«Сейфуллин оқулары – 18: « Жастар және ғылым – болашаққа көзқарас» халықаралық ғылыми -практикалық конференция материалдары = Материалы международной научно-практической конференции «Сейфуллинские чтения – 18: « Молодежь и наука – взгляд в будущее» - 2022.- Т.I, Ч.IV. - С. 27-30

ПЛАТФОРМА ОБНАРУЖЕНИЯ АНОМАЛЬНОГО ТРАФИКА ДЛЯ ИНТЕРНЕТА ВЕЩЕЙ (IOT)

Нургалиев Т.М., магистрант 2-го курса
Казахский агротехнический университет им С. Сейфуллина, г. Нур-Султан.

Введение

Невероятные достижения в повседневном использовании сетевых сервисов и электронных приложений привели к огромному прогрессу в коммуникационных сетях и появлению концепции интернета вещей (iot). Интернет вещей — это многообещающая парадигма, состоящая из распределенных сенсорных узлов, облачных серверов и программного обеспечения.

Объекты или вещи для восприятия и обработки в реальном времени, и они имеют возможность ощущать свою среду, связываться друг с другом и обмениваться данными через Интернет. Такая сеть приносит большие экономические выгоды, поскольку повышает эффективность генерации и использования данных. В последние годы IoT внес значительный вклад во многие области, такие как умные дома, здравоохранение, сельское хозяйство, транспорт и т. д.[1]–[3].По прогнозам, к 2020 году количество устройств IoT достигнет 20 миллиардов по сравнению с населением мира, составляющим более 7 миллиардов человек.

Растущее количество устройств IoT предоставит злоумышленникам много возможностей для их компрометации с помощью вредоносных электронных писем, атак по сговору и атак типа «отказ в обслуживании», а также многих других типов атак.[4].Многие устройства IoT подвергаются различным сетевым атакам, использующим различные уязвимости, такие как шифрование и безопасность паролей. Злоумышленники могут контролировать смарт-объекты с помощью кибератак и влиять на целые сети IoT, распространяя вредоносные приложения (например, вредоносные программы). Таким образом, кибербезопасность в настоящее время становится ключевым препятствием для более широкого внедрения услуг Интернета вещей. Обычно существует две основные категории методов обнаружения аномального трафика в IoT:

1) сигнатура и обнаружение аномалий [4]–[5].Каждая категория имеет свои преимущества и ограничения. Методы обнаружения на основе сигнатур обладают очень многообещающими характеристиками обнаружения известных аномалий. Однако им бросают вызов новые виды атак, так как особенности этих атак трудно идентифицировать. Обнаружение на основе

аномалий строит модель, содержащую образцы нормального поведения, и учитывает отклонения от модели для выявления подозрительного поведения или атак. Эти подходы позволяют выявлять новые типы атак, но могут приводить к увеличению ложных срабатываний, в основном из-за неравномерного распределения обучающих выборок.

2) Хотя по вопросам безопасности IoT было проделано несколько отличных работ, большинство существующих работ сосредоточено на вопросах безопасности на сетевых уровнях, таких как аутентификация, шифрование, управление ключами, согласованность данных и т. д. Следовательно, безопасность на уровне приложений в существующих работах в значительной степени игнорируется. В частности, протокол передачи гипертекста (HTTP) как универсальный протокол широко используется в приложениях IoT. Однако из-за открытости и разнообразия протоколов IoT-приложений уязвимости системы безопасности легко задействовать на этапах проектирования и развертывания. Злоумышленники могут использовать HTTP, что серьезно угрожает конфиденциальности личной информации и собственности пользователей в IoT.

Например, вредоносные коды запросов могут быть встроены в URL-адреса для запуска HTTP-атак и получения доступа к информации о разрешениях [1]. В этой статье, с последними достижениями в области периферийного интеллекта (EI)[2], [3], мы изучаем обнаружение аномалий HTTP для IoT. Вклад этой статьи резюмируется как следует.

1) Новая структура обнаружения аномалий HTTP предназначена для последовательного использования методов кластеризации и классификации, которые могут быстро и точно обнаруживать аномалии в HTTP-трафике для IoT.

2) В отличие от существующих работ, основанных на централизованном сервере для обнаружения аномалий, с последними достижениями в области EI и с учетом эффективности использования ресурсов, предлагаемая структура распределяет весь процесс обнаружения по разным узлам. Эта структура может эффективно уменьшить перегрузку сети и вычислительную нагрузку на централизованные серверы, а также раскрыть потенциал EI в IoT.

3) Предлагается новый метод обработки данных для разделения полей обнаружения данных HTTP, который может устранить избыточные данные и извлечь функции из полей заголовка HTTP. Кроме того, представлен одноклассовый HYBIRD-классификатор для повышения производительности системы обнаружения аномалий.

4) Результаты моделирования представлены, чтобы показать, что предлагаемая структура может значительно повысить скорость и точность обнаружения аномалий HTTP, особенно неизвестных аномалий.

Предлагаемая структура обнаружения

Структура полностью сочетает в себе преимущества EI для разделения всего процесса обнаружения на разные узлы. Не увеличивая время обнаружения, он эффективно снижает вычислительную нагрузку центра обработки данных. Кроме того, обнаружение кластеров в платформе

отфильтровывает часть обычного трафика, что эффективно снижает нагрузку на полосу пропускания.[6]

A. Обработка данных

Перед обнаружением необходимо обработать данные HTTP, как показано на рис. 3. Процесс в основном делится на три этапа: 1) обнаружение поля; 2) нормализация; 3) векторизация. Детали каждого шага описаны ниже.

*1) Обнаружение поля:*Трафик HTTP содержит информацию из множества разных полей, и каждое поле может содержать фактические атаки вторжения. Чтобы лучше обнаруживать атаки и избегать помех, вызванных бесполезной информацией в других полях, мы предлагаем полевое обнаружение. Поскольку каждое поле заголовка обнаруживается отдельно, заголовок необходимо сегментировать. Этот метод хорош для устранения избыточных данных и повышения точности обнаружения, а также может определить местоположение аномалии в точном поле.[5] Показывает пример различных полей в данных HTTP после сегментации.

Поскольку объем данных HTTP очень велик, чтобы обеспечить быстрый доступ к данным в процессе обнаружения, для представления данных используется структура словаря. Создается словарь с именем Dictm. Он имеет ряд ключей, и каждый ключ представляет собой поле заголовка HTTP. Каждый сегмент заголовка хранится в Dictm.

2) Нормализация: Чтобы упростить данные обнаружения и эффективно защитить конфиденциальность личной информации, мы нормализовали данные HTTP. Поскольку каждое поле имеет некоторые структурные особенности и структурную информацию, их можно упростить с помощью определенных символов. Здесь «C» используется для представления связанных с SQL ключевых слов в данных (например, «выбрать», «подсчитать», «откуда» и «где»). Все остальные символы, кроме символов, заменяются на «A». После нормализации поля заголовка HTTP-запроса преобразуются в sig. Это не только упрощает данные, но и сохраняет их структурную информацию.

3) Векторизация: На этапе векторизации sig преобразуется в вектор для облегчения последующих этапов кластеризации и классификации. В этом процессе анализ n-грамм используется для извлечения значимых признаков из сигнала. Его также можно рассматривать как подстроку длины n. Например, строка «ababc» содержит четыре подстроки с «ab», «ba», «ab» и «bc», которые имеют три уникальных 2-грамма «ab», «ba» и «bc». 2-граммовое «ab» встречается дважды с частотой 2/4. 2-граммовые «ba» и «bc» появляются только один раз с частотой 1/4. Список текстовых токенов может быть представлен вектором, состоящим из частот n-грамм. Вектор признаков, описывающий эту строку, будет Xababc

Выводы

Непрерывная эволюция кибератак и массовое использование IoT-приложений поставили перед обнаружением аномалий в IoT новые проблемы. Для решения этих проблем была предложена новая структура

обнаружения аномалий, основанная на последних достижениях в области EI, путем последовательной кластеризации и классификации HTTP-трафика, которая может эффективно и действенно обнаруживать неизвестные вторжения в сеть. Кроме того, мы представили метод обработки данных с обнаружением полей для устранения избыточных данных, который делит заголовок данных HTTP-трафика на несколько полей для обнаружения. Этот процесс может ускорить вычислений.

Список использованных источников

1. П. Дюссель, К. Гейл, У. Флегель, С. Дитрих и М. Мейер, «Обнаружение атак нулевого дня с использованием контекстно-зависимого обнаружения аномалий на уровне приложений», Int. Дж. Инф. Безопасность, том. 16, нет. 5, стр. 475–490, 2017.

2. М.Б. Сейяр, Ф.О. Чатак и Э. Гюль, «Обнаружение направленных на атаку сканирований из журналов доступа к HTTP-серверу apache», Appl. вычисл. Информ., вып. 14, нет. 1, стр. 28–36, 2018.

3. Л. Ни, Ю. Ли и К. Конг, «Пространственно-временная оценка сетевого трафика и обнаружение аномалий на основе сверточной нейронной сети в автомобильных специализированных сетях», IEEE Access, vol. 6, стр. 40168–40176, 2018.

4. А. Ювонен, Т. Сипола и Т. Хямяляйнен, «Онлайн-обнаружение аномалий с использованием методов уменьшения размерности для анализа журнала HTTP», Comput. Сеть., том. 91, с. 46–56, ноябрь 2015 г.

5. С. Ван, К. Ян, З. Чен, Б. Ян, К. Чжао и М. Конти, «Обнаружение вредоносных программ для Android, использующих текстовую семантику сетевых потоков», IEEE Trans. Инф. Криминалистическая безопасность, том. 13, нет. 5, стр. 1096–1109, май 2018 г.

6. Yan H. et al. Centralized duplicate removal video storage system with privacy preservation in IoT //Sensors. – 2018. – Т. 18. – №. 6. – С. 1814