УДК 004.7 : 004.31

Т.Н. Шипова[1], В.В. Босько[2], И.А. Березюк[2], Ю.М. Пархоменко[2]

[1] *Национальный технический университет «ХПИ», Харьков*
[2] *Кировоградский национальный технический университет, Кировоград*

АНАЛИЗ СОВРЕМЕННЫХ МЕТОДОВ ОБНАРУЖЕНИЯ ВТОРЖЕНИЙ В КОМПЬЮТЕРНЫЕ СИСТЕМЫ

В работе представлены результаты анализа методов обнаружения вторжений в компьютерные системы, выявлены их основные достоинства и недостатки. Предложены пути повышения эффективности эвристического анализа компьютерных систем с использованием свойства самоподобия в характеристиках информационного трафика. Определены приоритетные пути учета свойства самоподобия трафика при анализе состояния компьютерной системы в условиях возможных внешних воздействий злоумышленников.

Ключевые слова: *анализ трафика, самоподобие, аномалии, фрактальная размерность, показатель Херста.*

Введение

Постановка задачи. Стремительное развитие вычислительной техники привело к тому, что компьютерная сеть стала использоваться как полнофункциональное распределённое вычислительное устройство для обработки и передачи данных.

В настоящее время проводится большое количество исследований, имеющих целью выявление характерных свойств и параметров технологии передачи данных, а также поиск оптимально гарантированных способов обнаружения аномалий в работе компьютерной сети, которые могут каким-либо образом повлиять на протекающие в ней процессы.

На фоне стремительного развития и модификации вредоносного программного обеспечения (ПО) и повышения уровня хакерства, антивирусное ПО не всегда может в полной мере защитить пользователей компьютерных сетей от действий злоумышленников. Всё чаще наряду со статистическими методами, анализирующими соответствие конкретного действия в сети определённым шаблонам и записям журналов (брандмауэры), используется анализ поведения трафика во время работы сети. Анализируя определённые параметры, можно выявить в какой момент времени в поведении трафика появляются аномальные изменения (всплески). Исследование анализа трафика в компьютерной сети сегодня очень актуально, так как такой анализ может выявить атаку злоумышленника еще на этапе её подготовки.

Цель данной статьи – анализ основных методов обнаружения вторжений в компьютерные системы и сети.

Проведенный анализ литературы [1 – 5] показал, что в современных компьютерных сетях остро стоит вопрос о своевременном обнаружении вторжений.

Ряд научных статей [1, 2 – 5] посвящены исследованиям в области нахождения оптимальных методов обнаружения аномалий в сети, наличие которых обычно есть или предшественниками вредоносного действия, или же указывают на наличие уже совершающейся атаки, а так же могут указывать на неправильную работу (неисправность) оборудования.

Основной материал

В общем случае сетевые аномалии можно классифицировать по источнику их возникновения (рис. 1).

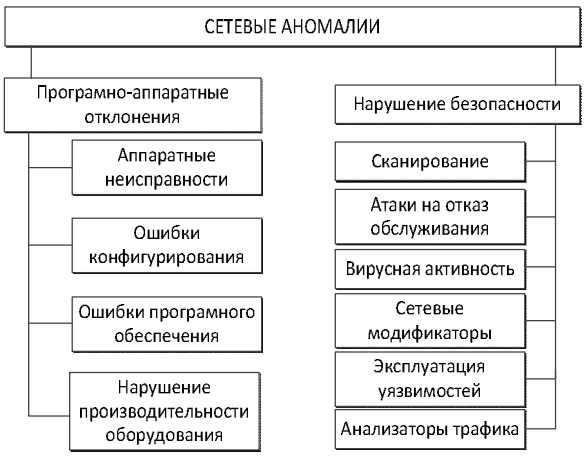

Рис. 1. Источники возникновения сетевых аномалий

Системы обнаружения вторжений можно разделить на системы, ориентированные на поиск:

– аномалии взаимодействия контролируемых объектов;

– сигнатур всех узнаваемых атак;

– искажение эталонной профильной информации.

Методы обнаружения аномалий направлены на выявление неизвестных атак и вторжений. В обнаружении аномалий значительную роль играет выбор оптимального количества учитываемых параметров оценки, а также определение общего показателя общего состояния аномальности в защищённой системе [4].

В поиске аномалий взаимодействия контролируемых объектов в качестве объекта наблюдения может выступать сеть в целом, отдельный компьютер, сетевая служба (FTP-сервер), отдельный пользователь и т.д.

При анализе взаимодействия контролируемых объектов для выявления аномалий (атак) учитываются следующие атрибуты:

– пороговые значения – количественные величины использования обращений к определённым службам и файлам, число неудавшихся попыток входа в систему, загрузка процессора, присущие по статистике конкретным пользователям;

– статистические меры – методом сбора статистики принимается решение о наличии или отсутствии аномалий (атак);

– параметрические – шаблон профиля «нормальной системы» для выявления «не нормальных» отклонений;

– не параметрические – профиль системы строится на основе наблюдения за объектом в период обучения;

– метод на основе правил (сигнатур) – в период обучения составляется представление о нормальном поведении объекта, которое записывается в виде специальных правил.

Смысл сигнатурных методов заключается в использовании специальной базы данных, которая содержит шаблоны (сигнатуры) атак, для поиска действий попадающих под определение «угроза». Таким образом, такой метод защиты может быть действенным, если вид угрозы уже занесён в базу данных, т.е. новый вид опасного воздействия может остаться незамеченным. Но и здесь есть определённые нюансы, например, если угроза исходит от вредоносного ПО, которое внедрено злоумышленником в приложение, разрешенное правилами метода. Так как разновидность вредоносного ПО и атак довольно таки обширна и с каждым днём прогрессирует, то такой метод не оптимален для защиты от действий злоумышленников.

Системы, ориентированные на поиск искажений эталонной профильной информации, основываются на сборе статистики и относятся к поведенческим методам определения нарушений в сети, они основаны на сопоставлении текущего состояния сетевой инфраструктуры с некими определёнными заранее признаками, характеризующими штатное функционирование сетевой инфраструктуры.

Недостатки этого метода могут быть:

– нечувствительность к последовательности возникновения событий, т.е. вероятность возникновения вторжения, если оно происходит в виде последовательности сходных событий;

– систему можно обучить таким образом, что аномальное поведение будет считаться нормальным;

– трудность с определением порога, выше которого аномалии можно рассматривать как вторжение;

– ограниченность к типам поведения, которые могут быть смоделированы [1].

В то же время похожий по принципу действия метод анализа трафика позволяет улучшить защиту вычислительной сети. В отличие от сигнатурного метода, когда разнообразие угроз иногда просто нереально полностью классифицировать, анализ поведения трафика в сети более поддается классификации, т.е. не зависимо от того какой тип атаки происходит, виды аномалий всё же не так разнообразны.

Стоит отметить, что анализ трафика жизненно важен для эффективного управления сетью. Он является источником информации о функционировании корпоративных приложений, которая учитывается при распределении ресурсов, планировании вычислительных мощностей, определении и локализации отказов, решении вопросов безопасности [4].

Для анализа трафика используют программы анализаторы (снифферы), которые могут выполнять:

– мониторинг сетевых интерфейсов и сетевого трафика;

– фильтрацию, т.е. выбор кокой-либо части трафика – вплоть до конкретного сайта или трафика с конкретной машины в течение какого-либо указанного времени;

– предоставление графиков активности сетевых соединений на основе выбранных фильтров;

– сбор статистики (от часа до года) с функцией экспорта;

– просмотр статистики удаленных компьютеров;

– оповещение и уведомление при определённом событии;

– возможность запуска как сервиса.

Перечисленные возможности могут присутствовать в программах мониторинга не обязательно в полном объёме.

Анализ трафика, прошедшего через сниффер, позволяет:

– обнаружить паразитный, вирусный и закольцованный трафик, наличие которого увеличивает загрузку сетевого оборудования и каналов связи;

– перехватить любой незашифрованный (иногда и зашифрованный) пользовательский трафик с целью получения паролей и другой информации;

– локализовать неисправность сети или ошибку конфигурации сетевых агентов.

Методы статистического анализа применительно к трафику сети основаны на различных динамических характеристиках, но имеют одинаковые базовые принципы. Особенностью нахождения аномалий являются серии наблюдений, а не конкретное значение. В пределах такой серии изменения ищутся в момент времени, в который статистические свойства наблюдаемой величины резко изменяются. Статистические данные могут сохраняться как в базе данных, так и в специальной структуре изменений – профайле.

Для анализа поведения трафика в первую очередь нужно определить множество варьируемых параметров сети, каким либо образом влияющих на её работу. Зафиксированные значения этих параметров должны быть сгруппированы в подмножества, после чего их можно использовать для анализа.

Предполагается, что показателем аномалии в поведении трафика является существенное изменение некоторых его характеристик. Причем показатели, выбранные для анализа трафика, должны быть достаточно чувствительны к его изменениям и неисправностям, которые вызваны законным и безвредным трафиком, иначе не исключены ложные тревоги [3].

Современная сетевая инфраструктура настолько велика, что отследить правильность (безаномальность) движения всей информации в ней практически невозможно. Сетевой трафик представляет собой сложный динамический процесс и является суперпозицией многих потоков с множественными взаимосвязями, которые генерируются различными потоками.

При исследованиях компьютерных систем и сетей применяют как аналитические (основанные на математических расчётах) модели, так и имитационное моделирование на основе уже готовых программ эмуляторов созданных с помощью универсальных языков программирования. Системы моделирования могут быть как узконаправленными (специализированными), так и универсальными, позволяющими имитировать сети самых различных типов [5].

Результат исследований экспериментальных данных показал, что трафик современных компьютерных сетей обладает особой структурой, которая проявляет эффект самоподобия. Этот эффект проявляется в том, что статистические характеристики трафика как бы «масштабируются» при усреднении значений взятых за разные промежутки времени. Другими словами, под самоподобием подразумевается повторяемость распределения нагрузки во времени при различных масштабах (рис. 2) [4].

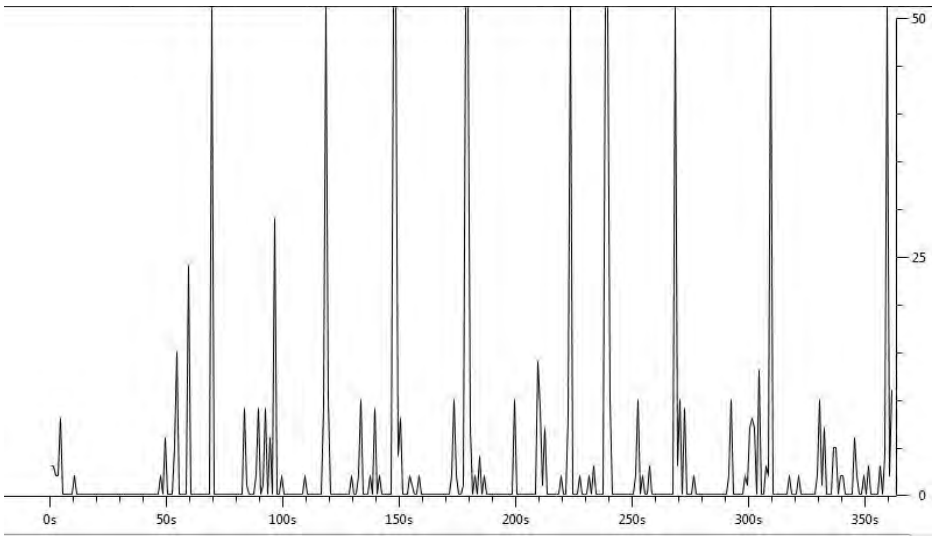

Рис. 2. Нормальная нагрузка трафика при беспроводном соединении

Появление эффекта самоподобия связано со свойствами TCP/IP протокола передачи данных, который на сегодняшний день используется в компьютерных сетях. Имеется в виду особенность передачи данных в виде пакетов, которые приходят на узел коммутации целыми пачками, а не случайным образом.

Трафик в таких сетях имеет явно выраженный всплесковый характер, что повышает вероятность перегрузок в узлах сети, которые ведут к переполнению буферов и вызывают потери и /или задержки. Пульсации приводят к перепадам скорости информационных потоков, при которых отношение максимального значения скорости к минимальному составляет десятки раз.

Исследования самоподобия показали, что это явление значительно ухудшает качество трафика через сеть [4].

Однако, несмотря на то, что проявление самоподобия оказывает не очень положительное влияние на передачу данных в сети (оно есть и с ним нет способов бороться), из этого явления можно извлечь выгоду.

Свойство самоподобия ассоциируется с определением фрактала, то есть, при изменении шкалы корреляционная структура самоподобного процесса остается неизменной.

Самоподобные (фрактальные) модели более точно характеризуют поведение нагруженного сетевого потока, чем пуассоновские модели, важной задачей стала разработка инструментальных средств для понимания самоподобных процессов, и для синтеза искусственного сетевого трафика, который отражает основные характеристики этих процессов.

Фрактальный анализ трафика дает более близкие к правде результаты, чем применение классических методов, основанных на методах Пуассона, которые дают неоправданно оптимистические результаты, приводящие к недооценке нагрузки, а, следовательно, к сбоям в работе сети. На данный момент фрактальность трафика широко используется для прогнозирования поведения компьютерной сети на этапе её проектирования, а так же для выявления аномалий поведения трафика в существующей сети.

Центральными понятиями, которые используют для фрактального анализа, являются фрактальная размерность (D) и показатель Херста (H).

Фрактальная размерность множества (по Хаусдорфу) определяется:

$$D = -\lim_{\varepsilon \to 0} \frac{\lg[N(\varepsilon)]}{\lg[\varepsilon]},$$

где $N(\varepsilon)$ – число непустых кубов размером ε, покрывающих заданное множество.

Показатель Херста характеризует степень самоподобия процесса:

1) $0 < H < 0.5$ – случайный процесс, который не обладает самоподобием, характеризуется стремлением к среднему значению;

2) $H = 0.5$ – полностью случайный процесс без выраженной тенденции;

3) $H > 0.5$ – трендоустойчивый процесс, который обладает длительной памятью и является самоподобным.

Фрактальная размерность напрямую связана с показателем Херста соотношением: $D = 2 – H$.

Это соотношение справедливо, когда структура кривой, описывающей фрактальную функцию, исследуется с высоким разрешением, то есть в локальном пределе [5].

Алгоритм оценки безопасности трафика можно разделить на пять этапов:

1) сбор трафика;

2) статистический анализ;

3) оценка показателя Херста;

4) обнаружение аномалий;

5) оценка безопасности.

Для уменьшения влияния на нормальное функционирование сети трафик дублируются на специальный сервер, занимающийся сбором трафика. Из пакетов, принятых от маршрутизатора, извлекается информация о типе пакета, а также четыре метрики трафика: общее число пакетов, число TCP пакетов, UDP пакетов, ARP пакетов в единицу времени. Вычисляются показатели Херста для четырех метрик трафика итеративным методом оценки в режиме реального времени.

Эти значения используются для обнаружения аномалий и обновления модели нормального трафика.

Текущее вычисленное значение показателя Херста сравнивается со значением из нормальной модели поведения трафика.

Если значение выходит за пределы допустимого, трафик считается аномальным. Нормальная модель трафика строится путем анализа нормальной работы сети в течение определенного промежутка времени.

Модель включает нормальное значение показателя Херста и доверительный интервал, и может быть обновлена при обнаружении аномалий. Критерием оценки безопасности является уровень риска, вычисляемый методом средневзвешенных величин, который учитывает результаты обнаружения аномалий от четырех метрик трафика. Уровень риска предоставляет администраторам текущее состояние передачи данных в сети с точки зрения безопасности [5].

Для синтеза и анализа алгоритмов обнаружения атак часто используют базу данных KDD-99, кото-

рая содержит около 5 миллионов записей о сетевых соединениях.

При этом представленные в этой базе 22 вида атак делятся на 4 основные категории:

DoS – отказ в обслуживании, генерация большого объёма трафика, что приводит к перегрузке и блокированию сервера;

U2R – предполагает получение зарегистрированным пользователем привилегий локального суперпользователя (администратора);

U2L – характеризуется получением доступа незарегистрированного пользователя к компьютеру со стороны удалённой машины;

Probe – заключается в сканировании портов с целью получения конфиденциальной информации.

Атаки, попадающие под описание перечисленных категорий, в обязательном порядке будут вызывать аномальные изменения в работе сети [1].

Выводы

Таким образом, анализ методов выявления аномалий и вторжений показал, что для надежной передачи данных в современных компьютерных сетях анализ трафика жизненно необходим, так как посредством него существует возможность производить прогнозирование поведения процессов сети.

Прогнозирование поведения трафика применимо как при проектировании, так и во время обеспечения безопасности.

Причем в таком прогнозировании не последнюю роль играет явление фрактальности трафика (несмотря на отрицательное влияние эффекта самоподобия на нагрузку сети), которое обусловлено свойствами используемого протокола, а также изменчивостью характеристик при появлении аномалий в сети, что позволяет использовать методы фрактального анализа для обнаружения атак.

Основным требованием к любым методам обнаружения вторжений является возможность обнаружения произвольных типов аномалий, в том числе, распределенных во времени.

Список литературы

1. Басараб М.А. Обнаружение аномалий в информационных процессах на основе мультифрактального анализа / М.А. Басараб, И.С. Строганов // Вопросы кибербезопасности. – 2014. – №4(7). – С. 30-40 [Электронный ресурс]. – Режим доступа к ресурсу: http://cyberrus.com/wp-content/uploads/2015/01/vkb_05_04.pdf.

2. Левоневский Д.К. Разработка системы обнаружения аномалий сетевого трафика / Д.К. Левоневский, Р.Р. Фаткиева // Научный вестник НГТУ. – 2014. – № 3, том 56. – С. 108-114.

3. «Фрактальная катастрофа» TCP/IP [Электронный ресурс]. – Режим доступа к ресурсу: http://itc.ua/articles/_fraktalnaya_katastrofa_tcp_ip_5571/.

4. Шелухин О.И. Обнаружение вторжений в компьютерные сети (сетевые аномалии): учебное пособие для вузов / О.И. Шелухин, Д.Ж. Сакалема, А.С. Филинова. – М.: Горячая линия-Телеком, 2013. – 220 с.

5. Shaabany A. Network traffi c deviation detection based on fractal dimension / A. Shaabany, F. Jamshidi // Journal of Computing and Information Technology - CIT 20. – 2012. – 1. – P. 27-32 [Электронный ресурс]. – Режим доступа к ресурсу: http://cit.srce.unizg.hr/index.php/CIT/ article/view/2007/1522.

Поступила в редколлегию 13.11.2015

Рецензент: д-р техн. наук, с.н.с. С.Г. Семенов, Национальный технический университет «ХПИ», Харьков.

АНАЛІЗ СУЧАСНИХ МЕТОДІВ ВИЯВЛЕННЯ ВТОРГНЕНЬ В КОМП'ЮТЕРНІ СИСТЕМИ

Т.Н. Шипова, В.В. Босько, І.А. Березюк, Ю.М. Пархоменко

У роботі представлені результати аналізу методів виявлення вторгнень в комп'ютерні системи, виявлені їх основні достоїнства і недоліки. Запропоновані шляхи підвищення ефективності евристичного аналізу комп'ютерних систем з використанням властивості самоподобия в характеристиках інформаційного трафіку. Визначені пріоритетні шляхи обліку властивості самоподобия трафіку при аналізі стану комп'ютерної системи в умовах можливих зовнішніх дій зловмисників.

Ключові слова: *аналіз трафіку, самоподобие, аномалії, фрактальна розмірність, показник Херста.*

ANALYSIS OF MODERN METHODS OF FINDING OUT INTRUDING IN COMPUTER SYSTEMS

T.N. Shipova, V.V. Bos'ko, I.A. Berezyuk, Yu.M. Parkhomenko

The results of analysis of methods of finding out intruding are in-process presented in the computer systems, their basic dignities and failings are exposed. The ways of increase of efficiency of heuristic analysis of the computer systems are offered with the use of property of самоподобия in descriptions of informative traffic. The priority ways of account of property of самоподобия of traffic are certain at the analysis of the state of the computer system in the conditions of possible external influences of malefactors.

Keywords: *analysis of traffic, self-similarity, anomalies, fractal dimension, index of Kherst.*