**Введение**

Один из подходов к выявлению сетевых атак - анализ сетевого трафика. В этом случае анализируются только заголовки сетевых пакетов, в отличие от анализа контента, при котором анализируется содержимое пакетов. Конечно, наиболее полный контроль информационных взаимодействий может быть обеспечен только анализом всего содержимого сетевых пакетов, включая их заголовки и области данных. Однако с практической точки зрения эта задача является трудновыполнимой из-за огромного объема данных, которые приходилось бы анализировать. Современные СОВ начинают испытывать серьезные проблемы уже в сетях с производительностью 100 Мбит/с. Поэтому в большинстве случаев целесообразно использовать для выявления атак методы анализа сетевого трафика, в некоторых случаях сочетая их с анализом контента.

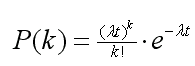
Сигнатура сетевой атаки концептуально практически не отличается от сигнатуры вируса. Она представляет собой набор признаков, позволяющих отличить сетевую атаку от других видов сетевого трафика.

**Теоретическая часть**

**Способы описания трафика рядами экспериментальных данных**

Пуассоновский поток данных.

Классической моделью трафика в информационных сетях является Пуассоновский (простейший) поток. Он характеризуется набором вероятностей P(k) поступления k сообщений за временной интервал t:



где k=0,1,… - число сообщений; λ - интенсивность потока.

Заметим, что интервал времени измерения количества сообщений t и интенсивность потока λ являются постоянными величинами.

Семейство Пуассоновских рапределений P(k) в зависимости от λ изображено на рис.1. Большее значение λ соответствует более широкому и симметричному графику плотности вероятности.

Математическое ожидание (среднее) и дисперсия Пуассоновского потока равны λt.

Зная вероятность поступления данных за период, можно получить распределение интервала τ между соседними событиями:

https://nag.ru/goodies/articles/sst/003.gif

Отсюда вывод: пуассоновский поток характеризуется экспоненциальным распределением интервалов между событиями.

Основным свойством пуассоновского потока, обусловливающим его широкое применение при моделировании, является аддитивность: результирующий поток суммы пуассоновских потоков тоже является пуассоновским с суммарной интенсивностью:

https://nag.ru/goodies/articles/sst/004.gif

При моделировании Пуассоновский поток можно получить мультиплексированием совокупности ON/OFF источников, которые называются Марковскими процессами.

Тяжелохвостое распределение.

Считается, что случайная величина имеет распределение с тяжелым (весомым) хвостом (РТХ или Heavy Tailed), если:

https://nag.ru/goodies/articles/sst/006.gif

То есть хвост распределения затухает по степенному закону, в отличие от, например, Гауссовского распределения с экспоненциальным убыванием хвоста.

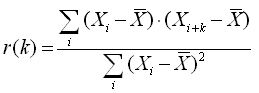
Считается, что сетевой трафик во многих случаях лучше всего описывается именно тяжелохвостым распределением, например распределением Парето.

Автокорреляционная функция (АКФ).

Автокорреляция – корреляционная связь между значениями одного и того же случайного процесса в разнесенные моменты времени. Автокорреляционная функция (АКФ) характеризует эту связь.

В общем случае АКФ характеризует внутреннюю зависимость между временным рядом и тем же рядом, но сдвинутым на некоторый промежуток (сдвиг) времени, который называется лагом.

Вычисления АКФ проводятся по классической формуле:



где X с чертой - выборочное среднее, k=0,1,2, … - лаг.

Медленно убывающая зависимость (МУЗ).

Считается, что процесс обладает медленно убывающей зависимостью если он характеризуется АКФ, которая убывает гиперболически при увеличении лага. В противоположность МУЗ существует понятие быстро убывающей зависимости.

С понятием МУЗ связано важнейшее прогнозирующее свойство – продолжительная память, характеризующаяся зависимостью текущих параметров процесса от предыдущих.

Самоподобие трафика.

Многочисленные современные исследования Интернет трафика свидетельствуют о том, что он обладает свойством самоподобия .

Простейшим самоподобными объектами являются фракталы. Согласно определению Мандельброта: "фрактал – структура, состоящая из частей, которые в каком-то смысле подобны целому" . Поэтому самоподобные процессы часто называют фрактальными.

Неформально самоподобный процесс определяетcя как случайный процесс, статистические характеристики которого проявляют свойства масштабирования.

Строго самоподобный в широком смысле процесс (ССШП) характеризуется инвариантностью АКФ при изменении уровня агрегирования при условии МУЗ.

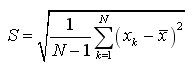
В отличие от пуассоновских процессов самоподобные характеризуются наличием последействия: вероятность поступления следующего (очередного) события зависит не только от времени, но и от предыдущих событий (предыстории). Это означает, что число текущих событий может зависеть от числа предыдущих событий в отдаленные промежутки времени. Поэтому одним из основных свойств самоподобного процесса (self similar) является МУЗ (long range dependency).

Важнейшим параметром, характеризующим степень самоподобия, является параметр Херста (Hurst) H, определяемый для временного ряда Xk,k=**1,2,K** N из соотношения

https://nag.ru/goodies/articles/sst/011.gif

где

R=**max(Xk) - min(Xk)** - размах отклонения;

 - стандартное отклонение,

N - число членов временного ряда, а - константа.

Используя значение показателя Херста H, выделяют три типа случайных процессов:

1. 0<=H<=0,5 - случайным процесс является антиперсистентным, или эргодическим, рядом, который не обладает самоподобием;
2. H = 0,5 - полностью случайный ряд, аналогичный случайным смещениям частицы при классическом броуновском движении;
3. H > 0,5 - персистентный (самоподдерживающийся) процесс, который обладает длительной памятью и является самоподобным.

Дополнительно следует отметить, что самоподобный процесс часто носит взрывной (burst) характер, что выражается в возможности наличия выбросов во время относительно низкой скорости поступления событий.

Применительно к трафику самоподобие выражается в неизменности поведения при изменении временных масштабов наблюдения и сохранения склонности к всплескам при усреднении по шкале времени.

**Обзор программ-анализаторов (снифферов) сетевого трафика**

Анализатор трафика, или сниффер, — сетевой анализатор трафика, программа или программно-аппаратное устройство, предназначенное для перехвата и последующего анализа либо только анализа сетевого трафика, предназначенного для других узлов.

Анализ прошедшего через сниффер трафика позволяет:

• обнаружить паразитный, вирусный и закольцованный трафик, наличие которого увеличивает загрузку сетевого оборудования и каналов связи (снифферы здесь малоэффективны; как правило, для этих целей используют сбор разнообразной статистики серверами и активным сетевым оборудованием и её последующий анализ);

• перехватить любой незашифрованный (а порой и зашифрованный) пользовательский трафик с целью получения паролей и другой информации;

• локализовать неисправность сети или ошибку конфигурации сетевых агентов (для этой цели снифферы часто применяются системными администраторами).

Поскольку в «классическом» сниффере анализ трафика происходит вручную, с применением лишь простейших средств автоматизации (анализ протоколов, восстановление TCP-потока), то он подходит для анализа лишь небольших его объёмов.

**Виды атак**

*Сетевая атака Portsweep* — наблюдатель изучает множество

портов для определения, какие сервисы поддерживаются на единичном хосте.

*Сетевая атака Neptune*  — пример DoS атаки, которая выполняет рассылку флуда (SYN пакетами) на один или множественные порты. Нормальное установление TCP-соединения происходит в соответствии со схемой ≪трехстороннего рукопожатия≫:

клиент отправляет SYN, сервер отвечает SYN-ACK, клиент отвечает ACK. Если со стороны клиента отправлять большое количество SYN-пакетов, то сервер будет вынужден каждый раз инициировать соединение, которое не будет установлено. Количество таких соединений ограничено и когда оно достигнет предела, сервер перестанет отвечать на какие-либо запросы.

*Сетевая атака Nmap*  — составление карты сети сиспользованием утилиты nmap. Способы исследования сети могут меняться — возможно использование SYN пакетов.

Особенностью данной атаки является продолжительность — более 12 ч . В связи с наличием фонового трафика, который скрывает саму атаку, масштаб был взят наиболее большим, однако при данной длительности атаки появляется проблема хранения и обработки большого объема данных (шаг 1 с).

*Mailbomb* — почтовые бомбы ---один из простейших видов сетевых атак. Злоумышленником посылается на компьютерпользователя или почтовый сервер компании одно огромное сообщение или множество (десятки тысяч) почтовых сообщений на SMTP- порт mail-сервера жертвы, что приводит к выводу системы из строя, если не установлено специальных блокировок.

*Stealth scan* --- Скрытое сканирование реализуется отсылкой пакета FIN на порт. В соответствии со стандартом RFC 793[28 Postel, Transmission Control Protocol, RFC 793, Sep. 1981] корректный ответ закрытого порта будет ответный пакет RST, тогда как открытый порт просто отбросит пакет и ничего не пошлет в ответ. Такое сканирование заметить сложнее.

*Ipsweep* --- Существует два типа сканирования портов: вертикальное, когда один хост сканируется по всем открытым портам, и горизонтальное, когда группа хостов сканируется на открытый определенный порт. В большинстве случаев на порт посылается SYN-пакет, и если порт открыт, он ответит SYN-ACK.

**Методы обнаружения аномилий**

Методы обнаружения аномалий направлены на выявление неизвестных атак и вторжений. Для защищаемой системы СОВ на основе совокупности параметров оценки формируется ≪образ≫ нормального функционирования. В современных СОВ выделяют несколько способов построения ≪образа≫:

*1)*накопление наиболее характерной статистической информации для каждого параметра оценки;

*2)*обучение нейронных сетей значениями параметров оценки;

*3)*событийное представление.

Легко заметить, что в обнаружении очень значительную роль играет множество параметров оценки. Поэтому в обнаружении аномалий одной из главных задач является выбор оптимального множества параметров оценки.

Другой, не менее важной задачей является определение общего показателя аномальности. Сложность заключается в том, что эта величина должна характеризовать общее состояние «аномальности» в защищаемой системе.

Основной недостаток сигнатурных методов обнаружения сетевых атак, связанный с неспособностью системы обнаруживать атаки неизвестных типов, может быть устранен применением методов, основанных на выявления аномалий сетевой активности. Такие методы основаны на предположении, что для вычислительной системы существует свой профиль нормального состояния и любые значительные отклонения от него являются вероятным кандидатом на возможную атаку. Основное достоинство такого метода – возможность выявления новых, неизвестных ранее видов атак. Для построения

базового профиля системы используется набор данных, свободный от аномалий, или статистические методы.

Методы обнаружения изменений могут классифицироваться в соответствии со следующими критериями:

*Онлайн — офлайн*: регистрация онлайн изменений повторяет тесты с поступающими новыми наблюдениями. Цель состоит в том, чтобы обнаружить изменения с низкой задержкой и длительным временем между ложными тревогами.

Офлайн регистрация изменений анализирует серию наблюдений фиксированной длины. Цель состоит в том, чтобы определить, существует ли изменение в любой момент времени. Если да, то время или величина изменения может быть оценена.

*Байесовский — небайесовский*: байесовская регистрация из менений использует априорную информацию о распределении времени изменения для того, чтобы улучшить качество обнаружения. Небайесовская регистрация изменений не рассматривает такую информацию, т. е. вероятность изменения предполагается независимой от времени.

*Параметрический — непараметрический*: в случае парамет-

рического метода вероятность распределения наблюдаемых переменных должна следовать параметрической модели. Кроме того, изменение предполагает действие с параметром, а не с самой параметрической моделью. С другой стороны, непараметрическая регистрация изменений не делает никаких предположений о распределении наблюдаемой величины до и после изменения. Следовательно, вместо рассмотрения параметров модели непараметрический метод рассматривает статистические свойства наб- людений, такие, как среднее, дисперсия, корреляция и т. д., которые предположительно будут затронуты изменением.

*Известное — неизвестное изменение*: если модель системы и параметры после изменения априорно известны, процедура испытаний может базироваться на максимально вероятном решении, которое определяет образцы вероятностей для *H*0 и *H*1 и выбирает

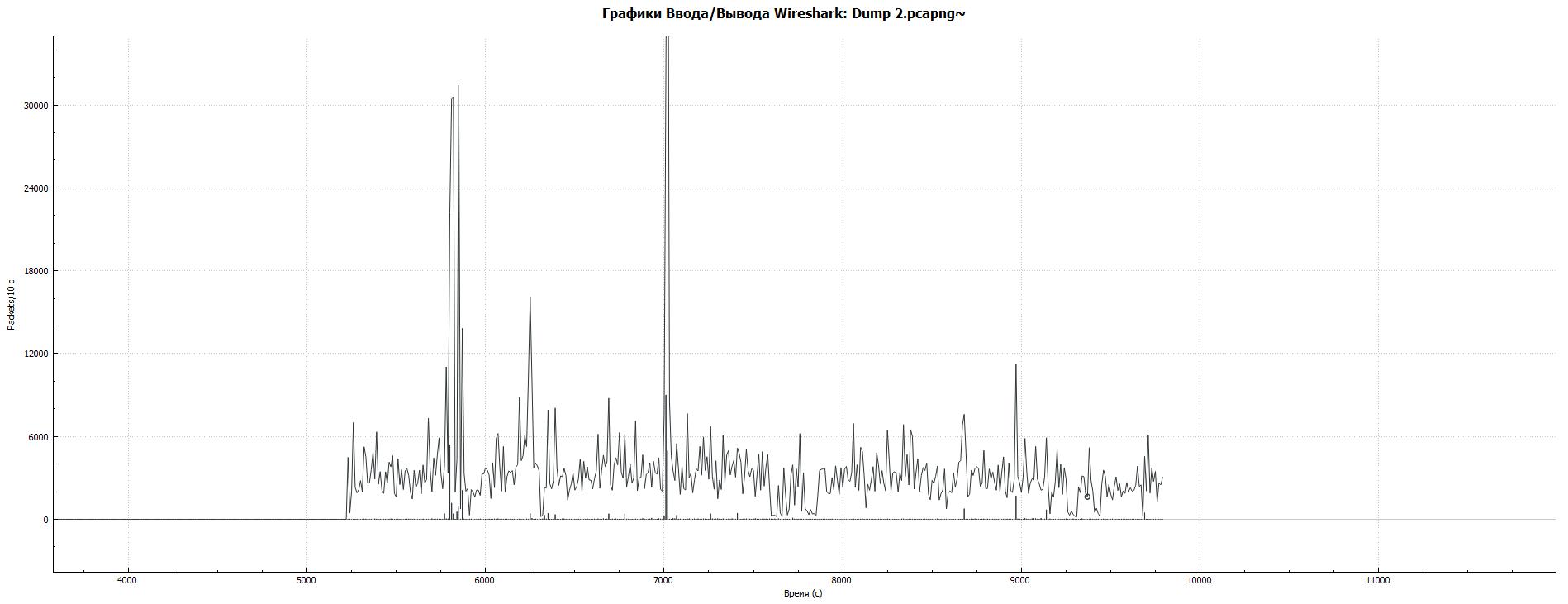
гипотезы, которые наиболее вероятны. Если величина изменения априорно не известна, *H*0 может быть только отклонена по отношению к данному уровню значимости.

*Одномерный — многомерный*: одномерная регистрация изменений принимает решение по одной переменной. Многомерная регистрация изменений рассматривает несколько переменных и корреляции между ними.

В общем случае большинство априорных знаний доступно, что облегчает регистрацию изменений с высокой точностью. К примеру, параметрический метод более сильный, чем непараметрический, что позволяет обнаруживать настоящие аномалии среди некоторого

уровня ложных тревог (т. е. вероятность тревоги в отсутствие каких-либо значительных изменений). Однако, если предложенная модель некорректна, параметрический метод теряет свою силу и может привести к неправильным решениям. В случае обнаружения аномального трафика мы обычно не знаем конкретного распределения наблюдаемых переменных, таким образом, метод параметрической регистрации изменений должен быть применен осторожно, зная то, что таблицы параметризации обычно действительны для нормально распределенных переменных. В общем случае непараметрический метод регистрации изменений более подходящий. Кроме того, изменения должны быть зарегистрированы очень быстро (т. е. онлайн) без каких-либо априорных знаний об их величине, так как такая информация обычно недоступна. Кроме того, мы ограничимся небайесовскими методами и не будем считать, что изменения происходят чаще в определенные моменты времени.

**Практическая часть**

В ходе работы помимо сбора и анализа информации о методах представления и сбора данных трафика сети, поиске аномалий а так же таповых атаках и формах их проявления так же мной был получен дамп трафика с домашней сети. Для сбора дампа трафика была использована программа – анализатор сетевого трафика Whireshark. С ее помощью мной были сняты несколько дампов трафика с домашней сети, с целью обучения в определении базового параметра трафика и тренировке в нахождении аномалий с его помощью. 

После применения программы - анализатора был получен текстовый файл с массивом данных о числе пакетов принятых за единицу отсчета. Среди них находится первая минимальная последовательность отсчетов с различным числом пакетов, размер которой является уникальной для всего файла (больше не встречается в таком порядке нигде в файле). Эта процедура повторяется до тех пор, пока все данные файла не будут преобразованы к такому виду. Затем, уже исходя из преобразованного файла можно проводить дальнейшее исследование, к примеру определять критическое изменение .

На момент сдачи отчета в разработке находится программа, которая будет автоматически анализировать полученный от программы анализатора файл и самостоятельно разбивать его на подобные ячейки.