05.00.00 Engineering Science

05.00.00 Технические науки

UDC 004.738.5.057.4

IDENTIFICATION OF NETWORK ABNORMALITIES USING METHODS OF STATISTICAL ANALYSIS

[1] German V. Babenko
[2] Sergey V. Belov

[1] Astrakhan State Technical University
16, Tatischeva street, Astrakhan, 414025
PhD student
[2] Astrakhan State Technical University
16, Tatischeva street, Astrakhan, 414025
PhD (technical), Assistant of professor
E-mail: babenko.german@mail.ru

The paper deals with the application of statistical analysis techniques to identify unusual conditions in the network traffic flow and to integrate them into a comprehensive approach for monitoring network infrastructure.

Keywords: network traffic, statistical methods, security.

Современный вектор развития общества тесно связан с процессом информатизации и совершенствования инфокоммуникационных технологий. Одновременно с всеобщей информационной интеграцией стремительно возрастает количество информации передаваемой с использованием сетевых технологий, что влечет за собой рост количества угроз информационной безопасности, возникающих при межсетевом взаимодействии [1]. Для решения задачи выявления неправомерных действий в основном применяют сигнатурный и поведенческий подход, имеющие как, достоинства, так и явные недостатки. Рассмотрим один из так называемых поведенческих методов выявления атак – статистический анализ.

Статистические методы обнаружения попыток нарушения безопасности в сети основаны на изменении некоторых статистических характеристик потока пакетов. Так, например, в случае Flood-атаки возрастает трафик, при сканировании сети увеличивается доля пакетов определённого типа (SYN, ACK, FIN) и т.п. Стоит отметить, что одним из основополагающих преимуществ применения поведенческих методов является их способность выявлять впервые реализуемые атаки, что является одной из глобальных целей работы. Для решения задачи применения статистических методов анализа TCP/IP трафика, необходимо выделить основные показатели, характеризующие штатное функционирование сетевой инфраструктуры и осуществлять динамический контроль над их состоянием. В качестве таких показателей должна выступать информация, по которой можно проанализировать историю межсетевого взаимодействия. В этом случае методы обнаружения нарушений основываются на сравнении текущих, локальных характеристик потока пакетов с обработанными за продолжительный промежуток времени, глобальными характеристиками.

Из всего спектра методик статистического анализа для реализации в работе были выбраны пороговая методика и методика средних и среднеквадратических

отклонений [2]. Основой идеей методик является выявление отклонений в установленных интервалах на основе математического ожидания и дисперсии или же превышение заранее установленного обоснованного абсолютного порогового значения в анализируемых параметрах. Если локальные характеристики сильно отличается от соответствующих глобальных характеристик, то это свидетельствует об аномальном поведении потока пакетов и вполне вероятна попытка сканирования сети или сетевой атаки. Для решения этой задачи необходимо реализовать эффективные методы вычисления локальных статистических характеристик в течение некоторого ограниченного интервала времени и определение порога отклонения локальных характеристик от глобальных статистических характеристик потока.

Рассмотрим данные, которые могут быть проанализированы при захвате трафика TCP/IP. К ним относятся поля заголовков протоколов IP, TCP, UDP, ICMP и содержимое полей данных [3]. Для определения локальных и глобальных характеристик потока сетевого трафика, необходимо определить взаимодействующие логические сущности. Установим под логической сущностью потока IP-адреса и порты отправителя и получателя пакетов. Таким образом, методы статистического анализа будут применяться на определенном временном интервале для сетевых пакетов, имеющих одинаковые логические сущности, что позволит более детально проанализировать статистику сетевых взаимодействий.

Для реализации эффективного метода вычисления локальных статистических характеристик в работе предлагается набор весовых функций. В качестве статистических характеристик потока пакетов в сети применены выборочное среднее, выборочная дисперсия и критерий согласия χ^2 потока в сети и связь между этими признаками при использовании различных статистических характеристик.

Пусть числовая величина $x_i, x_{\min} < x_i < x_{\max}$ представляет некоторое событие из потока событий произошедшее в момент времени t_i. Весь набор событий характеризуется средним значением и дисперсией величины x_i. Общее количество событий N определяется интервалом времени, в течение которого ведется наблюдение. При увеличении числа событий среднее арифметическое стремится к математическому ожиданию величины и может быть использовано в качестве глобальной, долговременной характеристики потока. В случае появления отклонения к общему «нормальному» фону функционирования подмешиваются значения, отличающиеся по абсолютной величине, что вызывает дифференциацию данных и является причиной изменения характеристик.

Результаты анализа позволяют полагать, что применение статистическим методов анализа отдельных характеристик проявляет постоянство результатов на «нормальных» данных, а по их изменению возможно зафиксировать нарушения. Для каждой из используемых статистических характеристик можно сформулировать свой критерий присутствия аномалии в потоке событий.

При реализации вышеописанных методов в автоматизированной системе были созданы хранилища данных, содержащие неадаптированные глобальные характеристики сетевого потока. Для улучшения визуального восприятия получаемых результатов анализа разработана двумерная динамическая характеристика сетевой активности (рис. 1).

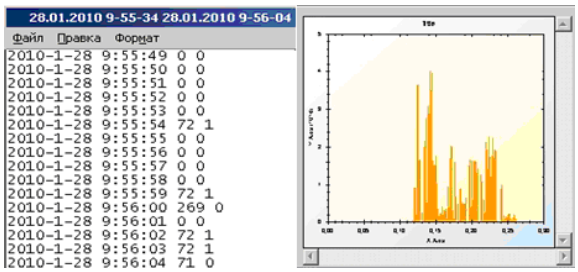

Рис. 1 Характеристика TCP/IP: объемная

А так же частотная характеристика появления типов пакетов (рис. 2).

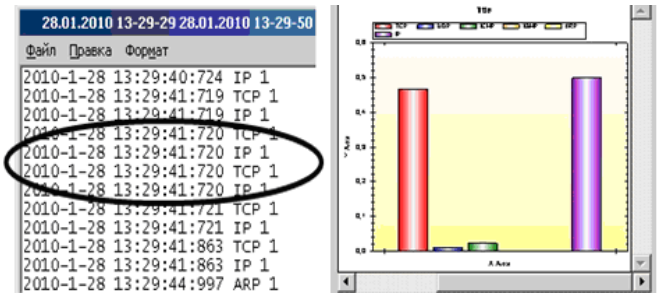

Рис. 2 Характеристика TCP/IP: частотная

Сопоставляя результаты, можно сделать вывод, что диапазоны значений характеристик, не содержащих следов отклонений, имеют относительное постоянство оценок математических ожиданий и дисперсий. При возникновении нетипичной ситуации аналогичные оценки по различным характеристикам изменяются существенно. Стоит отметить, что эффективность применения данных методов зависит от временного периода анализа, составляющего порядка 30 мин на итерацию. Для уменьшения вероятности появления ошибок первого рода, рекомендуется применять иные компонента независимые методы анализа трафика. Иной особенностью применения данных методов является возможность определения проблем не только внешнего воздействия, но и вопросов корректного функционирования компонентов сетевой инфраструктуры.

Примечания:

1. Бабенко Г.В. Анализ современных угроз безопасности информации, возникающих при сетевом взаимодействии. // Вестник Астраханского государственного технического университета. Серия: "Управление, вычислительная техника и информатика", 2010. №2. С. 149-152.

2. Андронов А.М., Копытов Е.А. Теория вероятностей и математическая статистика. СПб.: «Питер», 2004. 461 с.

3. Олифер В.Г., Олифер Н.А. Компьютерные сети. Принципы, технологии, протоколы. Учебник для вузов. 4-е изд. СПб.: Питер, 2010. 944 с.

УДК 004.738.5.057.4

ВЫЯВЛЕНИЕ СЕТЕВЫХ ОТКЛОНЕНИЙ СТАТИСТИЧЕСКИМИ МЕТОДАМИ АНАЛИЗА

[1] Герман Валерьевич Бабенко
[2] Сергей Валерьевич Белов

[1] Астраханский государственный технический университет
414025, г. Астрахань, ул. Татищева, 16
Аспирант
[2] Астраханский государственный технический университет
414025, г. Астрахань, ул. Татищева, 16
Кандидат технических наук, доцент
E-mail: babenko.german@mail.ru

В работе рассматриваются вопросы применения статистических методов анализа с целью определения нетипичных состояний в потоке сетевого трафика и интеграции их в комплексный подход мониторинга сетевой инфраструктуры.

Ключевые слова: сетевой трафик, статистические методы, безопасность.