infotecs

Тестовое задание для стажера по направлению «Разработчик Machine Learning для информационной безопасности» в Центр научных исследований и перспективных разработок ИнфоТеКС

Цель:
Разработать программный макет, предназначенный для обнаружения аномальной сетевой активности на хосте с использованием методов машинного обучения.

Программный макет – программа на языке Python (предпочтительно) или C++, которая запускается на подключенном к сети конечном устройстве (например, персональном компьютере) и выполняет в автоматическом режиме следующие **задачи:**

Обучение

1. Сбор трафика на сетевом интерфейсе в виде пакетов или Netflow-статистики (на усмотрение исполнителя) и сохранение в удобном формате. Должна быть возможность задать период сбора обучающей выборки (ОВ). По умолчанию он может быть равен 24 часа. ОВ накапливается и хранится на жестком диске компьютера.
2. Оцифровка собранной ОВ и преобразование к виду двумерной матрицы, каждая строка которой соответствует элементу анализа на предмет аномальности (пакет, поток или сессия – на усмотрение исполнителя).
3. Обучение нейронной сети (НС) с использованием ОВ. Тип НС – на усмотрение исполнителя.
4. Выбор порога аномальности на основании ошибок восстановления, полученных для ОВ.

Обнаружение аномалий

5. Непрерывный сбор на сетевом интерфейсе, оцифровка и проверка на предмет аномальности тестовых данных с помощью обученной НС. *Вариант выполнения подзадачи* – вначале сбор тестовой выборки (ТВ) и сохранение на жесткий диск ПК в удобном формате, и затем оцифровка и проверка.
6. Формирование отчетов об обнаруженных аномалиях (на консоли и в файле).

Должна быть возможность при запуске программы подгрузить и использовать обученную ранее и сохраненную на диск НС.

Требования к присылаемым решениям:
Присылаемое решение должно содержать:

- созданный программный макет и исходный код к нему;
- описание программного макета, инструкция по установке и использованию;
- результаты тестирования, примеры обнаружения аномалий.

Максимальное время на выполнение задания: 3 недели.