УДК 004.722.4

ОТСЛЕЖИВАНИЕ ИЗМЕНЕНИЙ В СТРУКТУРЕ СЕТИ И РЕШЕНИЕ ЗАДАЧ ПОВЫШЕНИЯ БЕЗОПАСНОСТИ НА ОСНОВЕ АНАЛИЗА ПОТОКОВ ДАННЫХ

М.Б. Будько, М.Ю. Будько

Рассматриваются методы обнаружения широковещательных штормов и динамического построения топологии сети на основе анализа потоков данных. Рассмотрен механизм обнаружения широковещательного шторма в сети, его источника и области поражения. Производится сравнение критериев, обеспечивающих поиск похожих шаблонов трафика.

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

Введение

Методы обнаружения широковещательных штормов и динамического построения топологии сети актуальны в современных компьютерных сетях, построенных с использованием коммутаторов. В настоящее время широкое распространение получили сети передачи данных, построенные на основе технологии Ethernet. Несмотря на то, что физическая топология таких сетей представляет собой дерево, достаточно большие 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 [1]. В связи с этим возникает задача динамического анализа потоков данных с целью обнаружения источников дестабилизирующего воздействия на сеть и определения области поражения.

Существующие способы обнаружения широковещательных штормов сводятся к определению интенсивности передачи широковещательных пакетов через конкретный порт коммутатора. При превышении некоторого порога пакеты начинают отбрасываться. Однако не все коммутаторы поддерживают такие функции, и, как показывает опыт, они не всегда работают корректно.

Эффективность решения задачи анализа потоков данных зависит от средств и способов мониторинга сетевой инфраструктуры. Как правило, в крупных сетях в качестве основного источника информации используется протокол SNMP. С его помощью собираются сведения о загрузках сетевых интерфейсов коммутирующего оборудования. Сложность анализа потоков данных состоит в том, что устройства опрашиваются системой мониторинга не синхронно, т.е. существует разница во времени между запросом статистики у первого и последнего устройства в списке мониторинга. Для сниже-

ния влияния задержки данные интерполируются до того момента времени, когда начался опрос первого устройства. Это приводит к несоответствию показаний статистики 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 реальных значений:

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

где $t = (t_1, t_2, \ldots, t_n)$ – вектор значений временных меток, во время которых инициируется процесс опроса устройств, $\Delta t = (t_i - t_{i+1})$ – период опроса устройств, n – количество значений в выборке; $y = y(t_1, t_2, \ldots, t_n)$ – вектор показаний трафика на порту, сохраненных в системе мониторинга; $f = f(t_1, t_2, \ldots, t_n)$ – реальные значения трафика на порту; $\varepsilon = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)$ – вектор случайных компонент, образовавшихся вследствие несинхронного опроса устройств и последующей интерполяции данных.

Поиск связей между устройствами в сети

Проведено исследование, в рамках которого оценивались следующие критерии обнаружения одинаковых последовательностей трафика:

– коэффициент, вычисленный на основе сумм абсолютных значений остаточных разностей,

$$s_{ab} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_a(i) - y_b(i)}{\max(y_a(i), y_b(i))}\right|;$$

– коэффициент, использующий сумму квадратов отклонений,

$$s_{ab} = \sqrt[2]{\frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_a(i) - y_b(i)}{\max(y_a(i), y_b(i))}\right|^2};$$

– коэффициент, использующий сумму остаточных разностей, возведенную в третью степень,

$$s_{ab} = \sqrt[3]{\frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_a(i) - y_b(i)}{\max(y_a(i), y_b(i))}\right|^3};$$

– выборочный коэффициент корреляции Пирсона

$$r = \frac{\sum_{i=1}^{n}\left(x_i - \frac{1}{n}\sum_{i=1}^{n}x_i\right)\times\left(y_i - \frac{1}{n}\sum_{i=1}^{n}y_i\right)}{\sqrt{\sum_{i=1}^{n}\left(x_i - \frac{1}{n}\sum_{i=1}^{n}x_i\right)^2 \times \sum_{i=1}^{n}\left(y_i - \frac{1}{n}\sum_{i=1}^{n}y_i\right)^2}};$$

– выборочный ранговый коэффициент корреляции Кендалла [2]

$$r_K = \frac{2}{n(n-1)}\sum_{i=1}^{n}\sum_{j=i+1}^{n} sign\left(q_{(j)} - q_{(i)}\right).$$

В качестве исходных данных использовались показания загрузки интерфейсов сетевых устройств в распределенной сети. При этом делалась попытка на основе показаний статистики найти связанные друг с другом порты оборудования, так как очевидно, что трафик между ними должен быть одинаковым. Результаты исследования рассмотрены в [3]. Из них можно сделать вывод о том, что наиболее достоверным способом обнаружения одинаковых последовательностей трафика является использование коэф-

фициента ранговой корреляции Кендалла. Его единственным ограничением является требование к объему анализируемой выборки (не менее 15 значений).

Обнаружение широковещательных штормов

Следующим этапом является применение похожего подхода для обнаружения аномалий трафика 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. Пример широковещательного трафика приведен на рис. 1.

Из рис. 1 видно, что широковещательный шторм имеет свойства, которые могут использоваться для его обнаружения:

– распространяется по всем портам коммутирующего оборудования в рамках одного сегмента;
– является исходящим трафиком для источника и входящим для всех остальных узлов сегмента.

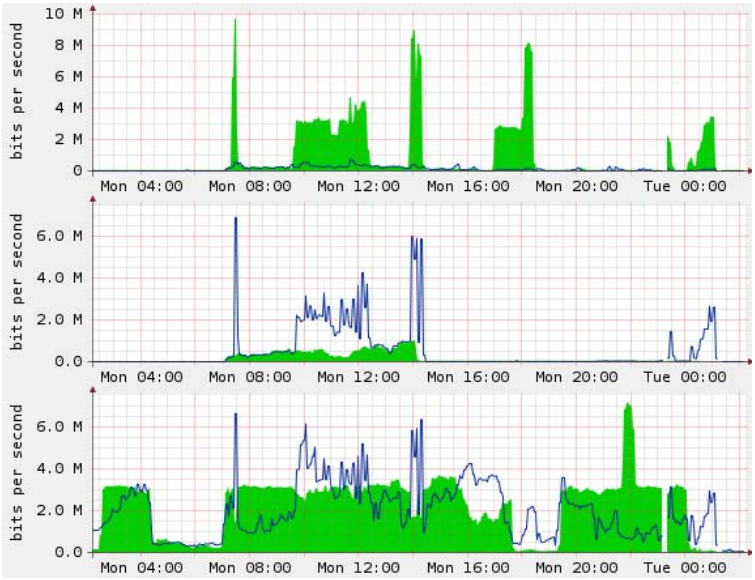

Рис.1. Распространение широковещательного шторма в сети

Следовательно, можно определить последовательность действий по обнаружению широковещательного шторма.

1. Определение уровней иерархии устройств в сети с помощью предварительного построения структуры сети. Построение осуществляется динамически путем сравнения интенсивностей трафика на портах коммутирующих устройств [4]. При этом в качестве коэффициента похожести используется ранговый коэффициент корреляции Кендалла.
2. Анализ трафика на устройствах, находящихся на самом нижнем уровне каждой ветки дерева структуры. Это позволит, с одной стороны, проанализировать все широковещательные домены и, с другой стороны, исключить из рассмотрения сетевое оборудование уровня распределения, так как обнаружить шторм на этом уровне существенно сложнее, чем на уровне доступа.
3. Определение списка устройств, у которых разница между средними на порт значениями трафика предыдущих и последующих отсчетов превысила порог, установленный администратором сети:

$$d = \frac{1}{n_1}\sum y(i) - \frac{1}{n_2}\sum y(i-1),$$

где d – разница между средними значениями трафика для конкретного устройства, $y(i)$ – значение отсчета с номером i на одном из активных портов, n_1 и n_2 – количество активных портов в моменты времени, соответствующие отсчетам с номерами i и i–1.

4. Широковещательный трафик воздействует на все активные интерфейсы коммутаторов. Поэтому, несмотря на повышение средней загрузки, ее распределение по портам должно быть равномерным. Для определения этого вычисляем среднеквадратичные отклонения интенсивностей трафика:

$$d_s = \sqrt{\frac{1}{n_1}\sum\left(y(i) - \frac{1}{n_1}\sum y(i)\right)^2} - \sqrt{\frac{1}{n_2}\sum\left(y(i-1) - \frac{1}{n_2}\sum y(i-1)\right)^2},$$

где d_s – разница между среднеквадратичными отклонениями, $y(i)$ – значение отсчета с номером i на каком-либо активном порту, n_1 и n_2 – количество активных портов в моменты времени, соответствующие отсчетам с номерами i и i–1. Если разница среднеквадратичных отклонений будет небольшой, то фиксируется начало широковещательного шторма.

5. Момент окончания шторма определяется, как и начало, только при этом фиксируется уменьшение среднего трафика.
6. Форма шторма определяется по усредненным на порт значениям трафика за время шторма. Это возможно потому, что изменение формы широковещательного трафика будет в большей степени отражаться на суммарном трафике, чем изменение трафика на каком-либо одном интерфейсе.
7. Для поиска источника шторма просматриваем исходящий трафик на всех сетевых интерфейсах в сети. При этом осуществляется сравнение с шаблоном широковещательного шторма. В качестве критерия соответствия используем ранговый коэффициент корреляции Кендалла, который оказался эффективным для обнаружения похожих последовательностей трафика.

Заключение

Использование рассмотренных методов анализа статистики позволяет повысить информированность администраторов о процессах, которые происходят в сети. Средства динамического построения структуры 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.

Литература

1. Библиотека I2R [Электронный ресурс]. Классификация атак. – Режим доступа: http://i2r.ru, свободный.
2. Минько А.А. Статистический анализ в MS Excel. – М.: Издательский дом «Вильямс», 2004. – 448 с.
3. Будько М.Ю. Сравнение эффективности критериев для обнаружения связей между устройствами в сети // Труды XV Всероссийской научно-методической конференции «Телематика'2008». – 2008. – Том 1.
4. Пат. 5,926,462 USA, МКИ[6] 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. David Schenkel, Michael Slavitch, Nicholas Dawes, 16.11.1995, 20.07.1999.

5. IEEE 802.1AB. IEEE Standard for Local and metropolitan area networks. Station and Media Access Control Connectivity Discovery. – New York, 2005.
6. Казиев В.М. Введение в математику и информатику. – СПб: БИНОМ. Лаборатория знаний; Интернет-университет информационных технологий – ИНТУИТ.ру, 2007. – 304 с.

Будько Михаил Юрьевич — Санкт-Петербургский государственный университет информационных технологий, механики и оптики, инженер, bmu@mail.ru

Будько Марина Борисовна — Санкт-Петербургский государственный университет информационных технологий, механики и оптики, преподаватель, budkomb@mail.ru