УДК 519.854

ЛИСТРОВОЙ С.В., д.т.н., профессор,
МОЦНЫЙ С.В., аспирант (УкрГАЖТ)

Анализ сетевого трафика стека протоколов TCP/IP на основе динамической модели

В данной статье рассматривается динамическая модель анализа сетевого трафика стека протоколов TCP/IP, разработка которой является крайне важной в условиях постоянно обновляющейся сетевой инфраструктуры. Проведен анализ существующих подходов, с помощью которых возможно обеспечить приемлемый уровень сетевой безопасности. Рассмотрена практическая возможность реализации алгоритмов нахождения минимального вершинного покрытия при построении динамической модели.
Ключевые слова: *динамическая модель, компьютерные сети, VLAN, сетевой трафик, вершинное покрытие, стек протоколов TCP/IP, статистический анализ.*

Постановка задачи на исследование

Активное развитие сетевых технологий и расширение объема информационных услуг обусловливает постоянный прирост новых пользователей, который носит явно выраженный динамический характер. При этом также наблюдается увеличение объемов сетевого трафика. Согласно проведенным исследованиям [1], приблизительная динамика роста трафика, передаваемого посредством Всемирной сети, составляет 70-150% в год (за последние несколько лет), т.е. в среднем каждый год количество информации, предназначенной для передачи, удваивается.

Однако, наравне с увеличением трафика, наблюдается стремительный рост различных информационных угроз, возникающих при сетевом взаимодействии. По данным антивирусной лаборатории И. Данилова, только за один месяц в вирусную базу добавляется более 7 тыс. записей. По мере того, как трафик заполняет внутрикорпоративные сети, становится очевидным определяющее влияние динамического анализа возможных угроз безопасности на экономическую эффективность телекоммуникационной отрасли.

Последние работы в области исследования информационной безопасности в компьютерных сетях [2] наглядно демонстрируют тот факт, что независимо от используемых протоколов, сетевой топологии, инфраструктуры сетевого взаимодействия, в основе проектируемых систем лежат одни и те же принципы передачи данных, а, следовательно, имеются очень схожие проблемы безопасности. Анализ сетевого трафика при этом является основным элементом, входящим в различные методики исследований.

Среди существующих на сегодняшний день подходов анализа трафика (стека протоколов TCP/IP) можно выделить следующие направления: на основе подготовленной заранее статистической модели, подходы с использованием методики допустимого порога и отклонения характеристик и др. Все они имеют свои плюсы и минусы. Однако, при использовании того или иного подхода часто приходится выбирать между точностью и эффективностью анализа и производительностью системы. В связи с этим необходима разработка новой усовершенствованной методологии анализа сетевого трафика, которая бы учитывала важность динамической составляющей концепции современных компьютерных систем.

Анализ последних исследований и публикаций

Рассматривая исследования в области анализа сетевого трафика несложно заметить, что сетевые технологии в значительной степени опережают в темпе своего развития как аналитическое, так и теоретическое понимание сетевых взаимодействий.

Несмотря на то, что узкоспециализированные телекоммуникационные задачи прошлых лет хорошо изучены и математически формализованы (к примеру, основные положения теории массового обслуживания [3]), подобные методы не соответствуют необходимым требованиям современной динамической структуры существующих систем. Традиционные принципы не позволяют достаточно точно предсказывать такие характеристики, как сетевые задержки, длины очередей и т.д. Экспериментальный анализ сетевых процессов набирает сегодня все большую популярность по сравнению с классическими математическими моделями прошлых лет.

Среди наиболее широко распространенных технологий анализа выделяются так называемые

RBID-системы, основанные на правилах (Rule-Based Intrusion Detection)[4]. Данные системы с целью выявления зловредного потока используют сравнение сигнатур с заранее подготовленной вирусной базой. После обнаружения атаки происходит анализ ее характеристик, а затем создается новое правило, которое в будущем обеспечит защиту от данного вида вторжения.

Другой подход к обнаружению информационных угроз имеют статистические системы (SBID) [5]. Статистический анализ относят к поведенческим методам определения неисправностей в компьютерных сетях. Данный подход основан на сопоставлении текущего состояния сети с определёнными заранее параметрами, описывающими правильное функционирование всей системы в целом.

Методы статистического анализа сетевого трафика используются в качестве инструментов прогнозирования загруженности каналов связи, диагностики искажений трафика, потерь информации и т.д., при этом они имеют различные интерпретации, основанные на различных характеристиках сетевого трафика. Главным преимуществом методов статистического анализа считается потенциальная возможность гарантировать безопасность не только от уже известных характеров атак, но и от предугаданных заранее.

Довольно крупную нишу занимает класс методов, основанных на маршрутизаторах. Важной составляющей данных методов является протокол простого сетевого мониторинга (SNMP)[6]. SNMP является частью протокола TCP/IP и позволяет осуществлять такие операции: сбор статистики по трафику, планирование роста сети, анализ производительности, обнаружение различных сетевых проблем и др. Для данного протокола существуют различные расширения, которые применяются в различных специфических отраслях. Среди них можно выделить удалённый мониторинг (RMON), который предоставляет возможность настраивать сигналы, диагностирующие сеть, основанные на определённом критерии, а также расширение Netflow (RFS 3954), использующееся в маршрутизаторах Cisco, которое предоставляет возможность собирать IP сетевой трафик и преобразовывать данные для экспорта.

Анализ сетевого трафика также представляется возможным обеспечить с помощью использования фильтров Блума, благодаря которым заметно повышается эффективность обнаружения информационных угроз.

Выделение нерешенных ранее частей общей проблемы

Примечательной особенностью практически всех существующих моделей является выраженный статический характер анализа, который является недостаточно эффективным для поддержки гарантировано высокого уровня безопасности в современных условиях постоянного динамического обновления сетевых инфраструктур. Также многие подходы значительно снижают скорость обмена данными, повышают стоимость и сложность обслуживания компьютерной сети.

Цель статьи. Проведенный анализ существующих подходов говорит о необходимости разработки динамической модели анализа сетевого трафика TCP/IP, которая бы учитывала характеристики и требования современных компьютерных систем.

Подходы к решению поставленной задачи

Как при проектировании сети, так и при ее динамической реконфигурации, возникает проблема соблюдения баланса между стоимостью применяемых систем анализа трафика и эффективностью их работы. Для решения этой проблемы предлагается использовать алгоритм нахождения минимального вершинного покрытия, с помощью которого находится оптимальный набор сетевых узлов с установленными системами обеспечения безопасности. Однако при этом необходимо обеспечить сохранность покрытия при динамическом функционировании компьютерной системы.

Рассмотрим произвольную компьютерную сеть, изображенную на рис. 1.

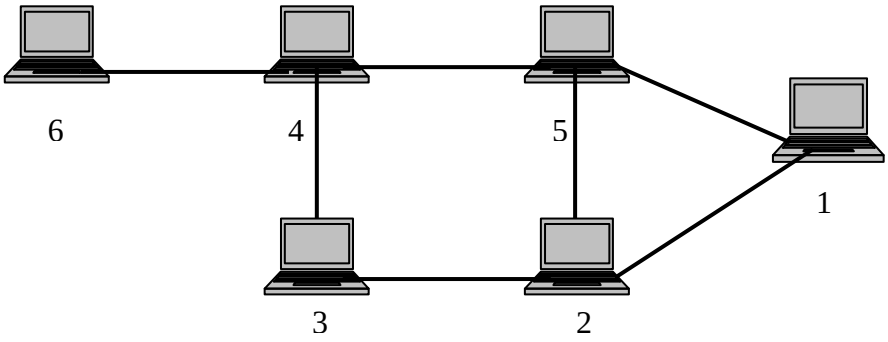

Рис. 1. Произвольная компьютерная сеть (минимальное вершинное покрытие подразумевается на узлах 2, 4, 5)

Предположим, что на данном этапе у нас уже имеется найденное при помощи оптимального алгоритма минимальное вершинное покрытие, которое в данном случае охватывает узлы {2, 4, 5}.

Необходимо определить, каким образом повлияет динамическая реконфигурация сети на структуру найденного покрытия, а также установить связанные с этим затраты.

Главными критериями, которые определяют эффективность переконфигурирования в данном случае, являются: уровень безопасности компьютерной сети, материальные затраты, временная сложность повторных расчетов минимального покрытия.

Для детального анализа и разработки подходов с соблюдением перечисленных критериев выделим основные варианты возможного динамического обновления сети:

- **Удаление одного и более узлов из сети**

В случае удаления из сети одного и более узлов возникает вопрос, сохранится ли приемлемый уровень безопасности? Другими словами, нужно ли будет снова решать задачу о нахождении наименьшего вершинного покрытия?

Исходя из самого определения вершинного покрытия, каждое ребро в нем должно быть инцидентно хотя бы одной вершине в данной сети. Учитывая специфику использования компьютерной сети, можно утверждать, что удаление любого узла также повлечет за собой удаление всех инцидентных с ним ребер. При этом никак не будет нарушена структура тех ребер, которые уже входили в покрытие каким либо своим концом.

Таким образом, становится очевидным, что после удаления одного или более узлов из сети нет необходимости заново пересчитывать алгоритм нахождения минимального покрытия, и, следовательно, уровень безопасности не снижается.

Пусть $A_{исх}$ – исходное множество узлов в компьютерной сети, $B_{уд}$ – множество удаленных узлов, V_{min} – минимальное вершинное покрытие. Если $B_{уд} \subseteq A_{исх}$, то в данном случае будет справедливо следующее выражение:

$$\forall (u, v), Vmin, \qquad (1)$$

где u – множество вершин, v – множество ребер.

- **Добавление узлов в сеть**

При добавлении узлов в компьютерную сеть предлагается использовать базовые идеи теории динамического программирования. В основе данной методики лежит способ решения сложной задачи путем разбиения ее на более простые составные части. Динамическое программирование позволяет снизить временную сложность благодаря запоминанию полученных решений. Другими словами, когда число простых задач экспоненциально велико, и при этом они с большой периодичностью повторяются в различных исходных задачах, пропадает необходимость их повторного перерасчета.

В случае динамического переконфигурирования компьютерной сети (рис. 1) представляется возможным разбить общую задачу на следующие части: нахождение исходного минимального вершинного покрытия, нахождение минимальных вершинных покрытий групп добавляемых узлов, объединение двух найденных множеств. Для первых двух пунктов достаточно использовать соответствующий алгоритм. Последний случай требует детального рассмотрения.

Для того чтобы в результате объединения множеств не нарушалось покрытие, необходимо следовать такому алгоритму:

- Если хотя бы один из объединяемых узлов входит в покрытие, значит, в структуре ничего изменять не нужно;

- Если оба узла не входят в покрытие, включаем первый из них.

Пусть $V_{min\ исх}$ – исходное минимальное вершинное покрытие, $V_{min\ найд}$ – найденное минимальное вершинное покрытие новой группы узлов. Тогда справедливо следующее выражение:

$$V_{min\ исх} \cup V_{min\ найд}, V_{min\ общ}, \qquad (2)$$

где $V_{min\ общ}$ – общее минимальное вершинное покрытие после добавления узлов в компьютерную сеть.

- **Случайное добавление связей среди существующих узлов**

При возникновении необходимости случайным образом добавить связи между существующими узлами компьютерной сети для сохранения всей структуры покрытия достаточно придерживаться принципов, исходящих из самого определения минимального вершинного покрытия. Т. е. необходимо последовательно добавлять связь таким образом, чтобы не образовалось ребра, инцидентного двум вершинам, не входящим в покрытие.

С целью построения эффективной динамической модели анализа сетевого трафика предлагается использование технологии под названием «Виртуальные сети» (VLAN) [7], а также применение программной реализации фильтров Блума HASH-AV [8]. Рассмотрим свойства и характеристики данных средств более подробно.

VLAN (Virtual Local Area Network) — группа устройств, взаимодействующие между собой на

канальном уровне сетевой модели, хотя физически они могут располагаться в разных частях географического пространства и могут быть подключены к разным сетевым коммутаторам. При этом устройства, находящиеся в разных VLAN'ах, невидимы друг для друга на канальном уровне. Они образуют отдельные широковещательные домены, даже когда подключены к одному коммутатору. Связь между этими устройствами возможна только на сетевом уровне (или более высоком).

Виртуальные сети возможно настроить на коммутаторах, маршрутизаторах и других сетевых устройствах. В современных компьютерных сетях это основной механизм для создания логической топологии, которая не зависит от физического расположения узлов. Также технология VLAN может использоваться для повышения сетевой безопасности (борьбы с ARP-spoofing'ом) и сокращения широковещательного трафика.

Для понимания принципа организации виртуальных сетей рассмотрим работу коммутатора. В каждом типичном коммутаторе (англ. Switch) хранится специальная таблица коммутации, при помощи которой производится передачи фреймов по сети. После первого включения эта таблица не содержит ни одной записи. Ее заполнение происходит автоматически по мере установления соединений между компьютерными узлами. Коммутатор формирует структуру записи таблицы, сопоставляя MAC-адрес узла-отправителя с номером порта, к которому было установлено соединение.

На коммутаторе, который изображен на рис. 2, показана организация двух виртуальных сетей: VLAN1 и VLAN2.

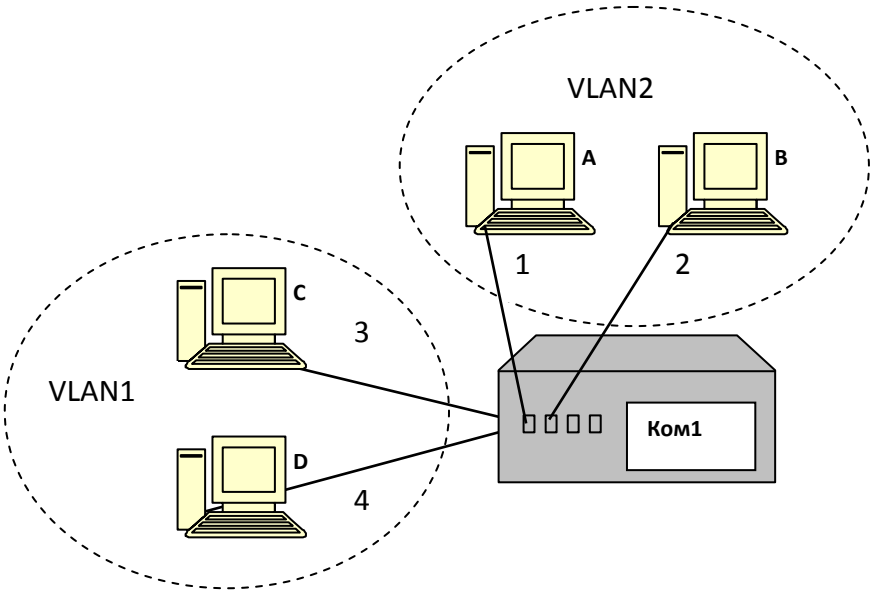

Рис. 2. Пример организации виртуальных подсетей на коммутаторе Ком 1:
A, B, C, D - MAC-адреса узлов; 1, 2, 3, 4 – соответствующие порты.

Порты коммутатора подразделятся на тегированные (англ. Tagged) и нетегированные (англ. Untagged). Под «нетегированными» поратми подразумеваются те порты, при помощи которых соединения устанавливаются стандартными способом. При помощи же «тегированных» портов появляется возможность передавать трафик различных виртуальных сетей между коммутаторами через один единственный порт. Пот тегом здесь понимается информация о принадлежности фрейма к той или иной виртуальной сети. Как правильно добавлять данный тег к фрейму описано в стандарте IEEE 802.1Q. Так как в нашем примере (рис. 2) используется всего один коммутатор, то порты на нем будут настроены как нетегированные.

Таким образом, благодаря технологии VLAN представляется возможным динамически добавлять узлы в компьютерную сеть, независимо от их географического расположения, что позволяет значительно оптимизировать общую сетевую конфигурацию. Перемещение средств анализа трафика с одного узла на другой позволяет упростить программное решение под названием Hash-AV.

В основе Hash-AV лежит использование хэш-функций и фильтров Блума. Главными преимуществами Hash-AV являются: низкое потребление памяти, высокая точность, необходимость выполнения минимального набора инструкций центрального процессора. Это достигается благодаря способности фильтров Блума достаточно быстро

определить принадлежность того или иного сетевого пакета к вирусной базе данных.

В программном обеспечении Hash-AV фильтр Блума представляет собой вектор размером N бит, значения которых формируются на основе вирусных сигнатур (первоначально всем битам присваивается значение 0). К каждым β байтовым блокам имеющейся сигнатуры применяется соответствующая хэш-функция $h_1(a)$, $h_2(a)$, ..., $h_k(a)$, значения которой лежат в пределах 1.....N (где a – значение сигнатуры). Затем биты, которые находятся в вычисленных позициях $h_1(a)$, $h_2(a)$, ..., $h_k(a)$, устанавливаются равными 1.

Программное обеспечение Hash-AV также способно обнаруживать полиморфные вирусы благодаря использованию эмуляции (исполнения набора инструкций в виртуальной среде).

Выводы

Описанные подходы позволяют по сравнению с существующими методиками соблюдать баланс между затратами и производительностью не только на этапе проектирования сети, но и при ее динамическом обновлении, сохраняя при этом высокий уровень безопасности. Учитывая показанную функциональность программного обеспечения Hash-AV, а также достаточную гибкость перекон-фигурирования компьютерной сети, которую обеспечивает технология виртуализации VLAN, можно сделать вывод о гарантированной эффективности применения данных технологий совместно с описанными алгоритмами для реализации динамической модели анализа трафика.

Литература

1. Odlyzko A. M. Internet traffic growth: Sources and implications. // Optical Transmission Systems and Equipment for WDM Networking 11. 2003. Vol. 5247. P. 1-15.
2. Бабенко Г.В. Анализ современных угроз безопасности информации, возникающих при сетевом взаимодействии, 2010. №2. –С. 149-152.
3. Claffy K. Internet Traffic Characterization / Ph.D. thesis. University of CaHfomia, San Diego. 1994.
4. Хогдал Дж. Скотт. Анализ и диагностика компьютерных сетей. AddisonWesley Longman, Inc., 2000.
5. Beran J. Statistical Methods for Data with Long-Range Dependence. //Statistical Science, Volume 7, Issue 4 (Nov., 1992), 404-416.
6. Harringson D., et al. An Architecture for Describing SNMP Management Frameworks. IETF STD 62, RFC 3411. — 2002. http://www.ietf.org/rfc/rfc3411
7. Эндрю Таненбаум, 2003, «Computer Networks», Pearson Education International, New Jersey.
8. O. Erdogan and P. Cao. Hash-av: Fast virus signature scanning by cache-resident filters. In http://crypto.stanford.edu/c˜ao/hash-av/, 2005.

Лістровий С.В., Моцний С.В. Аналіз мережевого трафіку стека протоколів TCP/IP на основі динамічної моделі. У даній статті розглядається динамічна модель аналізу мережевого трафіку стека протоколів TCP/IP, розробка якої є вкрай важливою в умовах постійного оновлення мережевої інфраструктури. Проведено аналіз існуючих підходів, за допомогою яких можливо забезпечити прийнятний рівень мережевої безпеки. Розглянута практична можливість реалізації алгоритмів знаходження мінімального верхового покриття при побудові динамічної моделі.

Ключові слова: динамічна модель, комп'ютерні мережі, VLAN, мережевий трафік, верхове покриття, стек протоколів TCP/IP, статистичний аналіз.

Listrovoy S.V., Motsnyi S.V. TCP/IP protocol stack network traffic analysis based on the dynamic model. The dynamic model of TCP/IP protocol stack network traffic analysis, the development of which plays a crucial role in the conditions of constantly updated network infrastructure, is considered in this article. The analysis of current approaches which provide acceptable level of network security has been conducted. Practical implementation of minimum vertex cover matching algorithm while dynamic model construction has also been considered.

Key words: dynamic model, computer networks, VLAN, network traffic, vertex cover, TCP/IP protocol stack, statistic analysis.

Рецензент к.т.н., доцент, профессор кафедры СКС Коновалов В.С. (УкрГАЖТ)

Поступила 29.05.2014г.