СОВРЕМЕННЫЕ МЕТОДЫ АНАЛИЗА СЕТЕВОГО ТРАФИКА

Климец А.А.

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

Перцев Д.Ю. – к.т.н., доцент

Представлен обзор современных подходов к глубокому анализу сетевых пакетов, лежащих в основе современных DPI-систем, используемых интернет-провайдерами и спецслужбами разных стран, а также причины их использования.

DPI – это технология проверки, фильтрации и накопления статистических данных о сетевом трафике по его содержимому или каким-то специфичным характеристикам протоколов. На данный момент на рынке представлено множество различных технологий DPI, они могут быть как standalone решениями, так и встроенными в маршрутизаторы. Среди наиболее крупных компаний, разрабатывающих DPI-системы, выделяются: Allot Communications, Cisco, Sandvine, Huawei. В данный их системы широко используют по всему миру.

В общем случае DPI способен фильтровать трафик на транспортном и сетевом уровнях модели OSI, однако наибольшая точность идентификации или классификации трафика происходит на сеансовом, представительском и прикладном уровнях, анализируя содержимое отдельных пакетов. Базовой проверки пакетов на данный момент недостаточно, так как количество генерируемого пользователями трафика, а также используемых протоколов неизменно растёт огромными темпами. Всё большее развитие потоковых приложений, обфусцированных протоколов, к примеру, QUIC версии 50, а также широкое использование VPN, создаёт значительные препятствия для работы DPI систем, не говоря уж об общей тенденции на улучшение безопасности пользовательских данных в сети Интернет. Всё это стало причиной появления новых техник распознавания трафика.

Если раньше во многом DPI системы сводились к обычной проверке содержимого 1-2 пакетов, чаще всего HTTP или HTTPS трафика, проверкой поля Host или SNI соответственно, на данный момент DPI системы широко используют альтернативные способы идентификации трафика по каким-либо косвенным признакам, присущим определённым сетевым программам и протоколам. Широко используется статистический и поведенческий анализ, в частности анализ встречаемости определённых символов, длины пакетов и др.

Для подобного анализа больших объёмов сетевого траффика нередко используются методы Data Mining, позволяющие обнаружить неявные зависимости между конкретными сетевыми соединениями и работающими в этот момент приложениями или какими-либо действиями пользователя при использовании веб-приложений. Применение подобных методов делает идентификацию трафика чрезвычайно точным, ведь современные приложения широко используют различные микросервисы и широко интегрируются с различными сторонними сервисами, что создаёт большое количество сетевых сессий, подписать которые бывает достаточно трудно.

Подходы к анализу трафика посредством DPI ограничиваются только тем, сколько различных параметров может проанализировать та или иная система, а также сложностью анализируемого протокола. Однако основная проблема DPI в том, что при определении принадлежности трафика к тому или иному приложению необходимо, чтобы устройство проверило оба направления сетевых сессий, т.е. необходимо проверять асимметричный трафик, который широко встречается у крупных операторов.

Существует несколько известных подходов к работе с таким траффиком: Cisco проверяет только одну часть сессии, Sandvine инкапсулирует асимметричный трафик в broadcast-фреймы и в таком виде прогоняет их через свои DPI устройства [1].

Существует много причин для использования DPI. Одна из причин – это экономия средств, так как внедрение DPI гораздо выгоднее, чем расширять аплинки, и позволяет осуществлять контроль утилизации каналов. Кроме того, DPI активно используется для реализации QoS, к примеру, давая приоритет VoIP-сервисам или стриминговым платформам, что актуально во время пандемии COVID-19, а также система может значительно снижать пропускную способность некоторых сервисов, тем самым осуществляя контроль загрузки сетей и для борьбы с нежелательным трафиком, выполнять полную блокировку тех или иных сайтов или приложений. И, наконец, DPI системы используются в маркетинговых целях, для сбора статистики по наиболее популярным у пользователей приложениям или тому, какие устройства используют пользователи. На этом фоне на данный момент стало чрезвычайно популярным у мобильных операторов делать тарифы, предоставляющие безлимитный доступ к мессенджерам или стриминговым сервисам.

Список использованных источников:

1. *Краткий обзор технологии DPI – Deep packet Inspection [Электронный Ресурс]. – Режим доступа:* https://habr.com/ru/post/111054/*. – Дата доступа: 21.03.2021*