ВЫДЕЛЕНИЕ СТРУКТУРНЫХ ЭЛЕМЕНТОВ СЕТЕВОГО ТРАФИКА РЕАЛЬНЫХ СЕТЕЙ В ЗАДАЧЕ ТЕСТИРОВАНИЯ КОММУНИКАЦИОННОГО ОБОРУДОВАНИЯ

А. В. Агафонов
(Екатеринбург, УрФУ)

В настоящее время для существующих и разрабатываемых компьютерных сетей характерно повсеместное применение такого коммуникационного оборудования (КО), как маршрутизаторы и коммутаторы. Характеристики сетевого трафика (СТ) обрабатываемого КО могут значительно различаться в конкретных сетях, что оказывает влияние на ключевые параметры, связанные с производительностью КО: пропускную способность, задержку пересылки, относительную долю потерь пакетов.

Необходимость обоснованного выбора КО в условиях отсутствия описания алгоритмов и принципов их работы приводит к актуализации задачи тестирования оборудования для экспериментального подтверждения заявленных разработчиком значений указанных параметров производительности в условиях конкретных сетей, в которых предполагается эксплуатация устройств.

Тестирование КО может быть осуществлено с использованием:

- реального СТ;
- синтезируемого СТ.

При создании системы тестирования КО синтезированный СТ оказывается предпочтительным в сравнении с реальным, так как обладает следующими важными преимуществами:

- отсутствием необходимости хранения больших массивов трафика;
- гарантированным отсутствием защищаемой законом информации и сетевых аномалий, которые могли иметься в трафике реальных сетей.

При использовании синтезированного трафика процедуры его подготовки разделяются на два этапа:

1. Измерения значимых характеристик трафика реальной сети;

2. Синтез тестового трафика с использованием измеренных характеристик.

При этом защита от утечки охраняемой законом информации обеспечивается заменой множества реальных сетевых адресов на сгенерированные случайным образом (с учетом топологии исходной сети), а также измерением лишь статистических характеристик трафика реальных сетей и игнорированием информационного наполнения сетевых пакетов, а реалистичность получаемого тестового трафика — обоснованным выбором данных характеристик на этапе измерения.

Реалистичным в контексте данной работы будем называть синтезированный трафик, отражающий следующие свойства реальной сети:

- количество взаимодействующих узлов и топологию;
- статистическое распределение логических соединений между взаимодействующими сетевыми узлами;
- статистические характеристики трафика, связанные с размером и временным распределением сетевых пакетов внутри каждого из логических соединений.

Используемая в данной работе модель рассматривает СТ как совокупность составляющих его *потоков*, где под термином *поток* понимается процесс обмена данными между двумя приложениями конечных узлов сети с использованием определенного протокола транспортного уровня в течение определенного интервала времени.

Модель выделяет следующие ключевые параметры потока:

- сетевые адреса:
  - ♦ узла, принимающего соединение – d;
  - ♦ узла-инициатора соединения – s;
- идентификатор протокола транспортного уровня (TCP, UDP) или подуровня управления сетью (ICMP) – t;
- номера портов (для протоколов транспортного уровня) или типы получаемых сообщений (для протоколов подуровня управления сетью):
  - ♦ узла, принимающего соединение – p_d;
  - ♦ узла-инициатора соединения – p_s.

В массиве СТ могут существовать несколько потоков, имеющих одинаковые значения указанных выше параметров и различающихся лишь моментами времени начала и окончания передачи данных. Все множество таких потоков может быть описано функцией распределения вероятности длительности отдельного потока – $F_{tf}(t_f)$.

Сетевые пакеты, передаваемые в рамках одного потока, в подавляющем большинстве случаев содержат данные, созданные определенными приложениями конечных узлов потока, и поэтому могут иметь схожие характеристики, такие как размер пакета или промежуток времени между началами передач текущего и следующего за ним пакетов, которые также могут быть описаны с использованием функций распределения вероятности:

- размера пакетов – $F_{lp}(l_p)$;
- промежутка времени между началами передач двух последовательных пакетов – $F_{tp}(t_p)$.

Таким образом, отдельный поток C может быть описан выражением (1):

$$C = \langle d, s, t, p_d, p_s, F_{tf}(t_f), F_{lp}(l_p), F_{tp}(t_p) \rangle. \quad (1)$$

Так как массив сетевого трафика является совокупностью составляющих его потоков, каждый из которых встречается в нем с определенной вероятностью, то данный массив M может быть описан в рамках рассматриваемой модели выражением (2), где $F_C(i)$ – функция распределения вероятности появления i-го потока в массиве тестового трафика:

$$M = \left\langle (C_i)_{i=1}^{n}, F_C(i) \right\rangle. \quad (2)$$

Таким образом, для измерения значимых характеристик трафика реальной сети в соответствии с приведенной моделью необходимо выделить внутри него отдельные потоки.

Особенностью процедуры выделения потоков является различие механизмов определения начала и окончания потоков для различных протоколов. Поэтому первым этапом обработки записанного в реальной сети массива СТ является его разделение на состав-

ляющие в зависимости от использования конкретного протокола транспортного уровня или подуровня управления сетью. Дополнительно пакеты сортируются в порядке их отправки при наличии необходимой для этого информации. Например, для TCP процедура реализуется путем анализа поля порядкового номера и номера подтверждения заголовков.

Выделение в записанном в трафике протокола TCP отдельных потоков реализуется с использованием программного автомата, в котором переходы между состояниями осуществляются по сигналам вида <узел1, узел2, флаги>, генерируемым на основе заголовка текущего обрабатываемого пакета.

Автомат использует в процессе работы следующие переменные, обозначающие сетевые адреса:

- o1 – узла, инициировавшего установление соединения;
- o2 – узла, с которым o1 устанавливает соединение;
- r1 – узла, инициировавшего разрыв соединения;
- r2 – узла, соединение с которым разрывается;
- f1 – узла, инициировавшего завершение соединения;
- f2 – узла, с которым f1 завершает соединение (рис. 1).

Символами «*, *» в сигналах обозначены пары адресов, где их порядок не важен.

Поле «флаги» сигнала отображает значения следующих флагов заголовка TCP: ACK, SYN, FIN, RST.

Автомат имеет следующие состояния:

- CLOSED – соединение не установлено;
- OPEN_1, OPEN_2 – промежуточные этапы установления соединения;
- ESTABLISHED – соединение установлено;
- FIN_1, FIN_WAIT, FIN_2 – промежуточные этапы завершения соединения.

Начальным состоянием автомата является CLOSED.

При переходе в каждое из состояний или при их сохранении на каждом шаге автомат осуществляет чтение очередного пакета из массива трафика и в зависимости от его параметров выполняет одну из следующих команд:

• nw() – удалить ранее использовавшиеся значения переменных, создать новую запись о потоке, сделать ее текущей и сохранить в нее пакет;

• w() – записать пакет в текущий поток.

Для корректной обработки соединений, не завершенных штатно, в каждом из состояний контролируется временная задержка между предыдущим обработанным и текущим пакетами. Если в текущем пакете установлен флаг SYN и рассматриваемая задержка больше предельного значения (для протокола TCP на практике применяется значение 7200 с [1]), то происходит переход в состояние OPEN_1 с выполнением команды nw().

Если текущий пакет не удовлетворяет правилу, указанному выше, а также ни одному из условий переходов, приведенных на рис. 1, то осуществляется переход в состояние CLOSED с выполнением команды w().

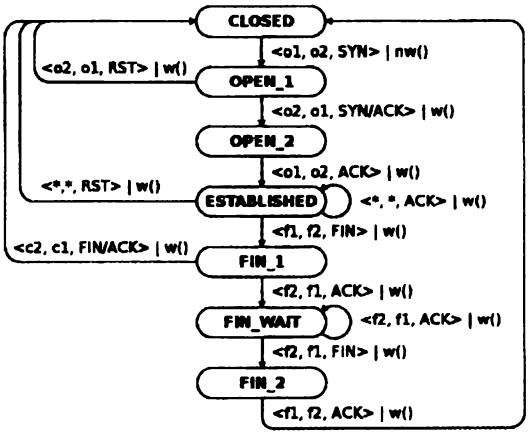

Рис. 1. Граф для выделения потоков TCP

Протокол UDP не предусматривает установления и закрытия соединения, в заголовках данного протокола отсутствует информация об очередности поступления пакетов и не содержится поле флагов, которое позволило бы указывать явно моменты начала и окончания потоков данных. С другой стороны, использование дан-

ного протокола обычно подразумевает непрерывную передачу данных, поэтому распространенным подходом при отслеживании потоков UDP является введение предельного значения задержки между приемом двух последовательных пакетов внутри соединения (на практике принимается равным 180 с).

Выделение потоков UDP показано на рис. 2:

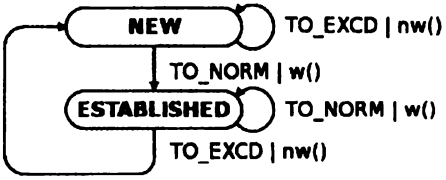

Рис. 2. Граф для выделения потоков UDP

На рис. 2 использованы следующие обозначения сигналов перехода:

- TO_NORM – задержка между предыдущим обработанным пакетом и текущим превышена;
- TO_EXCD – задержка не превышена.

Автомат имеет следующие состояния:

- NEW – обнаружен новый поток;
- ESTABLISHED – передача данных в рамках текущего потока.

Начальным состоянием автомата является NEW. При чтении первого пакета массива трафика задержка считается превышенной (TO_EXCD).

Протокол ICMP является одним из важнейших элементов стека TCP/IP, обеспечивающим правильную работу сети, поэтому трафик данного протокола также должен быть проанализирован. Как и UDP, ICMP не предусматривает создания устойчивых логических соединений, однако некоторые из сообщений предполагают взаимодействие узлов в режиме «запрос – ответ». Данные сообщения в рамках используемой модели считаются относящимися к одному потоку, а определение взаимосвязи полученного ответа с отправленным ранее запросом происходит путем анализа полей Identifier и Sequence Number [2].

После выполнения рассмотренных выше операций исходный массив сетевого трафика оказывается разбит на множество отдельных структурных элементов – потоков, после чего может быть произведено вычисление необходимых для используемой модели статистических характеристик как отдельных потоков, так и массива СТ в целом.

Одним из наиболее важных достоинств предлагаемого в данной работе механизма выделения структурных элементов СТ является его расширяемость: благодаря отсутствию привязки к конкретной реализации сетевой адресации, он легко может быть модернизирован для работы с трафиком IPv6, а разработка дополнительных подпрограмм для выделения потоков STCP и RTP позволит корректно обрабатывать трафик и этих протоколов транспортного уровня.

Библиографические ссылки

1. http://frozentux.net/documents/iptables-tutorial/
2. http://www.ietf.org/rfc/rfc792.txt

БРАНДМАУЭР ВЕБ-ПРИЛОЖЕНИЙ – ОБЗОР, НЕДОСТАТКИ

А. В. Адамов, А. В. Бабич
(Тюмень, ТюмГУ)

Введение

На сегодняшний день одной из самых быстроразвивающихся сфер деятельности человека является Интернет. Неудивительно, что наряду с его развитием теми же темпами развиваются и угрозы безопасности сервисов, работающих в среде Интернет. По данным статистики WASC (Web Application Security Consortium) [1] более 13 % сайтов могут быть скомпромитированы полностью автоматически, 80–96 % из которых имеют высокую степень уязвимостей, 86 % – среднюю степень уязвимостей, 37 % – низкую.

В данной работе хотелось бы поговорить о части Интернета, являющейся его непосредственным лицом и касающейся практи-