УДК 681.3

Р.Р. ФАТКИЕВА

КОРРЕЛЯЦИОННЫЙ АНАЛИЗ АНОМАЛЬНОГО СЕТЕВОГО ТРАФИКА

Фаткиева Р.Р. **Корреляционный анализ аномального сетевого трафика.**
Аннотация. В статье рассматриваются методы моделирования поведения информационной системы. Использование моделей динамических систем сопряжено с трудностями определения функциональной зависимости временного распределения от большого числа параметров, а также отсутствием начальных данных. Предложен подход к обнаружению атак, основывающийся на анализе отклонений от автокорреляционной функции.
Ключевые слова: сетевой трафик, атаки, автокорреляционная функция.

Fatkieva R.R. **Correlation analysis of abnormal network traffic.**
Abstract. The paper deals with the simulation methods of information system behavior. Attention is paid to the difficulties of dynamic system models application which are accounted for functional dependence of temporal distribution on a large number of parameters as well as for the lack of initial data. An approach to attack detection is proposed on the basis of analysis of deviation from autocorrelation function.
Keywords: network traffic, attacks, autocorrelation function.

1. Введение. Существующие работы по оценке сетевого трафика связаны как с реальными измерениями характеристик информационных систем (ИС), так с моделированием физических механизмов и свойств системы при их проектировании и эксплуатации. Данные о сетевом трафике анализируются с целью выявления аномального поведения, производятся количественные оценки характеристик трафика, определяются пороговые значения той или иной характеристики [1, 2]. При этом поведение ИС зависит от её структуры, аппаратных и программных решений, алгоритмов, заложенных в работе ИС, и количества информации, циркулирующей в тот или иной момент времени. Для оценки поведения ИС важно знать макрохарактеристики, зависящие от набора микросостояний на каждом отдельном промежутке времени. Например, получение определенного трафика за тактовую единицу времени можно назвать микросостоянием, а выделение тренда сетевого потока на множестве микросостояний – макросостоянием. Для формализации перехода от микросостояний к макросостоянию необходимо определить или задать функцию распределения состояний ИС. Статистическому моделированию поведения ИС посвящен ряд работ [1-7], в которых используется понятие параметров порядка, описывающих поведение ИС. Однако, при их использовании возникает проблема, связанная с установлением взаимосвязи с функцией распре-

деления. Исторически сложилось, что рассмотренные в 60-70 гг. XX века статистические оценки сетевого трафика хорошо известны и широко распространены, однако реалии сетевого трафика, характерного для 90-х гг. XX века и начала XXI века не вписывались в рамки классических математических моделей потоков. Связано это с тем, что реальные процессы в компьютерных системах как минимум нестационарны, т.е. не соответствуют условиям применимости классических моделей [3]. Для моделирования нестационарных процессов в компьютерных сетях стали применяться методы нелинейной динамики [4, 5], что позволило выявить в этих процессах закономерности степенного распределения и описать тяжелые хвосты распределений. Однако, использование моделей динамических систем сопряжено с трудностями определения функциональной зависимости временного распределения от большого числа параметров, а также с отсутствием начальных данных. Для упрощения задачи динамического моделирования применяют способ эвристического задания функции распределения, используя свойства реальных сетевых потоков. Так, в частности, исследования 2000-2012 гг. показывают, что в ряде случаев распределение сетевого трафика имеет фрактальную размерность [3].

Существует множество локальных состояний ИС, приводящих к одному и тому же распределению. Например, если каждое состояние описывать вероятностью перехода от одного состояния к другому, то, вычислив множество значений трафика в каждой точке траектории динамического поведения ИС, можно найти распределение вероятностей перехода и, соответственно, определить устойчивые состояния ИС. Динамическое описание устойчивости ИС позволяет дать качественную оценку текущего состояния процесса управления (например, «устойчивое», «рискованное», «неуправляемое»). При этом неустойчивое поведение ИС можно описать участками «аномального» поведения, которые, как правило, имеют короткую длительность. В то же время для проверки на «нормальность» необходимы большие массивы временных рядов. Может возникнуть ситуация, когда внутри интервала, используемого для проверки, окажется большое количество участков с разным поведением. При расчете тестовых значений характеристики персистентных и антиперсистентных участков компенсируются, и итоговые значения оказываются близкими к нормальным. Этим и объясняется различие в полученных авторами распределениях при анализе сетевого трафика. Не исключено и предположение о наличии перехода из одного вида распределения в другое при возникновении несанкционированной активности того или иного процесса, что и по-

рождает необходимость дальнейшего исследования характеристик сетевого трафика.

2. Характеристики объекта анализа и алгоритм процесса. Для применения статистических методов анализа сетевого трафика необходимо выделить основные характеристики штатного функционирования ИС и осуществлять динамический анализ их состояния. К данным, которые могут быть исследованы при анализе сетевого трафика, можно отнести: поля заголовков протоколов, содержимое полей данных, длина пакета (средняя, максимальная, минимальная), средняя длительность пикового трафика, средняя длительность сеанса связи, интенсивность запросов, количество потерянных пакетов. Численными характеристиками, определяющими внутреннее состояние системы, являются: уровень загрузки процессора, уровни загрузки оперативной памяти (физической и файла подкачки) и каналов.

Сравнение характеристик работы ИС в штатном режиме и режиме под воздействием атаки показало, что распределения значений частот обращения к разным ресурсам на одном и том же временном отрезке различаются. В качестве наглядной интерпретации результатов можно представить гистограммы одних и тех же величин, измеренных в штатном режиме и в режиме HTTP-flood (рис. 1).

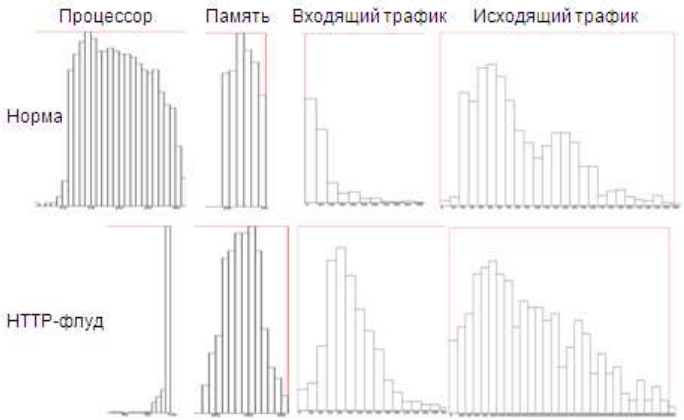

Рис.1. Гистограммы загрузки.

Различие плотностей распределения трафика на рис.1 показывает наличие в некоторых случаях качественного изменения функций (т. е. изменения закона распределения), особенно заметного на уровне за-

грузки процессора. Также наблюдается повышение уровня загрузки памяти и изменение средних значений сетевого трафика, что позволяет идентифицировать несанкционированную активность. В работах [1,6,7] показана возможность определения несанкционированной активности на основе оценки порогового значения, математического ожидания и среднеквадратичного отклонения выборки. В процессе обнаружения аномального поведения ИС на основе среднеквадратичного отклонения анализ экспериментальных данных показал возможность выявления несанкционированной активности также и по отклонениям от автокорреляционной функции, представленным на рис. 2, 3.

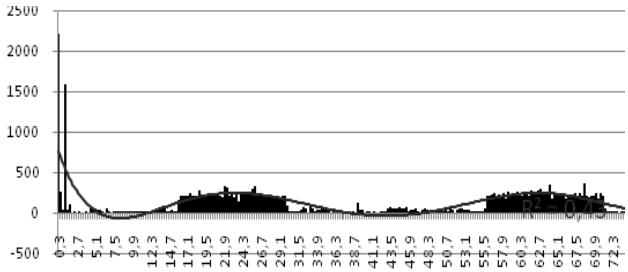

Рис.2.Вид автокорреляционной функции при штатном функционировании.

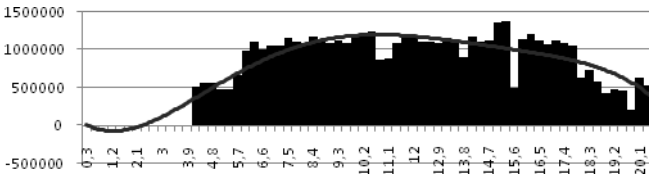

Рис.3. Расчет автокорреляционной функции трафика с атакой TCP flood.

Анализ автокорреляционной функции (рис. 2-4) позволяет не только определить наличие несанкционированной активности, но и сделать предположение о виде атаки.

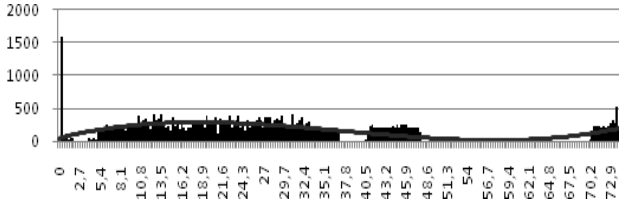

Рис.4. Расчет автокорреляционной функции трафика с атакой SYN flood.

Так, при атаке возникает уменьшение автокорреляции, что свидетельствует о нарушении стационарности процесса. С другой стороны, фрактальный анализ поведения ИС показал увеличение показателя Херста, представленного в таблице. Возможно, это связано с увеличением количества «однотипных пакетов» в сетевом трафике.

Характеристики сетевого трафика

Вид трафика	СКО	M	N	Dmax	Показатель Херста
Нормальный	0,314	125,712	250	4.795	0,53
SYN-flood	0,531	164,06	250	4.795	0,95
TCP-flood	0,738	200,02	250	4.795	0,79

Следует заметить, что наличие больших всплесков на фоне относительно низкого среднего уровня трафика при функционировании ИС в штатном режиме можно принять за атаку. Во избежание ошибок второго рода (ложное срабатывание систем защиты) при формировании реакции на атаку необходимо учитывать в том числе самоподобие сетевого трафика.

3. Заключение. Предложенный метод определения несанкционированной активности по отклонениям от автокорреляционной функции позволяет определить не только наличие несанкционированной активности, но и предположить вид атаки. На основании полученной информации можно предусмотреть использование соответствующих способов противодействия, направленных на борьбу с конкретной атакой.

Литература

1. *Бабенко Г.В., Белов С.В.* Анализ трафика TCP/IP на основе методики допустимого порога и отклонения как инструмент определения информационной безопасности // Интернет-журнал "Технологии техносферной безопасности" (http://ipb.mos.ru/ttb), Выпуск № 5 (39).
2. *Воробьев В.И., Евневич Е.Л., Фаткиева Р.Р.* Моделирование сетевого трафика методом Монте-Карло // Вестник Бурятского государственного университета. 2010. №9. С. 258–262.
3. *Шелухин О.И., Осин А.В., Смольский С.М.* Самоподобие и фракталы. Телекоммуникационные приложения/ Под. Ред. О.И. Шелухина.-М.:Физматлит, 2008.-328 С.
4. *Владимирский Э.И, Исмайлов Б.И.* Синергетические методы управления хаотическими системами. Баку, «EML» 2011.-240 с.
5. *Шишкин В,М.* Степенное распределение и управление рисками критических систем//Труды Института системного анализа Российской академии наук. 2007. Т. 31. С. 42–63.
6. *Фаткиева Р.Р.* Модель обнаружения атак на основе анализа временных рядов // Труды СПИИРАН. 2012. Вып. 21. С. 71–79.
7. *Мишин К.Н., Фаткиева Р.Р.* Природа сетевых аномалий и их полунатурное моделирование // Труды СПИИРАН. 2007. Вып. 5. С. 260–267.

Фаткиева Роза Равильевна — канд. техн. наук; старший научный сотрудник лаборатории информационно-вычислительных систем СПИИРАН. Область научных интересов: моделирование информационных систем. Число научных публикаций — 26. rikki2@yandex.ru; СПИИРАН, 14-я линия В.О., д. 39, г. Санкт-Петербург, 199178, РФ; р.т. +7(812)328-4369, факс +7(812)328-4450.

Fatkieva Rosa Ravilievna — senior researcher, Laboratory of Information and Computer Systems, SPIIRAS. Research interests: modeling of information systems. Number of publications — 26. rikki2@yandex.ru; SPIIRAS, 39, 14-th Line V.O., St. Petersburg, 199178, Russia; office phone +7(812)328-4369, fax +7(812)328-4450.

Рекомендовано лабораторией информационно-вычислительных систем СПИИРАН, заведующий лабораторией Воробьев В.И, д-р техн. наук, проф.
Статья поступила в редакцию 08.10.2012.

РЕФЕРАТ

Фаткиева Р.Р. **Корреляционный анализ аномального сетевого трафика.**

Моделирование поведения информационной системы сопряжено с трудностями определения функциональной зависимости временного распределения от большого числа параметров, а также с отсутствием начальных данных.

Показано, что для применения статистических методов анализа сетевого трафика необходимо выделить основные характеристики штатного функционирования информационной системы и осуществлять динамический анализ их состояния. Исследование изменения характеристик работы ИС в штатном режиме и режиме под воздействием атаки показало, что распределения значений частот обращения к разным ресурсам на одном и том же временном отрезке различаются. Различие плотностей распределения трафика указывает на наличие в некоторых случаях качественного изменения функций (т. е., изменения закона распределения).

Предложен метод определения несанкционированной активности по отклонениям от автокорреляционной функции, позволяющий определить не только наличие несанкционированной активности, но и предположить вид атаки.

На основании полученной информации можно предусмотреть использование соответствующих способов противодействия, направленных на борьбу с конкретной атакой.

SUMMARY

Fatkieva R. R. **Correlation analysis of abnormal network traffic.**

Simulation of information system behavior is associated with the problem of determination of functional dependence of temporal distribution on a large number of parameters as well as with that of the lack of initial data.

It is shown that in order to apply statistical techniques to network traffic analysis it is essential to determine principal characteristics of normal functioning of information system and to fulfill their dynamic analysis. Comparison of the system characteristics in normal mode and under attack demonstrates the difference of distributions of values of frequencies of access to various resources within the same temporal interval. Difference of traffic distribution densities is an evidence of qualitative changes of functions (i.e., changes in distribution law).

Technique is proposed for unauthorized activities detection by deviations from autocorrelation function, unauthorized activity occurrence is being revealed along with the identification of a type of the attack.

Taking into account the above information it is possible to provide appropriate counteraction measures to the specific attack type.