ОИПИ НАН БЕЛАРУСИ

Государственное научное учреждение
«Объединенный институт проблем информатики
Национальной академии наук Беларуси»

ОБНАРУЖЕНИЕ АНОМАЛИЙ СЕТЕВОГО ТРАФИКА В ТЕЛЕКОММУНИКАЦИОННЫХ СЕТЯХ С ПОМОЩЬЮ СТРУКТУРНО-СПЕКТРАЛЬНОГО МЕТОДА

Дмитриев В.А.
Касанин С.Н.
Максимович Е.П.

ОИПИ

При решении задач, связанных с диагностикой и защитой сетевых ресурсов, центральным вопросом является оперативное обнаружение состояния телекоммуникационной сети, приводящее к потере полной или частичной ее работоспособности, уничтожению, искажению или утечке информации, являющихся следствием отказов, сбоев случайного характера или результатом получения злоумышленником несанкционированного доступа к сетевым ресурсам. Раннее обнаружение таких состояний позволит своевременно устранить их причину, а также предотвратит возможные катастрофические последствия.

Воздействие кибератак приводит к появлению в сетевом трафике телекоммуникационной сети аномальной активности. Обнаружение и классификация аномалий предполагает непрерывный процесс в телекоммуникационной сети.

Для постоянного мониторинга сетевого трафика и обнаружения в нем аномальной активности, идентификации и классификации атак, а также выявления в нем ложных изменений необходимо учитывать наличие большого количества параметров, характеризующих проявление новых свойств трафика. Все это является побудительным мотивом для поиска новых методов обнаружения аномалий сетевого трафика телекоммуникационной сети.

Одним из наиболее перспективных направлений исследований по определению признаков наличия или отсутствия аномалий является спектральный анализ входных и выходных сетевых трафиков телекоммуникационной сети, которые рассматриваются как отрезки временных рядов. Главные компоненты временного ряда выбираются таким образом, чтобы они соответствовали наибольшей изменчивости исходного сетевого трафика. Остальные компоненты рассматриваются как составляющие шума.

Анализ спектральной плотности мощности позволяет не только определить наличие несанкционированной активности (аномалий сетевого трафика), но и сделать предположение о виде атаки.

В данной работе используется аппроксимационный подход для обнаружения аномалий сетевого трафика телекоммуникационной сети.

Аппроксимационный подход обобщает принципы, методы и технологии получения информации о явлениях экспериментальным путем по результатам наблюдений и контроля с использованием аналитических моделей функциональных характеристик этих явлений, выбранных на основе априорной информации с учетом целей проводимых исследований.

Математическая модель при этом выступает как системная интерпретационная конструкция, объединяющая цели и условия измерений, контроля и испытаний. Иными словами аппроксимационный подход – это подход, использующийся для поиска приближённого решения оптимизационной задачи.

Реальный сетевой трафик описывается нестационарным временным рядом.

В случае, когда $X_\alpha(t)$ представляет собой нестационарную случайную функцию, т.е. когда $M[X_\alpha(t)]$ меняется с течением времени, вместо $X_\alpha(t)$ рассматривают разность $X_\alpha(t) = X_\alpha(t+\tau) - X_\alpha(t)$. При не слишком большом τ медленные изменения функции $X_\alpha(t)$ будут мало сказываться на значениях $X_\alpha(t)$, и тем меньше, чем они медленнее. В результате подавления компонент с очень большими периодами приращение $X_\alpha(t)$ станет стационарным.

Для подтверждения того факта, что приращение нестационарного сетевого трафика телекоммуникационной сети является стационарным, используется тест Дики-Фуллера.

К преимуществам структурно-спектрального метода обнаружения аномалий сетевого трафика относится высокая точность и малая вероятность ложного срабатывания.

При анализе нестационарных случайных процессов со стационарными приращениями используется структурная функция, которая играет ту же роль, что и корреляционная функция при анализе стационарных случайных процессов.

Структурная функция определяется следующим образом:

$$S_\alpha(\tau) = M\{[X_\alpha(t+\tau) - X_\alpha(t)]^2\} \quad (1)$$

Спектральное разложение структурной функции имеет следующий вид:

$$S_\alpha(\tau) = 2 \cdot \int_{-\infty}^{+\infty} [1 - cos(\omega\tau)] dG_\alpha(\omega). \quad (2)$$

Спектральная плотность мощности определяется следующим образом:

$$1)\ G_\alpha(\omega) = \frac{1}{2\pi\omega} \cdot \int_0^{+\infty} sin(\omega\tau) \cdot \frac{\partial S_\alpha(\tau)}{\partial \tau} \mathrm{d}\tau, \quad (3)$$

если $\lim\limits_{\tau \to +\infty} \frac{\partial S_\alpha(\tau)}{\partial \tau} = 0$, $\lim\limits_{\tau \to +\infty} \tau^2 \cdot \frac{\partial S_\alpha(\tau)}{\partial \tau} = 0$.

$$2)\ G_\alpha(\omega) = \frac{1}{2\pi\omega} \cdot \int_0^{+\infty} cos(\omega\tau) \cdot \frac{\partial^2 S_\alpha(\tau)}{\partial \tau^2} \mathrm{d}\tau, \quad (4)$$

если $\lim\limits_{\tau \to +\infty} \frac{\partial^2 S_\alpha(\tau)}{\partial \tau^2} = 0$, $\lim\limits_{\tau \to +\infty} \tau \cdot \frac{\partial^2 S_\alpha(\tau)}{\partial \tau^2} = 0$

Структурную функцию можно также разложить в ряд по базису ортогональных функций:

$$S_\alpha(\tau) = \sum_{k=0}^{m} \beta_k \cdot \psi_k(\tau, \alpha), \quad (4)$$

где $\boldsymbol{m}$ – число членов разложения,

$\boldsymbol{\beta_k}$ – коэффициенты разложения,

$\boldsymbol{\psi_k(\tau, \alpha)}$ – семейство ортогональных функций ($\boldsymbol{\alpha}$ – параметр масштаба ортогональных функций).

Для модели структурной функции (4) коэффициенты разложения, обеспечивающие минимум квадратической погрешности аппроксимации:

$$\theta = \int_0^{+\infty} S_x^2(\tau) \ - \sum_{k=0}^{m} \beta_k^2 \cdot \|\psi_k(\tau, \alpha)\|^2 = min, \qquad (5)$$

где $\|\psi_k(\tau, \alpha)\|^2 = \int_0^{+\infty} \psi_k^2(\tau, \alpha) d\tau$ – норма ортогональной функции определяются следующим образом:

$$\beta_k = \frac{1}{\|\psi_k(\tau,\alpha)\|^2} \cdot \int_0^{+\infty} S_\alpha(\tau) \cdot \psi_k(\tau, \alpha) d\tau. \qquad (6)$$

Чаще используется относительная погрешность аппроксимации, которая определяется следующим образом:

$$\delta = \frac{\theta}{\int_0^{+\infty} S_\alpha^2(\tau) d\tau}. \qquad (7)$$

Ортогональный базис	δ_{min}
Лагерра	0,06852
Лежандра	0,09338
Якоби	0,09356
Сонина-Лагерра	0,3514

Полином Лагерра определяется следующим образом:

$$L_k(\tau,\alpha) = \sum_{s=0}^{k}(-1)^s \cdot C_k^s \cdot \frac{(\alpha\cdot\tau)^s}{s!} \cdot exp\left(-\frac{\alpha\cdot\tau}{2}\right), \quad (8)$$

$$C_k^s = \frac{k!}{s!\cdot(k-s)!}.$$

Структурная функция определяется следующим образом:

$$S_\alpha(\tau) = \sum_{k=0}^{m} \beta_k \cdot L_k(\tau, \alpha), \quad (9)$$

$$\|\psi_k(\tau, \alpha)\|^2 = \frac{1}{\alpha}, \quad (10)$$

$$\alpha = \frac{0{,}8}{\tau},$$

где τ – промежуток времени между соседними отсчетами.

Используя (9) для спектральной плотности мощности структурной функции получим следующее выражение:

$$G_\alpha(\omega) = \frac{cos\varphi}{\pi\alpha} \cdot \sum_{k=0}^{m} (-1)^k \cdot \beta_k \cdot cos[(2k+1)\varphi], \quad (11)$$

$$\varphi = arctg\left(\frac{2\omega}{\alpha}\right). \quad (12)$$

Коэффициент аномальности сетевого трафика телекоммуникационной сети определяется следующим образом:

$$K = \frac{G_\alpha(\omega)}{\sum_{k=0}^{m} \beta_k \cdot \psi_k(0,\alpha)} = \frac{G_\alpha(\omega)}{\sum_{k=0}^{m} \beta_k}. \quad (13)$$

Уровень аномальности	**Интервалы для K**
Низкий	$0 \leq K < 0{,}7$
Высокий	$0{,}7 \leq K < 1{,}0$

Аномалии присутствуют, если $\boldsymbol{K \geq 0{,}7}$.


Внешний канал передачи данных
Клиентская часть/
Мультифункциональные распределенные агенты
Модуль захвата данных
Модуль предобработки и формирования данных
Модуль шифрования
Модуль формирования и передачи пакета
Модуль логирования
Модуль ответа на вторжения
Локальная БД
Локальный канал передачи данных
Административная часть/Консоль администратора
Модуль управления клиентской частью
Модуль управления серверной частью
Модуль загрузки и обновления данных в БД
Доверенный канал передачи данных
Серверная часть/распределенная серверная часть
Модуль дешифрации и предобработки
Модуль формирования очереди
Модуль логирования
Модуль анализа/ансамбль классификаторов
Сигнатурные методы
Статистические методы
Методы вейвлет-анализа
Спектральные методы
Нейронные сети
Иммунные системы
Модуль ответа
Модуль хранения состояния
БД

СПАСИБО ЗА ВНИМАНИЕ !