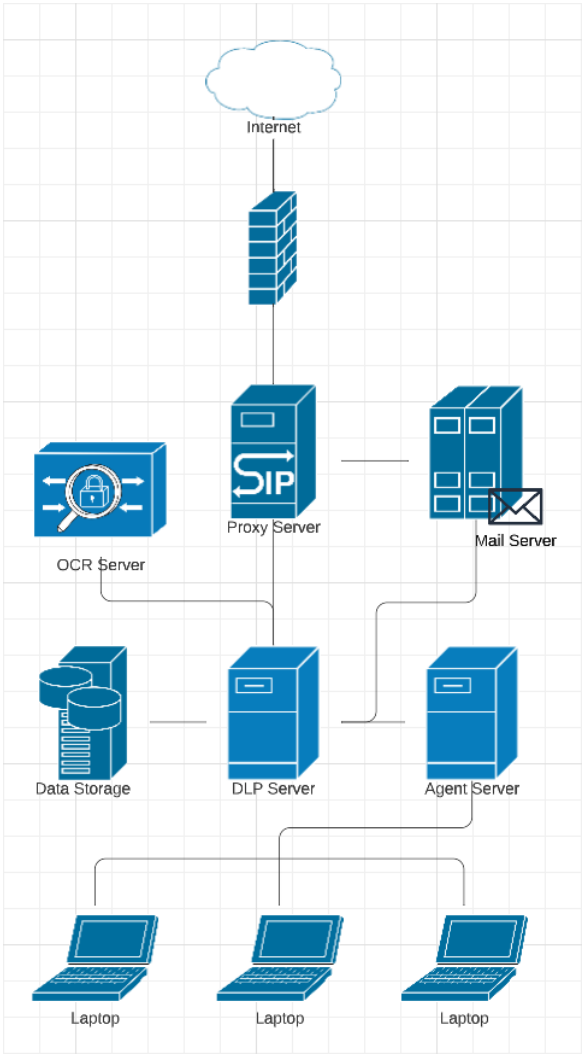

Рисунок 5. Архитектура системы в шлюзовом варианте

Происходит модификация на извлечение метаданных и запись в тело письма (при отсутствии прямого шифрования). В итоге, при отправке сообщения в виде теневой копии в системе мы будем видеть адрес отправителя, адрес получателя, название вложений (отсутствие извлеченного текста движком существующей DLP-системы) и прикрепленное текстовое сообщение с извлеченным на собственном OCR-сервере (рисунок 6).

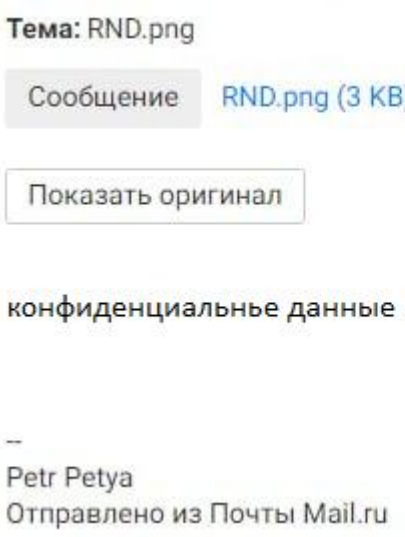

Рисунок 6. Передача сообщения на сервер DLP-системы

Заключение

В ходе тестирования на сгенерированных упомянутым ранее механизмом посимвольной замены текстах, система показала достаточно хороший результат, успешно извлекая в среднем 97% символов (тестирование проводилось на однострочных предложениях с длиной от 3 до 20 слов), учитывался не процент слов без ошибок, а процент правильно извлеченных символов, стоит учитывать, что для обработки по политикам достаточно 80% правильности слова. Поскольку это тестовый вариант работы и интеграции модуля, то он способен обрабатывать только однострочные выражения, что формирует одну из ветвей развития данной технологии. Вторым существенным вектором развития будет расширение библиотеки схожих символов: при упоминании неизвестных в изображении символов (неизвестными будут считаться символы, для которых не было подходящего замещающего, то есть наибольшее совпадение с символом из текущего атласа меньше 80%) будет добавлен механизм анализа изображений. Учитывая специфику обработки изображений незначительных

по размеру, стоит учитывать их специфику при графической обработке и в таком случае воспользоваться методами среднеквадратического отклонения пикселей одного изображения от другого и методом сравнения изображений по пикселям [10].

Литература

1. TADVISER: Утечки данных в России. URL: tadviser.ru/index.php/Статья:Утечки_данных_в_России (дата обращения: 08.06.2023).

2. КАСЛ ЦЛС Прогресс: Статистика утечки гостайны и конфиденциальных данных. URL: licenziya-fsb.com/statistika-utechki-gostajny-i-konfidentsialnyh-dannyh (дата обращения: 08.06.2023).

3. InfoWatch: Глобальное исследование утечек конфиденциальной информации в первом полугодии 2019 года. URL: infowatch.ru/sites/default/files/report/analytics/russ/Global_Data_Leaks_Report_2019_half_year.pdf (дата обращения: 10.06.2023).

4. Афанасьева Н.С., Елизаров Д.А., Мызникова Т.А., Классификация фишинговых атак и меры противодействия им // Инженерный вестник Дона. 2022. №5. URL: ivdon.ru/uploads/article/pdf/IVD_33__5_Afanaseva_Elizarov_Myznikova.pdf_c3fcdcfefd.pdf

5. Subhedar, M. S., & Mankar, V. H.. Current status and key issues in image steganography: A survey. Computer Science Review, 2014, pp: 95–113.

6. Nouh M., Almaatouq A., Alabdulkareem A., Vivek K. S., Shmueli E., Alsaleh M., Alarifi A., Alfaris A., Pentland A.. Social Information Leakage: Effects of Awareness and Peer Pressure on User Behavior. 2014. URL: link.springer.com/content/pdf/10.1007/978-3-319-07620-1_31.pdf (дата обращения 12.06.2023)

7. InfoWatch: Технологии InfoWatch для анализа и защиты информационных активов компании. URL: infowatch.ru/sites/default/files/products/tme/infowatch_technology_booklet_web.pdf (дата обращения: 12.06.2023)

8. SecurityLab: Агентская DLP: взгляд на проблему изнутри. URL: securitylab.ru/blog/company/securityinform/119654.php (дата обращения: 12.06.2023).

9. Prowell, S., Kraus, R., & Borkin, M.. Man-in-the-Middle. Seven Deadliest Network Attacks, 2010, pp: 101–120.

10. Филиппов М.Ю., Королев И.Д. Методика сравнения малоинформативных изображений Инженерный вестник Дона. 2023. №1. URL: ivdon.ru/ru/magazine/archive/n3y2023/8287

References

1. TADVISER: Utechki dannykh v Rossii [Data leakage in Russia]. URL: tadviser.ru/index.php/Stat'ya:Utechki_dannykh_v_Rossii (Date accessed: 08.06.2023).

2. KASL TSLS Progress: Statistika utechki gostayny i konfidentsial'nykh dannykh [Statistics of the leakage of state secrets and confidential data]. URL: licenziya-fsb.com/statistika-utechki-gostajny-i-konfidentsialnyh-dannyh (Date accessed: 08.06.2023).

3. InfoWatch: Global'noye issledovaniye utechek konfidentsial'noy informatsii v pervom polugodii 2019 goda [Global Privacy Leak Survey H1 2019]. URL: infowatch.ru/sites/default/files/report/analytics/russ/Global_Data_Leaks_Report_2019_half_year.pdf (Date accessed: 10.06.2023).

4. Afanas'yeva N.S., Yelizarov D.A., Myznikova T.A. Inzhenernyj vestnik Dona. 2022. №5. URL: ivdon.ru/uploads/article/pdf/IVD_33__5_Afanaseva_Elizarov_Myznikova.pdf_c3fcdcfefd.pdf.

5. Subhedar, M. S., Mankar, V. H.. Current status and key issues in image steganography: A survey. Computer Science Review. 2014, pp. 95–113.

6. Nouh M., Almaatouq A., Alabdulkareem A., Vivek K. S., Shmueli E., Alsaleh M., Alarifi A., Alfaris A., Pentland A.. Social Information Leakage: Effects of Awareness and Peer Pressure on User Behavior. 2014. URL: link.springer.com/content/pdf/10.1007/978-3-319-07620-1_31.pdf (Date accessed: 12.06.2023).

7. InfoWatch: Tekhnologii InfoWatch dlya analiza i zashchity informatsionnykh aktivov kompanii [InfoWatch technologies for analysis and protection of information assets of companies]. URL: infowatch.ru/sites/default/files/products/tme/infowatch_technology_booklet_web.pdf (Date accessed: 12.06.2023)

8. SecurityLab: Agentskaya DLP: vzglyad na problemu iznutri [Agent DLP: a look at the problem from the inside]. URL: securitylab.ru/blog/company/securityinform/119654.php (Date accessed: 12.06.2023).

9. Prowell, S., Kraus, R., & Borkin, M.. Man-in-the-Middle. Seven Deadliest Network Attacks, 2010, pp. 101–120.

10. Filippov M.YU., Korolev I.D. Inzhenernyj vestnik Dona. 2023. №1. URL: ivdon.ru/ru/magazine/archive/n3y2023/8287