004

Практическая работа

Информационно-аналитические технологии поиска угроз инорфмационной безопасности

Исследование метаданных DNS трафика

Цель работы

1. Зекрепить практические навыки использования языка программирования R для обработки данных
2. Закрепить знания основных функций обработки данных экосистемы `tidyverse` языка R
3. Закрепить навыки исследования метаданных DNS трафика

Общая ситуация

Вы исследуете подозрительную сетевую активность во внутренней сети Доброй Организации. Вам в руки попали метаданные о DNS трафике в исследуемой сети. Исследуйте файлы, восстановите данные, подготовьте их к анализу и дайте обоснованные ответы на поставленные вопросы исследования.

Задание

Используя программный пакет dplyr, освоить анализ DNS логов с помощью языка программирования R.

Ход работы

Для выполнения предложенного задания Вам необходимо последовательно проделать следующие шаги:

Подготовка данных

1. Импортируйте данные DNS – https://storage.yandexcloud.net/dataset.ctfsec/dns.zip

> **Что за формат данных**
>
> Данные были собраны с помощью сетевого анализатора **zeek**

2. Добавьте пропущенные данные о структуре данных (назначении столбцов)
3. Преобразуйте данные в столбцах в нужный формат
4. Просмотрите общую структуру данных с помощью функции `glimpse()`

Анализ

4. Сколько участников информационного обмена в сети Доброй Организации?
5. Какое соотношение участников обмена внутри сети и участников обращений к внешним ресурсам?
6. Найдите топ-10 участников сети, проявляющих наибольшую сетевую активность.
7. Найдите топ-10 доменов, к которым обращаются пользователи сети и соответственное количество обращений
8. Опередилите базовые статистические характеристики (функция `summary()`) интервала времени между последовательными обращениями к топ-10 доменам.
9. Часто вредоносное программное обеспечение использует DNS канал в качестве канала управления, периодически отправляя запросы на подконтрольный злоумышленникам DNS сервер. По периодическим запросам на один и тот же домен можно выявить скрытый DNS канал. Есть ли такие IP адреса в исследуемом датасете?

> **Tip**
>
> Дополнительные материалы можно найти в Telegram https://t.me/datadrivencybersec

Обогащение данных

10. Определите местоположение (страну, город) и организацию-провайдера для топ-10 доменов. Для этого можно использовать сторонние сервисы, например http://ip-api.com (API-эндпоинт – http://ip-api.com/json).

> **Отчет**
>
> Для оформления отчета используйте следующие материалы:
>
> 1. https://i2z1.ddslab.ru/posts/lab_recommendations/
> 2. https://i2z1.quarto.pub/checklab/criteria.html
> 3. https://github.com/i2z1/Report_template

Сайт курса

https://i2z1.ddslab.ru/IAMCTH