КОМПЬЮТЕРНЫЕ НАУКИ 2018. Т. 28. Вып. 3

УДК 517.28, 530.181

© *А. Е. Краснов, Е. Н. Надеждин, Д. Н. Никольский, Д. С. Репин, В. С. Галяев*

ДЕТЕКТИРОВАНИЕ DDOS АТАК НА ОСНОВЕ АНАЛИЗА ДИНАМИКИ И ВЗАИМОСВЯЗИ ХАРАКТЕРИСТИК СЕТЕВОГО ТРАФИКА

В работе усовершенствован подход к обнаружению DDoS-атак на основе использования оператора эволюции динамических систем, разработанный ранее авторами. В предложенном подходе сетевому трафику ставятся в соответствие различные характеристики — признаки его временной структуры, формируемые по адресным и нагрузочным параметрам заголовков пакетов данных трафика. Предполагается, что различным состояниям трафика (нормальное состояние, атаки разных типов) соответствуют различные временные структуры характеристик, которые генерируется неизвестными линейными динамическими операторами. Связь между значениями различных характеристик в различных дискретных временных отсчетах устанавливается оператором эволюции. Основная рабочая гипотеза исследования заключается в том, что различным состояниям трафика соответствуют различные динамические операторы, а следовательно, и операторы эволюции. Приведен общий вид матрицы оператора эволюции трафика, реконструированной по значениям его наблюдаемых характеристик. Матричные элементы оператора эволюции определяют взаимосвязь характеристик трафика, давая целостное описание его динамической структуры. Введено понятие среднего значения оператора эволюции трафика, на основе которого формируются специальные хеш-функции и их статистические распределения для различных состояний трафика. В вычислительном эксперименте формировались адресные и нагрузочные хеш-функции, причинно соответствующие адресным и нагрузочным параметрам заголовков пакетов данных трафика. Результаты вычислительного эксперимента подтвердили возможность точной классификации трех состояний трафика: нормального и двух аномальных (HTTP flood атака и SlowLoris атака).

Ключевые слова: сетевой трафик, DDoS-атака, обнаружение, динамический оператор, оператор эволюции, хеш-функция, классификация.

DOI: 10.20537/vm180310

Введение

Проблема разработки мер противодействия DDoS-атакам описана во многих источниках, например, в работах [1–3] подробно рассмотрены различные механизмы DDoS-атак и применяемые контрмеры, проведен тщательный анализ сильных и слабых сторон различных предлагаемых механизмов защиты. Для обнаружения DDoS-атак предлагается ряд методов, включая методы выявления и описания динамики и возможных внутренних отношений между параметрами трафика — корреляционный анализ, спектральный анализ, а также вейвлет-анализ. Хотя все перечисленные методы имеют некоторые недостатки.

В корреляционном анализе обычно предполагается некоторая стационарность поведения трафика на отдельных последовательных интервалах наблюдения. Вычисляемые для нормального трафика корреляционные функции рассматриваются далее как шаблоны, с которыми сравниваются текущие корреляционные функции исследуемого трафика [4, 5]. Значение меры сравнения, в предположении о нормальности ее распределения, используется как информативный признак для обнаружения атаки [4]. Показано, что обнаружение аномальных вариаций временных рядов трафика может быть сделано с заданной вероятностью идентификации и заданной вероятностью ложной тревоги. Однако отмечается, что шаблоны, установленные в течение периода обнаружения, могут не соответствовать другим периодам обнаружения. Для борьбы с этим явлением необходимы дальнейшие исследования по адаптации метода.

Результаты обнаружения DDoS-атак с высокой точностью в режиме реального времени обеспечивает многомерный корреляционный анализ, при котором коррелируют несколько параметров трафика [6]. Однако авторы указали некоторые ограничения для применения данного метода. Во-первых, нет гарантии, что выбранные параметры достаточны для обнаружения

DDoS-атак различных типов. Во-вторых, не ясно как выбирать соответствующий временной интервал корреляции.

В других корреляционных методах сами функции корреляции выбранных информативных признаков используются как меры сходства потоков трафика [7]. Однако данный подход весьма чувствителен к статистическим вариациям трафика и изменению временного масштаба чередования пакетов. Более общий подход к использованию корреляционного анализа в больших данных приведен в [8]. При обнаружении DDoS-атак наряду с корреляционным анализом используют родственный ему спектральный анализ [9–12]. Основным ограничением здесь, так же как и в корреляционном анализе, является нестационарность трафика, что несовместимо с необходимостью выбора достаточно длинных временных интервалов формирования спектральных данных для повышения частотного разрешения. Для того чтобы уменьшить временные интервалы формирования спектральных описаний, используют вейвлет-анализ, который позволяет подбирать более короткие (по сравнению со спектральным анализом) временные интервалы разложения информативных признаков трафика по базисным вейвлет-функциям [13–16]. Однако данная модель также имеет определенные проблемы: она зависит от регулярных циклов трафика (дневные, недельные и месячные циклы) и выбранных базисных вейвлет-функций [14]. Ограничением является также и то, что используется лишь один параметр без учета его связи с динамикой других параметров [13].

Во многих работах отмечается, что для точного обнаружения DDoS-атак необходимо анализировать динамику практически всех параметров сетевого трафика и учитывать их взаимосвязь. Например, как показано в [17], дополнительное рассмотрение взаимосвязи параметров адреса обеспечивает более эффективное обнаружение DDoS-атак и более низкие значения ложных тревог. Для этого в работе [18] предложен подход, основанный на описании динамической структуры трафика с учетом связи всех изменяющихся параметров, получаемых из заголовков пакетов данных. Для этого было введено конечно-разностное дифференциальное уравнение, определяющее эволюцию вектора информативных признаков трафика на основе гипотетического динамического оператора, описывающего некоторые взаимодействия потоков пакетов данных. Это позволяет классифицировать трафик как нормальный или относящийся к определенному типу атаки.

Целью настоящей работы является дальнейшее развитие анализа динамической / временно́й структуры трафика на основе его оператора эволюции, построение хеш-функций информативных признаков трафика, формируемых из параметров адресной и нагрузочной частей его пакетов данных, исследование их статистических распределений. Основными задачами работы являются:

(1) описание оператора эволюции потока информативных признаков трафика;

(2) построение хеш-функций информативных признаков трафика, формируемых по адресным и нагрузочным параметрам его пакетов данных, и их использование для классификации трафика;

(3) построение и исследование статистических распределений хеш-функций, необходимых для классификации DDoS-атак типа HTTP flood и SlowLoris, путем последовательного анализа пакетов данных.

§ 1. Описание оператора эволюции трафика

Предположим, что динамика непрерывного вектора $F(t) = [f_1(t), f_2(t), \ldots, f_M(t)]^T$, состоящего из M информативных признаков трафика $f_m(t)$ $(m = 1, \ldots, M)$, описывается дифференциальным уравнением:

$$\frac{d}{dt}F(t) = \frac{1}{\triangle t}H(t)F(t), \tag{1.1}$$

где $H(t)$ — неизвестный нестационарный динамический оператор, который определяет $F(t)$, а $\triangle t$ — параметр размерной нормализации. Будем полагать, что различным состояниям трафика на конечных интервалах времени соответствуют различные виды динамического оператора

$H(t)$. Поскольку трафик состоит из набора потоков пакетов данных, то по аналогии с физическим процессом, можно считать, что гипотетический оператор $H(t)$ описывает взаимодействие этих потоков. Например, вытеснение потока легитимных пакетов потоком, порожденным DDoS-атакой.

Решением (1.1) является выражение:

$$F(t) = S(t,\tau)F(\tau), \quad S(t,t) = E, \tag{1.2}$$

где $S(t,\tau)$ — матрица Коши [19], называемая в теоретической физике оператором эволюции или оператором временного сдвига [20]. Связь оператора эволюции $S(t,\tau)$ системы (1.1) с ее динамическим оператором впервые описана Пеано [19,21], а в теоретической физике ее описывают оператором Дайсона [20,22].

С использованием хронологического упорядочивания Дайсона получим:

$$S(t,\tau) = E + \frac{1}{\triangle t}\int_\tau^t H(t_1)\,dt_1 + \frac{1}{2!}\left(\frac{1}{\triangle t}\right)^2 \int_\tau^t dt_1 \int_\tau^t dt_2\, \mathfrak{J}[H(t_1)H(t_2)] + \ldots, \tag{1.3}$$

где $\mathfrak{J}$ — оператор хронологического упорядочивания

$$\mathfrak{J}[H(t_1)H(t_2)] = \begin{cases} H(t_1)H(t_2), & \text{если } t_1 > t_2 \\ H(t_2)H(t_1), & \text{если } t_1 < t_2 \end{cases}$$

В качестве преимуществ использования оператора эволюции по сравнению с динамическим оператором можно выделить следующие:

(1) нет необходимости знать точную форму динамического оператора, поскольку с помощью наблюдений $F(t)$ можно реконструировать оператор эволюции из (1.2);

(2) в уравнении (1.2) можно использовать дискретные отсчеты времени t, что необходимо для дальнейшего анализа.

В работе [18] была рассмотрена связь между оператором эволюции $S(t,\tau)$ и динамическим оператором $H(t)$, определяемая уравнением (1.3) с использованием приближения первого порядка. Однако нам нужно найти выражение для оценивания оператора эволюции $S(t,\tau)$ трафика из (1.2) по ряду наблюдаемых значений информативных признаков $F(t)$. После этого оценим некоторые статистики динамического оператора, используя непрямую аналогию связи $H(t)$ с $S(t,\tau)$ из уравнения (1.3).

§ 2. Оценка матрицы оператора эволюции и соответствующих статистик

Запишем решение уравнения (1.2) относительно $S(t,\tau)$ в виде проекционного оператора [23]:

$$S(t,\tau) = \frac{F(t)F^T(\tau)}{F^T(\tau)F(\tau)} = \frac{\begin{Vmatrix} f_1(t)f_1(\tau) & \ldots & f_1(t)f_M(\tau) \\ \vdots & & \vdots \\ f_M(t)f_1(\tau) & \ldots & f_M(t)f_M(\tau) \end{Vmatrix}}{\sum\limits_{m=1}^{M} f_m(\tau)f_m(\tau)}. \tag{2.1}$$

Убедиться, что (2.1) удовлетворяет (1.2), можно простой подстановкой.

На основании (1.1)–(2.1) можно сделать обобщение, позволяющее рассматривать не один оператор эволюции, определяемый вектором F, а семейство операторов эволюции, определенных на различных комбинациях компонент вектора F. Каждый из операторов эволюции будет описывать взаимодействие только соответствующих наборов компонентов. Например, для комбинаций компонентов $f_1(t), f_2(t), \ldots, f_{M-2}(t)$ и $f_{M-1}(t), f_M(t)$ мы получаем два оператора эволюции $S_{1,\ldots,M-2}(t,\tau)$ и $S_{M-1,M}(t,\tau)$, соответствующие матричные элементы которых определяются аналогично (2.1).

Используя семейство операторов эволюции можно сформировать различные статистики трафика. Рассмотрим один из подходов к формированию таких статистик на основе матрицы P, диагональные матричные элементы $p_{m,m}$ которой связаны соотношением нормировки $\sum\limits_{m=1}^{M} p_{m,m} = 1$. Данные матричные элементы определяют статистические веса связей однородных информационных признаков $f_m(t)$ и $f_m(\tau)$. На практике величины $p_{m,m}$ вводятся, как значимости соответствующих информативных признаков трафика ($m = 1, \ldots, M$)). Недиагональные матричные элементы $p_{m,n}$ определяют статистические веса связей разнородных информационных признаков $f_m(t)$ и $f_n(t)$. Для независимых разнородных информативных признаков трафика определим недиагональные элементы соотношением $p_{m,n} = p_{m,m}p_{n,n}$ ($m \neq n$). Далее будем рассматривать равнозначные информативные признаки $p_{m,m} = 1/M$.

Определим среднее значение оператора эволюции трафика $S(t,\tau)$ выражением

$$\overline{S(t,\tau)} = Tr\left[PS(t,\tau)\right] = \frac{1}{M}\frac{\left[\sum\limits_{m=1}^{M} f_m(t)f_m(\tau) + \frac{1}{M}\sum\limits_{m=1,m\neq n}^{M}\sum\limits_{n=1}^{M} f_m(t)f_n(\tau)\right]}{\sum\limits_{m=1}^{M} f_m(\tau)f_m(\tau)}. \tag{2.2}$$

В дальнейшем для упрощения вычислений будем рассматривать значения $\overline{S(t,\tau)}$, лежащие в интервале $(0;1)$. Далее мы применим $\overline{S(t,\tau)}$ для оценивания статистик, связанных с динамическим оператором $H(t)$, используя непрямую аналогию связи $S(t,\tau)$ с $H(t)$ из (1.3).

§ 3. Построение хеш-функций трафика и их использование

В любом реальном трафике значение t и τ неизвестны, а значения компонент его информативных признаков случайны. Поэтому средние значения $\overline{S}(t,\tau)$ также являются случайными величинами. Будем в дальнейшем рассматривать значения случайной хеш-функции Hash трафика, вычисляемой из уравнения

$$\overline{S(t,\tau)} = 1 + Hash(t,\tau) + \frac{1}{2}Hash^2(t,\tau) + \frac{1}{6}Hash^3(t,\tau) + \ldots . \tag{3.1}$$

Будем также рассматривать плотности распределений вероятности $w(Hash)$ значений хеш-функций, определяемых для различных состояний трафика по соседним значениям t и τ. Далее мы рассмотрим две хеш-функции $Hash_{Addr}$ и $Hash_{Load}$, которые определяются в соответствии с (2.2) и (3.1), но на основе информативных признаков компонент $f_1(t), f_2(t), \ldots, f_{M-2}(t)$ и $f_{M-1}(t), f_M(t)$, формируемых отдельно для адресной и нагрузочной частей заголовков пакетов данных трафика.

Шаблоны состояний трафика определим в виде условных плотностей распределения вероятностей $w(Hash_{Addr}|r)$ и $w(Hash_{Load}|r)$ адресных и нагрузочных хеш-функций соответственно, где r — тип трафика ($r = 0, 1, \ldots, R$). Для определенности будем считать, что $r = 0$ маркирует нормальный трафик, а остальные R типов относятся к аномальным типам трафика.

Используя первый порядок приближения в (3.1), вычислим значения H_1 всех адресных и нагрузочных хеш-функций:

$$H_1 = \overline{S} - 1 = \frac{1}{\triangle t}\int_{\tau}^{t} \overline{H(t_1)}\, dt_1. \tag{3.2}$$

Используя третий порядок приближения в (3.1), вычислим значения H_3 всех адресных и нагрузочных хеш-функций:

$$H_3 = \left(-(1-3\overline{S}) + \left[1 + (1-3\overline{S})^2\right]^{1/2}\right)^{1/3} + \left(-(1-3\overline{S}) - \left[1 + (1-3\overline{S})^2\right]^{1/2}\right)^{1/3}. \tag{3.3}$$

Из (3.2) и (3.3) следует, что значения H_1 и H_3 меняются в следующих диапазонах: $-1 < H_1 < 0$ и $-1 < H_3 < 1$.

Обнаружение атак и идентификация их типов будет строиться на основе последовательного анализа Вальда и Байесовской статической теории принятия решений [24]. В последовательном анализе Вальда будем использовать отношения правдоподобий:

$$L = \frac{w(Hash_{Addr}(1)|r)w(Hash_{Load}(1)|r)\dots\ w(Hash_{Addr}(n^*)|r)w(Hash_{Load}(n^*)|r)}{w(Hash_{Addr}(1)|0)w(Hash_{Load}(1)|0)\dots w(Hash_{Addr}(n^*)|0)w(Hash_{Load}(n^*)|0)}, \tag{3.4}$$

где n^* — количество последовательных наблюдений (пакетов или агрегатов пакетов в трафике), при котором реализуются заданные вероятности ошибок первого и второго родов. Байесовские апостериорные вероятности типов атак будут рассчитываться по известной схеме [24]:

$$W\Big(r|Hash_{Addr}(1), Hash_{Load}(1), \dots, Hash_{Addr}(n^*), Hash_{Load}(n^*)\Big) =$$
$$= \frac{w(Hash_{Addr}(1)|r)w(Hash_{Load}(1)|r)\dots w(Hash_{Addr}(n^*)|r)w(Hash_{Load}(n^*)|r)}{\sum\limits_{r=1}^{R} w(Hash_{Addr}(1)|r)w(Hash_{Load}(1)|r)\dots w(Hash_{Addr}(n^*)|r)w(Hash_{Load}(n^*)|r)}, \tag{3.5}$$

для $r = 1, 2, \dots, R$.

Из (3.4) и (3.5) следует, что для эффективной идентификации типов атак необходимо отличие распределений $w(Hash_{Addr}|r)$ и $w(Hash_{Load}|r)$ адресных и нагрузочных хеш-функций для различных r $(r = 0, 1, \dots, R)$.

§ 4. Экспериментальные исследования

В компьютерном эксперименте анализировался трафик, поступающий по протоколу NetFlow. В таблице 1 приведен пример содержания заголовков поступающих пакетов данных. Каждая строка таблицы 1 содержит значения соответствующих полей адресной и нагрузочной частей заголовков пакетов данных.

Набор информативных признаков трафика был сформирован следующим образом: параметры из адресных полей были преобразованы в двоичные значения $f_m(t)$ $(m = 1, 2, \dots, 6)$, представляющие собой индикаторы изменения (характеристика равна 1, если параметр в пакете отличается от предыдущего пакета); параметры из нагрузочных полей $f_m(t)$ $(m = 7, 8)$ были взяты без изменений. Временные интервалы между моментами получения пакетов не учитывались. Таким образом, для набора параметров в таблице 1 мы получим информативные признаки, представленные в таблице 2.

Таблица 1. Пример структуры заголовков пакетов, агрегируемых NetFlow протоколом

	Адресные поля							Нагрузочные поля	
№	SrcAddr	DstAddr	SrcPort	DstPort	Protocol	TCPFlags	№	Octets	Packets
0	192.168.17.16	192.168.66.16	6	39474	80	0x1b		513	8
1	192.168.18.16	192.168.64.16	6	38462	80	0x1b	1	513	8
2	192.168.64.16	192.168.18.16	6	80	38462	0x1b	2	13622	12
3	192.168.66.16	192.168.17.16	6	80	39474	0x1b	3	4544	6
4	192.168.64.16	192.168.20.16	6	80	56204	0x1b	4	13622	12
5	192.168.20.16	192.168.64.16	6	56204	80	0x1b	5	513	8
6	192.168.64.16	192.168.19.16	6	80	50810	0x1b	6	2154	5
7	192.168.19.16	192.168.64.16	6	50810	80	0x1b	7	409	6
8	192.168.19.16	192.168.66.16	6	44848	80	0x1b	8	2125	39
9	192.168.20.16	192.168.64.16	6	56206	80	0x1b	9	409	6

Все шаблоны трафика формировались по хронологически соседним пакетам трафика. Для построения использовалось 4000 пакетов для каждого типа трафика (Normal, HttpFlood, SlowLoris) на этапе обучения.

На рисунке 1 показаны распределения $w(Hash_{Addr}|r)$ адресных хеш-функций H_1, вычисленные по формуле (3.1) с применением приближения первого порядка. Они были рассчитаны

Таблица 2. Информативные признаки трафика

	Адресные поля							Нагрузочные поля	
№	f_1	f_2	f_3	f_4	f_5	f_6	№	f_7	f_8
1	1	1	0	1	0	0	1	513	8
2	1	1	0	1	1	0	2	13622	12
3	1	1	0	0	1	0	3	4544	6
4	1	1	0	0	1	0	4	13622	12
5	1	1	0	1	1	0	5	513	8
6	1	1	0	1	1	0	6	2154	5
7	1	1	0	1	1	0	7	409	6
8	0	1	0	1	0	0	8	2125	39
9	1	1	0	1	0	0	9	409	6

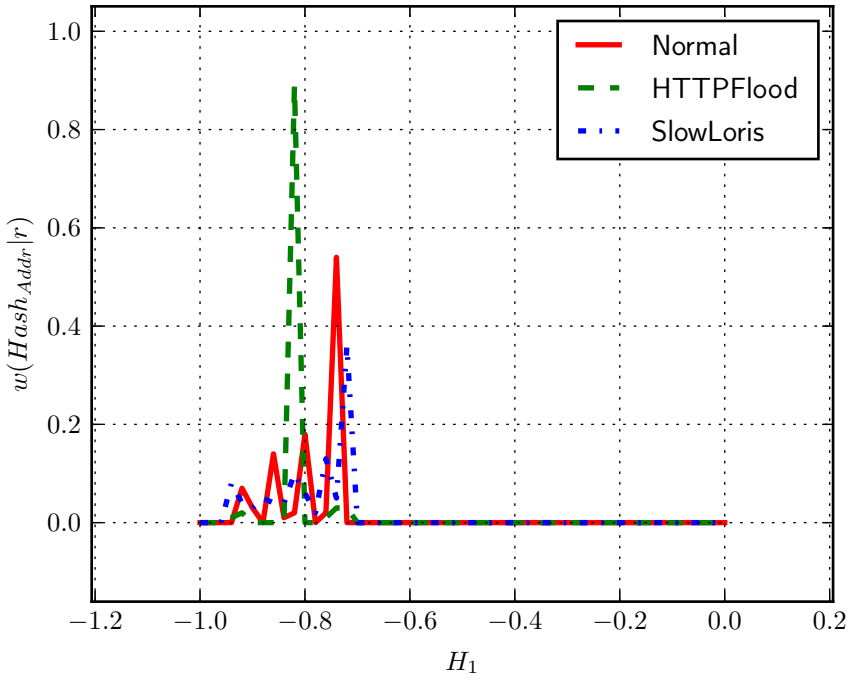

Рис. 1. Распределения адресных хеш-функций, сформированных на основе значений H_1

на основании шести информативных признаков $(f_1, \ldots, f_6)$ для трех типов трафика $(r = 0, 1, 2)$: нормального, двух аномальных (HttpFlood, SlowLoris). Распределения представляют собой полученные шаблоны состояния трафика.

На рисунке 2 показаны распределения $w(Hash_{Addr}|r)$ адресных хеш-функций H_3, вычисленные по формуле (3.1) с применением приближения третьего порядка. Они были рассчитаны на основании шести информативных признаков $(f_1, \ldots, f_6)$ для трех типов трафика $(r = 0, 1, 2)$: нормального, двух аномальных (HttpFlood, SlowLoris).

На рисунке 3 показаны распределения $w(Hash_{Load}|r)$ нагрузочных хеш-функций H_1, вычисленные по формуле (3.1) с применением приближения первого порядка. Они были рассчитаны на основании двух информативных признаков $(f_7,\ f_8)$ для трех типов трафика $(r = 0, 1, 2)$: нормального, двух аномальных (HttpFlood, SlowLoris).

На рисунке 4 показаны распределения $w(Hash_{Load}|r)$ нагрузочных хеш-функций H_3, вычисленные по формуле (3.1) с применением приближения третьего порядка. Они были рассчитаны на основании двух информативных признаков $(f_7,\ f_8)$ для трех типов трафика $(r = 0, 1, 2)$: нормального, двух аномальных (HttpFlood, SlowLoris).

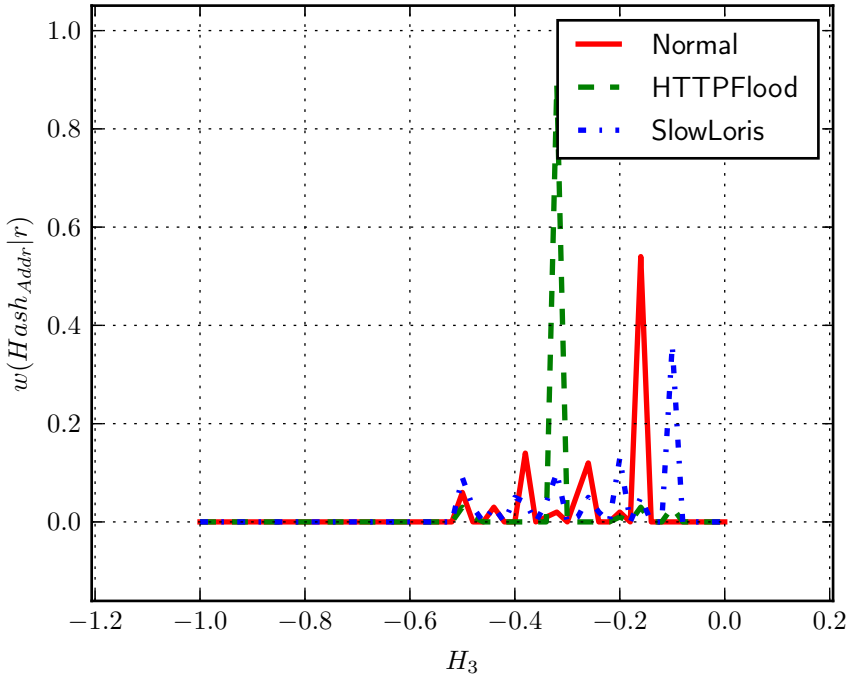

Рис. 2. Распределения адресных хеш-функций, сформированных на основе значений H_3

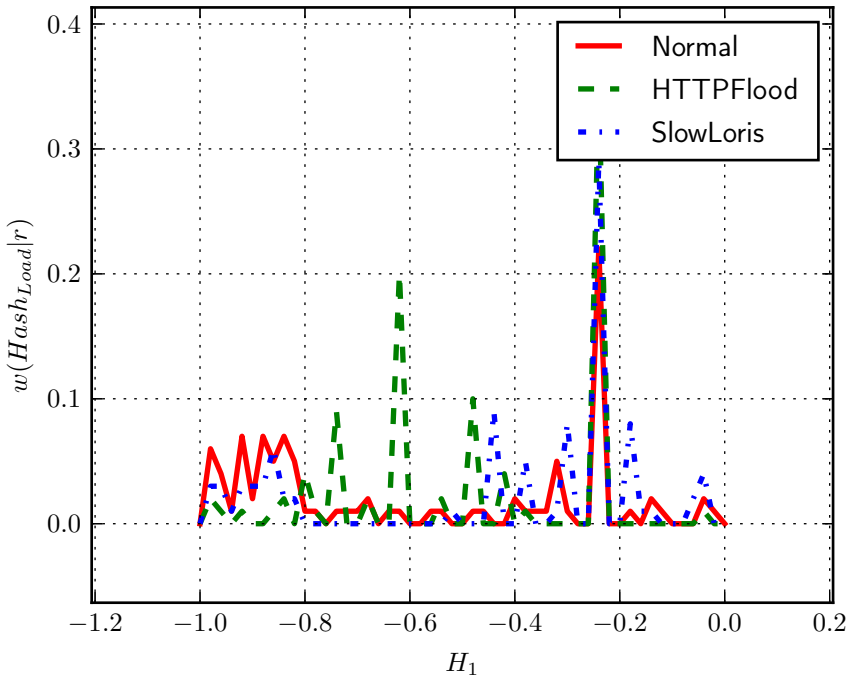

Рис. 3. Распределения нагрузочных хеш-функций, сформированных на основе значений H_1

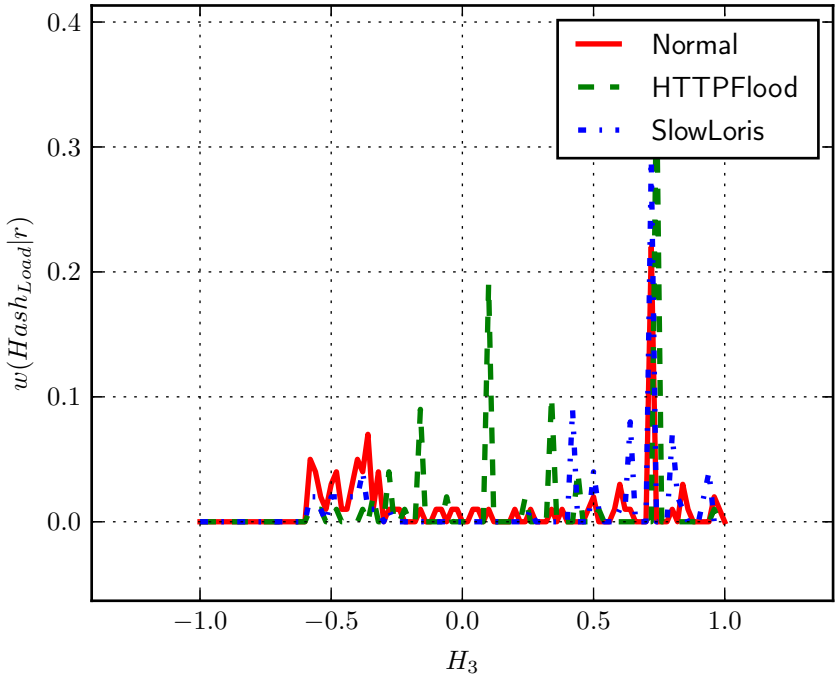

Рис. 4. Распределения нагрузочных хеш-функций, сформированных на основе значений H_3

Обсуждение результатов

Внимательно анализируя распределения рисунков 1–4, заметим, что все распределения заметно различаются, независимо от того сформированы они по средним значениям H_1 или H_3. Хотя распределения, сформированные по значениям H_1 или H_3, могут выглядеть схожими, они имеют специфические особенности. Например, на рисунке 4 видны максимальные пики распределения $w(Hash_{Load}|HttpFlood)$ и $w(Hash_{Load}|SlowLoris)$, а на рисунке 3 они перекрываются. Таким образом, хеш-функции, полученные с использованием приближения третьего порядка, более предпочтительны для классификации трафика.

Предлагаемый метод позволяет использовать небольшое количество пакетов NetFlow для классификации трафика по сравнению с корреляционным и спектральным анализом. Другими словами, используя то же количество пакетов для анализа, предлагаемый метод будет выполнять классификацию трафика с большей точностью.

Заключение

Работа посвящена новому методу структурного анализа трафика с использованием оператора эволюции и хеш-функций информативных признаков трафика. Основным преимуществом предлагаемого метода является возможность использования небольшого количества пакетов для формирования хеш-функций, используемых для классификации типов трафика. В частности, при компьютерном моделировании значения хеш-функций для адресных и нагрузочных параметров заголовков NetFlow-пакетов данные были сформированы из пары хронологически последовательных пакетов. Тем не менее показано, что распределения хеш-функций для трех типов трафика (Normal, HttpFlood, SlowLoris) существенно отличаются друг от друга, что позволяет эффективно применять в дальнейшим разработанный алгоритмы классификации трафика. В будущем авторы намерены исследовать спектральные характеристики различных семейств операторов эволюции в применении к задаче классификации различных типов атак.

Финансирование. Работа выполнена при финансовой поддержке Минобрнауки РФ в рамках федеральной целевой программы «Исследования и разработки по приоритетным направле-

ниям развития научно-технологического комплекса России на 2014—2020 годы» по теме: «Разработка эффективных алгоритмов обнаружения сетевых атак, основанных на выявлении отклонений в трафике сверхбольших объемов, поступающем на пограничные маршрутизаторы сети передачи данных, и создание на их основе образца программного комплекса обнаружения и предотвращения угроз безопасности информации, направленных на отказ в обслуживании». Уникальный идентификатор работ (проекта) RFMEFI57817X0261

СПИСОК ЛИТЕРАТУРЫ

1. Zeb K., Baig O., Asif M.K. DDoS attacks and countermeasures in cyberspace // 2015 2nd World Symposium on Web Applications and Networking (WSWAN). IEEE, 2015. DOI: 10.1109/WSWAN.2015.7210322
2. Singh K., Dhindsa K.S., Bhushan B. Distributed defense: an edge over centralized defense against DDoS attacks // International Journal of Computer Network and Information Security (IJCNIS). 2017. Vol. 9. No. 3. P. 36–44.
3. Bhattacharyya D.K., Kalita J.K. DDoS attacks: evolution, detection, prevention, reaction, and tolerance. CRC Press, 2016. 312 p.
4. Li M. An approach to reliably identifying signs of DDoS flood attacks based on LRD traffic pattern recognition // Computers and Security. 2004. Vol. 23. No. 7. P. 549–558. DOI: 10.1016/j.cose.2004.04.005
5. Yu S., Zhou W., Jia W., Guo S., Xiang Y., Tang F. Discriminating DDoS attacks from flash crowds using flow correlation coefficient // IEEE Transactions on Parallel and Distributed Systems. 2012. Vol. 23. Issue 6. P. 1073–1080. DOI: 10.1109/TPDS.2011.262
6. Jin S., Yeung D.S. A covariance analysis model for DDoS attack detection // 2004 IEEE International Conference on Communications (IEEE Cat. No.04CH37577). IEEE, 2004. DOI: 10.1109/icc.2004.1312847
7. Wu Z., Wang M., Zhang H., Liu X. Correlation-based detection of LDoS attack // Journal of Software. 2012. Vol. 7. No. 10. DOI: 10.4304/jsw.7.10.2341-2348
8. Котенко И.В., Федорченко А.В., Саенко И.Б., Кушнеревич А.Г. Технологии больших данных для корреляции событий безопасности на основе учета типов связей // Вопросы кибербезопасности. 2017. №5 (24). С. 2–16. DOI: 10.21681/2311-3456-2017-5-2-16
9. Cheng C.M., Kung H.T., Tan K.S. Use of spectral analysis in defense against DoS attacks // Global Telecommunications Conference, 2002. GLOBECOM '02. IEEE. 2002. DOI: 10.1109/glocom.2002.1189011
10. Chen Y., Hwang K. Spectral analysis of TCP flows for defense against reduction-of-quality attacks // 2007 IEEE International Conference on Communications. IEEE, 2007. DOI: 10.1109/icc.2007.204
11. Fouladi R.F., Seifpoor T., Anarim E. Frequency characteristics of DoS and DDoS attacks // 2013 21st Signal Processing and Communications Applications Conference (SIU). IEEE, 2013. DOI: 10.1109/SIU.2013.6531200
12. Fouladi R.F., Kayatas C.E., Anarim E. Frequency based DDoS attack detection approach using naive Bayes classification // 2016 39th International Conference on Telecommunications and Signal Processing (TSP). IEEE, 2016. DOI: 10.1109/TSP.2016.7760838
13. Li L., Lee G. DDoS attack detection and wavelets // Telecommunication Systems. 2005. Vol. 28. Issue 3–4. P. 435–451. DOI: 10.1007/s11235-004-5581-0
14. Li M., Li M. A new approach for detecting DDoS attacks based on wavelet analysis // 2009 2nd International Congress on Image and Signal Processing. IEEE, 2009. DOI: 10.1109/CISP.2009.5300903
15. Salagean M., Firoiu I. Anomaly detection of network traffic based on analytical discrete wavelet transform // 2010 8th International Conference on Communications. IEEE, 2010. DOI: 10.1109/ICCOMM.2010.5509071
16. Dingde J., Wenda Q., Laisen N., Cheng Y., Rongfang L. Time-frequency detection algorithm of network traffic anomalies // International Proceedings of Computer Science and Information Technology. 2012. Vol. 36. P. 103–108. http://www.ipcsit.com/vol36/021-ICIIM2012-M0053.pdf
17. Cheng J., Yin J., Liu Y., Cai Z., Wu C. DDoS attack detection using IP address feature interaction // 2009 International Conference on Intelligent Networking and Collaborative Systems. IEEE, 2009. DOI: 10.1109/incos.2009.34
18. Galayev V.S., Krasnov A.E., Nikol'skii D.N., Repin D.S. The space of structural features for increasing the efficiency of the algorithms for detecting network attacks, based on the detection of anomalies in the traffic of extremely large volumes // International Journal of Applied Engineering Research. 2017. Vol. 12. No. 21. P. 10781–10790. http://www.ripublication.com/ijaer17/ijaerv12n21_35.pdf

19. Демидович Б.П. Лекции по математической теории устойчивости. М: Наука, 1967. 472 с.
20. Ситенко А.Г. Теория рассеяния (курс лекций). Изд. 2. Киев: Вища школа, 1975. 256 с.
21. Peano G. Integration par series des equations differentielles lineaires // Mathematische Annalen. 1888. Vol. 32. Issue 3. P. 450–456. DOI: 10.1007/BF01443609
22. Dyson F.J. The radiation theories of Tomonaga, Schwinger, and Feynman // Physical Review. 1949. Vol. 75. Issue 3. P. 486–502. DOI: 10.1103/physrev.75.486
23. Краснов А.Е., Надеждин Е.Н., Никольский Д.Н., Галяев В.С. Применение метода оператора эволюции к анализу многомерных временных рядов // Алгебра, теория чисел и дискретная геометрия: современные проблемы и приложения: Материалы XV Международной конференции, посвященной столетию со дня рождения профессора Николая Михайловича Коробова. Тула: ТГПУ им. Л.Н. Толстого, 2018. С. 300–303. http://www.mathnet.ru/ConfLogos/1304/Conference2018_1.pdf
24. Wald A. Sequential analysis. New York: J. Wiley & Sons, Inc., 1947. 212 p.
25. Krasnov A.E., Nadezhdin E.N., Galayev V.S., Zykova E.A., Nikol'skii D.N., Repin D.S. DDoS attack detection based on network traffic phase coordinates analysis // International Journal of Applied Engineering Research. 2018. Vol. 13. No. 8. P. 5647–5654. http://www.ripublication.com/ijaer18/ijaerv13n8_11.pdf

Поступила в редакцию 15.06.2018

Краснов Андрей Евгеньевич, д. ф.-м. н., профессор, главный научный сотрудник, Государственный институт информационных технологий и телекоммуникаций, 125315, Россия, г. Москва, ул. Часовая, 21Б.
E-mail: a.krasnov@informika.ru

Надеждин Евгений Николаевич, д. т. н., профессор, главный научный сотрудник, Государственный институт информационных технологий и телекоммуникаций, 125315, Россия, г. Москва, ул. Часовая, 21Б.
E-mail: e.nadezhdin@informika.ru

Никольский Дмитрий Николаевич, к. ф.-м. н., ведущий научный сотрудник, Государственный институт информационных технологий и телекоммуникаций, 125315, Россия, г. Москва, ул. Часовая, 21Б.
E-mail: d.nikolsky@informika.ru

Репин Дмитрий Сергеевич, к. т. н., заместитель директора, Государственный институт информационных технологий и телекоммуникаций, 125009, Россия, г. Москва, Брюсов пер., 21, строение 2.
E-mail: r_d_s@informika.ru

Галяев Владимир Сергеевич, к. ф.-м. н., старший научный сотрудник, Государственный институт информационных технологий и телекоммуникаций, 125315, Россия, г. Москва, ул. Часовая, 21Б.
E-mail: v.galiaev@informika.ru

A. E. Krasnov, E. N. Nadezhdin, D. N. Nikol'skii, D. S. Repin, V. S. Galyaev

Detecting DDoS attacks by analyzing the dynamics and interrelation of network traffic characteristics

Citation: *Vestnik Udmurtskogo Universiteta. Matematika. Mekhanika. Komp'yuternye Nauki*, 2018, vol. 28, issue 3, pp. 407–418 (in Russian).

Keywords: network traffic, DDoS attack, detection, dynamical operator, evolution operator, hash function, classification.

MSC2010: 90B20, 47A62

DOI: 10.20537/vm180310

This paper presents an improved approach previously developed by the authors for detection of DDoS attacks. It uses traffic evolution and dynamical operators, which makes it possible to take into consideration interrelations observed for data packets headers of traffic. It is assumed that each traffic state (normal state and anomalous attacked states) can be described by unique temporal patterns of characteristics generated by unknown linear dynamical operators. Interrelations between values of network traffic characteristics in different discrete time samples are determined by the evolution operator. The approach was applied for classification of three traffic states: normal and two abnormal (HTTP flood and SlowLoris DDoS attacks).

The results prove that it is possible to distinguish normal and abnormal traffic states by hash functions of address and load fields of traffic data packets.

Funding. The work was supported by the Ministry of Education and Science of Russian Federation by lot code 2017–14–579–0002 on the topic: "The development of effective algorithms for detection network attacks based on identifying of deviations in the traffic of extremely large volumes arriving at the border routers of the data network and creating a sample of software complex for detection and prevention of information security threats aimed at denial of service". The unique identifier of the work (project) is RFMEFI57817X0261.

REFERENCES

1. Zeb K., Baig O., Asif M.K. DDoS attacks and countermeasures in cyberspace, *2015 2nd World Symposium on Web Applications and Networking (WSWAN)*, IEEE, 2015. DOI: 10.1109/WSWAN.2015.7210322
2. Singh K., Dhindsa K.S., Bhushan B. Distributed defense: an edge over centralized defense against DDoS attacks, *International Journal of Computer Network and Information Security (IJCNIS)*, 2017, vol. 9, no. 3, pp. 36–44.
3. Bhattacharyya D.K., Kalita J.K. *DDoS attacks: evolution, detection, prevention, reaction, and tolerance*. CRC Press, 2016, 312 p.
4. Li M. An approach to reliably identifying signs of DDoS flood attacks based on LRD traffic pattern recognition, *Computers and Security*, 2004, vol. 23, no. 7, pp. 549–558. DOI: 10.1016/j.cose.2004.04.005
5. Yu S., Zhou W., Jia W., Guo S., Xiang Y., Tang F. Discriminating DDoS attacks from flash crowds using flow correlation coefficient, *IEEE Transactions on Parallel and Distributed Systems*, 2012, vol. 23, issue 6, pp. 1073–1080. DOI: 10.1109/TPDS.2011.262
6. Jin S., Yeung D.S. A covariance analysis model for DDoS attack detection, *2004 IEEE International Conference on Communications (IEEE Cat. No.04CH37577)*, IEEE, 2004. DOI: 10.1109/icc.2004.1312847
7. Wu Z., Wang M., Zhang H., Liu X. Correlation-based detection of LDoS attack, *Journal of Software*, 2012, vol. 7, no. 10. DOI: 10.4304/jsw.7.10.2341-2348
8. Kotenko I., Fedorchenko A., Saenko I., Kushnerevich A. Big data technologies for security event correlation based on event type accounting, *Voprosy Kiberbezopasnosti*, 2017, no. 5 (24), pp. 2–16 (in Russian). DOI: 10.21681/2311-3456-2017-5-2-16
9. Cheng C.M., Kung H.T., Tan K.S. Use of spectral analysis in defense against DoS attacks, *Global Telecommunications Conference, 2002. GLOBECOM '02. IEEE*, 2002. DOI: 10.1109/glocom.2002.1189011
10. Chen Y., Hwang K. Spectral analysis of TCP flows for defense against reduction-of-quality attacks, *2007 IEEE International Conference on Communications*, IEEE, 2007. DOI: 10.1109/icc.2007.204
11. Fouladi R.F., Seifpoor T., Anarim E. Frequency characteristics of DoS and DDoS attacks, *2013 21st Signal Processing and Communications Applications Conference (SIU)*, IEEE, 2013. DOI: 10.1109/SIU.2013.6531200
12. Fouladi R.F., Kayatas C.E., Anarim E. Frequency based DDoS attack detection approach using naive Bayes classification, *2016 39th International Conference on Telecommunications and Signal Processing (TSP)*, IEEE, 2016. DOI: 10.1109/TSP.2016.7760838
13. Li L., Lee G. DDoS attack detection and wavelets, *Telecommunication Systems*, 2005, vol. 28, issue 3–4, pp. 435–451. DOI: 10.1007/s11235-004-5581-0
14. Li M., Li M. A new approach for detecting DDoS attacks based on wavelet analysis, *2009 2nd International Congress on Image and Signal Processing*, IEEE, 2009. DOI: 10.1109/CISP.2009.5300903
15. Salagean M., Firoiu I. Anomaly detection of network traffic based on analytical discrete wavelet transform, *2010 8th International Conference on Communications*, IEEE, 2010. DOI: 10.1109/ICCOMM.2010.5509071
16. Dingde J., Wenda Q., Laisen N., Cheng Y., Rongfang L. Time-frequency detection algorithm of network traffic anomalies, *International Proceedings of Computer Science and Information Technology*, 2012, vol. 36, pp. 103–108. http://www.ipcsit.com/vol36/021-ICIIM2012-M0053.pdf
17. Cheng J., Yin J., Liu Y., Cai Z., Wu C. DDoS attack detection using IP address feature interaction, *2009 International Conference on Intelligent Networking and Collaborative Systems*, IEEE, 2009. DOI: 10.1109/incos.2009.34
18. Galayev V.S., Krasnov A.E., Nikol'skii D.N., Repin D.S. The space of structural features for increasing the efficiency of the algorithms for detecting network attacks, based on the detection of anomalies in the traffic of extremely large volumes, *International Journal of Applied Engineering Research*, 2017, vol. 12, no. 21, pp. 10781–10790. http://www.ripublication.com/ijaer17/ijaerv12n21_35.pdf

19. Demidovich B.P. *Lektsii po matematicheskoi teorii ustoichivosti* (Lectures on the mathematical theory of stability), Moscow: Nauka, 1967, 472 p.
20. Sitenko A.G. *Teoriya rasseyaniya (kurs lektsii)* (Theory of scattering (course of lectures)), Kiev: Vishcha Shkola, 1975, 256 p.
21. Peano G. Integration par series des equations differentielles lineaires, *Mathematische Annalen*, 1888, vol. 32, issue 3, pp. 450–456. DOI: 10.1007/BF01443609
22. Dyson F.J. The radiation theories of Tomonaga, Schwinger, and Feynman, *Physical Review*, 1949, vol. 75, issue 3, pp. 486–502. DOI: 10.1103/physrev.75.486
23. Krasnov A.E., Nadezhdin E.N., Nikol'skii D.N., Galyaev V.S. Application of the evolution operator method to the analysis of multidimensional time series, *Algebra, Number Theory and Discrete Geometry: modern problems and applications: Proceedings of XV International Conference dedicated to the centenary of the birth of Professor Nikolai Mikhailovich Korobov*, Tula State Pedagogical University, Tula, 2018, pp. 313–316 (in Russian). http://www.mathnet.ru/ConfLogos/1304/Conference2018_1.pdf
24. Wald A. *Sequential analysis*. J. Wiley & Sons, Inc., New York, 1947, 212 p.
25. Krasnov A.E., Nadezhdin E.N., Galayev V.S., Zykova E.A., Nikol'skii D.N., Repin D.S. DDoS attack detection based on network traffic phase coordinates analysis, *International Journal of Applied Engineering Research*, 2018, vol. 13, no. 8, pp. 5647–5654.
http://www.ripublication.com/ijaer18/ijaerv13n8_11.pdf

Received 15.06.2018

Krasnov Andrey Evgenievich, Doctor of Physics and Mathematics, Professor, Chief Researcher, State Institute of Information Technologies and Telecommunications, ul. Chasovaya, 21B, Moscow, 125315, Russia.
E-mail: a.krasnov@informika.ru

Nadezhdin Evgeniy Nikolaevich, Doctor of Engineering, Professor, Chief Researcher, State Institute of Information Technologies and Telecommunications, ul. Chasovaya, 21B, Moscow, 125315, Russia.
E-mail: e.nadezhdin@informika.ru

Nikol'skii Dmitrii Nikolaevich, Candidate of Physics and Mathematics, Leading Researcher, State Institute of Information Technologies and Telecommunications, ul. Chasovaya, 21B, Moscow, 125315, Russia.
E-mail: d.nikolsky@informika.ru

Repin Dmitrii Sergeevich, Candidate of Engineering, Deputy Director, State Institute of Information Technologies and Telecommunications, Bryusov per., 21, bld. 2, Moscow, 125009, Russia.
E-mail: r_d_s@informika.ru

Galyaev Vladimir Sergeevich, Candidate of Physics and Mathematics, Associate Professor, Senior Researcher, State Institute of Information Technologies and Telecommunications, ul. Chasovaya, 21B, Moscow, 125315, Russia.
E-mail: v.galiaev@informika.ru