СХЕМА АНАЛИЗА СЕТЕВОГО ТРАФИКА

Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь
Романенко О.А

Мухуров Н.И. – д.т.н.,профессор

Задача анализа сетевого трафика в настоящее время становится все более актуальной. Этому способствует не только развитие и внедрение новых сетевых технологий (и ,как следствие, увеличение объёма данных, передаваемых по сети), но и появление большого количества новых сетевых протоколов прикладного уровня. Основными практическими задачами анализа сетевого трафика являются: выявление проблем в работе сети; тестирование (отладка) сетевых протоколов; восстановление потоков данных («прослушивание»); предотвращение сетевых атак; сбор статистики.

Общая схема анализа сетевого трафика состоит из определённой последовательности шагов. Каждый из приведенных ниже шагов приводит к повышению уровня представления объекта анализа.

1.Захват пакетов, проходящих через контролируемое сетевое соединение. Результат шага – получение объекта анализа в виде сетевых пакетов. В зависимости от необходимой точности и скорости последующего анализа, а также доступных вычислительных мощностей могут использоваться различные подходы.

- Слайсинг. При использовании этого подхода анализу подвергаются не всё содержимое пакетов, а только некоторый префикс (n первых байт). В ряде исследований показано, что этот подход хорошо работает для последующей классификации трафика по протоколам.

- Сэмплинг. При использовании этого подхода перехватываются не все пакеты, а только их часть, которая может выбираться по различным критериям, в зависимости от потребностей. В процессе развития технологии было предложено большое число стратегий отбора. Например, для задач мониторинга типов трафика подходит вариант с выбором каждого n-го пакета (uniform sampling), где n может выбираться в зависимости от соотношения ширины канала и пропускной способности системы анализа.

- Для задач, в которых требуется максимально точный анализ трафика, например для систем обеспечения сетевой безопасности, требуется перехватывать все данные всего поступающего трафика без потерь – для обозначения этого подхода используется термин lossless capture или deep packet capture (DPC).

2. Агрегирование пакетов в потоки по некоторым адресным признакам (flow generaion), получение нового объекта для анализа – сетевого потока. Если при этом данные пакетов в дальнейшем анализе не учитываются, то такой вид анализа называется «анализ потоков» - flow based analysis (в отличие от packet-based анализа, при котором анализируются данные пакетов). Flow-based анализ широко используется в силу значительно меньших требований к мощности вычислителя и пропускной способности, за счёт значительного снижения объёма данных для обработки. Такой вид анализа может выполняться и локально, и удалённо от точки сбора данных. Для передачи собранных данных от точки сбора до точки анализа используется большое число протоколов, часть из которых стандартизирована в виде IPFIX. IP Flow Information Export – стандарт, современный вариант модели сетевых потоков netflow, который обеспечивает компактное и универсальное представление информации о сетевом трафике. Другая часть разработана отдельными производителями – Cisco NetFlow, Juniper Jflow. В рамках подхода записи, описывающие поток, могут содержать различный набор данных. Наиболее общим набором таких данных является: IP адреса источника и адресата, протокол транспортного уровня, в случае протоколов TCP/UDP – номера портов источника/адресата, набор счётчиков: количество переданных пакетов и байт, время создания и завершения потока.

Данный метод действительно значительно снижает требования к анализатору, тем не менее, он не является достаточно гибким, так как в отличие от слайсинга и сэмплинга не позволяет варьировать количество поступающих данных (зависит от входных данных).

3. Выполнение классификации по протоколу прикладного уровня или конкретному сетевому приложению. Результатом данного шага является получение нового объекта для анализа – сетевого потока конкретного протокола или приложения (в этом случае связанных потоков может быть несколько, например, в случае VoIP приложения это потоки SIP и RTP). После выполнения данной операции возможна следующая дополнительная обработка полученного объекта, конкретный вид которой зависит от решаемой прикладной задачи:

- разбор полей протокола (protocol parsing),
- сборка сессии протокола для протоколов с установлением соединения,
- извлечение данных приложения (content extraction) – страниц сайтов (HTML), файлов различных типов (исполняемые, изображения, текстовые документы, и т.д.), электронных писем, аудио-видео потоков
- разбор данных приложения (application content parsing).

Анализ сетевого трафика актуален из-за быстрого совершенствования сетевой отрасли. Он также жизненно важен для эффективного управления сетью, по результатам которого можно судить о качественных и количественных характеристиках работоспособности сети или её отдельных компонентов.

Список использованных источников:

1. Гетьман А. И., Евстропов Е.Ф., Маркин Ю. В. Анализ сетевого трафика в режиме реального времени: обзор прикладных задач, подходов и решений // Препринт ИСП РАН №28, 2015. – 16с.

2. Маркин Ю. В., Санаров А.С. Обзор современных инструментов анализа сетевого трафика // Препринт ИСП РАН №27, 2014.