ВЫЧИСЛИТЕЛЬНАЯ ТЕХНИКА

УДК 004.722.4

М. Ю. Будько

ПОВЫШЕНИЕ БЕЗОПАСНОСТИ РАБОТЫ КОМПЬЮТЕРНЫХ СЕТЕЙ НА ОСНОВЕ АНАЛИЗА ПОТОКОВ ДАННЫХ

Рассматриваются методы обнаружения широковещательных штормов и динамического построения топологии сети на основе анализа потоков данных. Описывается механизм обнаружения широковещательного шторма в сети, его источника и области поражения. Производится сравнение критериев, обеспечивающих поиск идентичных последовательностей значений интенсивности трафика. Разработано программное обеспечение для обработки статистических сведений о функционировании сети.

Ключевые слова: *безопасность сети, широковещательный шторм, топология сети.*

В настоящее время широкое распространение получили сети передачи данных, построенные с использованием технологии Ethernet. Несмотря на то, что физическая структура таких сетей представляет собой дерево, достаточно большие ее сегменты могут быть объединены на втором уровне модели OSI: это приводит к возникновению угроз безопасности, связанных с использованием широковещательных и групповых адресов для организации штормов в сети. В связи с этим возникает задача динамического анализа потоков данных в целях обнаружения источников дестабилизирующего воздействия на сеть и определения области поражения.

Сетевые атаки можно классифицировать по следующим признакам [1]:

— по расположению источника атаки относительно цели:
  — межсегментные, когда субъект и объект атаки находятся в разных сегментах сети; особенностью такой атаки является возможность обнаружения на границе сегмента;
  — внутрисегментные, когда субъект и объект атаки находятся в одном сегменте сети; такие атаки обнаружить сложнее, так как необходимо контролировать трафик между всеми узлами сети;
— по технологии обнаружения:
  — обнаружение известных сигнатур — наиболее распространенный способ в настоящее время, особенностями которого являются высокая надежность и достоверность обнаружения известных атак;
  — обнаружение аномалий — способ, основанный на формировании статистической модели нормального поведения сети и предусматривающий в случае его отклонения от модели формирование предупреждения об атаке; этот способ менее надежен, но в настоящее время приобретает популярность за счет своей универсальности; в частности, в некоторых случаях он незаменим: например, когда отсутствует возможность контроля трафика между пользователями одного сегмента сети.

При внутрисегментных атаках возникают некоторые трудности их обнаружения. Особенность использования метода аномалий для обнаружения внутрисегментных атак состоит в том, что для эффективного обнаружения самой атаки, выявления ее источника и области поражения необходимо знать структуру охраняемой сети. Следовательно, необходимы механизмы для автоматизации построения топологии сети.

Эффективность решения задачи анализа потоков данных зависит от средств и способов мониторинга сетевой инфраструктуры. В крупных сетях в качестве основного источника информации используется протокол SNMP, с помощью которого собираются сведения о загрузке сетевых интерфейсов коммутирующего оборудования. Считается, что этой информация недостаточно для обнаружения угроз безопасности, и, как правило, она используется только в качестве дополнительной при наблюдении за сетевыми ресурсами. Однако, как показали проведенные исследования, даже с использованием столь ограниченных сведений можно повысить безопасность сети и получить дополнительные сведения о ее функционировании.

Сложность анализа потоков данных состоит в том, что устройства в сети опрашиваются системой мониторинга несинхронно, т.е. существует разница во времени между опросом первого и последнего устройств из списка для мониторинга. По умолчанию во многих системах сбора статистических данных максимальный интервал времени задержки составляет 300 с. Для уменьшения влияния задержки при опросе устройств данные интерполируются до момента времени, соответствующего началу опроса первого устройства.

Это приводит к различию статистических данных для двух портов, даже если весь трафик с одного из них поступает на вход другого. Соответственно показания, считанные из базы данных системы мониторинга, являются функцией от их реальных значений:

$$y(t) = f(t) + \varepsilon,$$

где $t = t_1, t_2, \ldots, t_n$ — вектор значений временных меток, соответствующих началу процесса опроса устройств, где $\Delta t = (t_i - t_{i+1})$ — период опроса устройств, n — количество значений в выборке; $y = y(t_1, t_2, \ldots, t_n)$ — вектор сохраненных в системе мониторинга показаний интенсивности (I) трафика, проходящего через порт; $f = f(t_1, t_2, \ldots, t_n)$ — реальные значения интенсивности трафика, проходящего через порт; $\varepsilon = \varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ — вектор случайных компонент, образовавшихся вследствие несинхронного опроса устройств и последующей интерполяции данных.

Было проведено исследование, в рамках которого оценивались следующие критерии обнаружения одинаковых последовательностей значений I [2, 3]:

— коэффициент d, вычисленный на основе сумм абсолютных значений остаточных разностей [4];

— коэффициент r, использующий сумму квадратов отклонений;

— коэффициент q, использующий сумму остаточных разностей, возведенную в третью степень;

— выборочный коэффициент корреляции Пирсона k_P;

— выборочный коэффициент ранговой корреляции Кендалла k_K [5].

В качестве исходных данных использовались показания загрузки интерфейсов сетевых устройств в распределенной сети. При этом была предпринята попытка на основе статистических данных найти связанные друг с другом порты оборудования, так как очевидно, что трафик между ними должен быть одинаковым. Результаты приведены на рис. 1 и 2, где показаны степень совпадения обнаруженных связей (S) с реальной структурой сети и процент ошибочно обнаруженных соответствий (E) в зависимости от объема выборки (N).

Из рис. 2 видно, что наиболее достоверным способом обнаружения одинаковых последовательностей значений интенсивности I трафика является использование коэффициента ранговой корреляции Кендалла, при этом единственное ограничение — объем анализируемой выборки: $N \geq 15$.

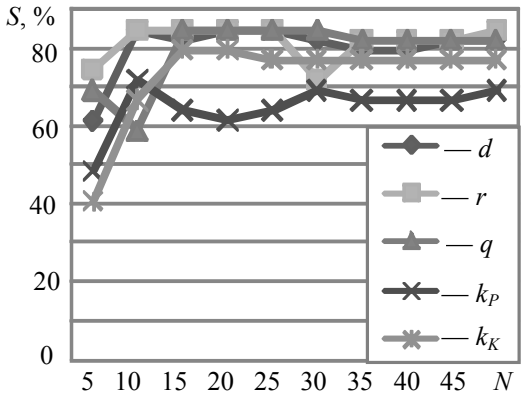

Рис. 1

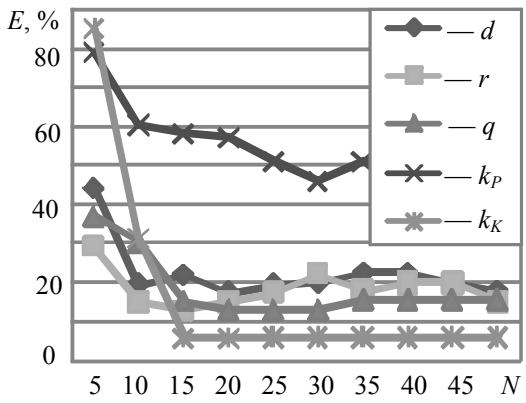

Рис. 2

Следующим этапом анализа является применение аналогичного подхода для обнаружения аномалий значений I на примере широковещательного шторма. Суть этого явления состоит в том, что несанкционированные действия какого-либо узла сети могут привести к увеличению нагрузки на широковещательный сегмент сети, а в некоторых случаях — к перегрузкам и снижению быстродействия других узлов. Это происходит вследствие того, что широковещательный трафик распространяется по всему сегменту и должен быть обработан каждым узлом сети.

Рассмотрим последовательность действий по обнаружению широковещательного шторма:

1) определение уровней иерархии устройств в сети с помощью предварительного построения ее структуры;

2) анализ трафика на устройствах нижнего уровня, так как именно они используются для подключения конечных пользователей;

3) составление списка устройств, у которых разность между средними для каждого порта значениями I предыдущих и последующих отсчетов превышает порог, установленный администратором сети; если разность среднеквадратичных отклонений при этом небольшая, то фиксируется начало широковещательного шторма;

4) определение момента окончания шторма — осуществляется так же как и определение начала шторма, только при этом фиксируется уменьшение средней интенсивности трафика;

5) определение формы шторма по усредненным для каждого порта значениям I за время шторма;

6) анализ исходящего трафика на абонентских портах сетевого оборудования в целях поиска источника шторма, при этом в качестве критерия соответствия используется коэффициент ранговой корреляции Кендалла вследствие его эффективности при обнаружении похожих последовательностей значений I.

Использование рассмотренных методов анализа статистических данных позволяет повысить информированность администраторов о процессах, которые происходят в сети. Средства динамического построения структуры сети можно интегрировать в систему управления и наблюдения за сетью и тем самым упростить ее администрирование. Возможность обнаружения широковещательных штормов позволит своевременно принимать меры для устранения источника дестабилизирующего воздействия.

СПИСОК ЛИТЕРАТУРЫ

1. Библиотека I2R [Электронный ресурс]: Классификация атак. 2002. <http://i2r.ru>.

2. IEEE, 802.1AB. IEEE Standard for Local and Metropolitan Area Networks. Station and Media Access Control Connectivity Discovery. N. Y., 2005.

3. Pat. 5,926,462 USA, МКИ6 H04L 12/28. Method of Determining Topology of a Network of Objects which Compares the Similarity of the Traffic Sequences/Volumes of a Pair of Devices / *D. Schenkel, M. Slavitch, N. Dawes.* 20.07.1999.

4. *Казиев В. М.* Введение в математику и информатику. СПб.: БИНОМ, 2007.

5. *Минько А. А.* Статистический анализ в MS Excel. М.: Изд. дом „Вильямс“, 2004.

Сведения об авторе

Михаил Юрьевич Будько — аспирант; Санкт-Петербургский государственный университет информационных технологий, механики и оптики, кафедра мониторинга и прогнозирования чрезвычайных ситуаций; E-mail: bmu@mail.ru

Рекомендована кафедрой мониторинга и прогнозирования чрезвычайных ситуаций

Поступила в редакцию 29.04.08 г.