Анализ и прогнозирование вредоносного сетевого трафика в облачных сервисах

М.В. Тумбинская
E-mail: tumbinskaya@inbox.ru

Б.И. Баянов
E-mail: bayanov_bulat@mail.ru

Р.Ж. Рахимов
E-mail: rahimov96@mail.ru

Н.В. Кормильцев
E-mail: kormiltcev@hotmail.com

А.Д. Уваров
E-mail: obg-96@mail.ru

Казанский национальный исследовательский технический университет им. А.Н. Туполева
Адрес: 420111, г. Казань, ул. К. Маркса, д. 10

Аннотация

В настоящее время решением одной из основных задач в обеспечении информационной безопасности облачных сервисов как крупных компаний, так и обычных пользователей является правильное определение и прогнозирование сетевого трафика злоумышленника. В статье представлена статистика угроз информационной безопасности, описана классификация угроз информационной безопасности облачных сервисов, определены цели злоумышленников, предложены контрмеры.

Остро стоит проблема определения наиболее эффективной методики, которая может быть применена в средствах защиты облачных сервисов от различного рода сетевых атак, а также для анализа сетевого трафика. В качестве одной из методик была выбрана и рассмотрена методика, основанная на аддитивной модели временных рядов, которая позволяет решить задачу прогнозирования неблагоприятного сетевого трафика. Для проверки работоспособности выбранной методики получены количественные показатели неблагоприятного сетевого трафика путем моделирования реализации сетевой атаки и фиксации эмпирических показателей, описывающих этот процесс. Для этого рассмотрена вредоносная программа, воспроизводящая сетевую атаку, и программа, предназначенная для получения и обработки необходимых для исследования эмпирических показателей.

На основе полученных исходных данных проанализирована эффективность применения методики, использующей аддитивную модель временных рядов. Также показано, что данная методика применима в исследованиях общей динамики числа сетевых атак, совершаемых в интернет-пространстве, что позволяет обнаруживать связи между динамикой числа совершаемых злоумышленниками сетевых атак и особенностями календарного периода. Результаты применения методики показывают, что, основываясь на показателях, описывающих сетевой трафик, можно обнаруживать, а затем и прогнозировать вредоносные действия злоумышленников.

Ключевые слова: прогнозирование; DDOS-атака; облачные сервисы; сетевой трафик; моделирование; аддитивная модель временных рядов; автокорреляционная функция; оценка погрешности.

Цитирование: Тумбинская М.В., Баянов Б.И., Рахимов Р.Ж., Кормильцев Н.В., Уваров А.Д. Анализ и прогнозирование вредоносного сетевого трафика в облачных сервисах // Бизнес-информатика. 2019. Т. 13. № 1. С. 71–81.

DOI: 10.17323/1998-0663.2019.1.71.81

Введение

На сегодняшний день с развитием инфраструктуры современных предприятий потребность в облачных технологиях сильно возросла, поскольку это удобно, экономично, мобильно, быстро и надежно. Облачные технологии позволяют использовать облачные сервисы. Облачный сервис [1] – это интернет-сервис, предполагающий передачу части объектов ИТ-инфраструктуры [2] на обслуживание сторонней организации (аутсорсинг). По статистике RightScale, в 2017 году 95% организаций использовали какую-либо из моделей развертывания облачных сервисов [3]. Рынок облачных сервисов в России, по мнению экспертов Orange Business Services, достигает порядка 24,6 млрд рублей [4]. Анализ работ [2, 5] показывает, что без использования облачных технологий современные ИТ-компании становятся неконкурентоспособными, что приводит к снижению их прибыли. Крупные компании давно используют облачные сервисы (Google Диск, iCloud от Apple, Облако mail.ru).

Облачные сервисы все больше требуют решения вопросов информационной безопасности, так как внедрение новых технологий ведет к появлению все большего количества видов угроз и уязвимостей в системе обеспечения информационной безопасности. По данным опроса «Лаборатории Касперского» [6], 13% российских компаний за год хотя бы однажды столкнулись с инцидентами, связанными с безопасностью облачной инфраструктуры. При этом около 32% компаний потеряли данные в результате этих инцидентов. Поэтому решение задачи обеспечения информационной безопасности в облачных сервисах является актуальной.

Научная новизна работы состоит в том, что представленная модернизированная методика анализа и прогнозирования сетевого трафика, основанная на аддитивной модели временных рядов и интегрированная в средства защиты, может обеспечить необходимый уровень безопасности облачных хранилищ, защищая их от сетевых атак различного рода. К сожалению, многие представленные на рынке методики обеспечения информационной безопасности данных не могут полностью решить задачу прогнозирования неблагоприятного сетевого трафика.

1. Возможность интеграции представленной методики в WAF

В работе [7], атаки большинства злоумышленников в основном строятся на типичных хакерских средствах, доводя их до автоматизма. Исходя из этого, необходимо определить механизмы, включающие в себя постоянный процесс обучения, для постепенного ухода от сигнатурного анализа. В статье [7] также отмечается, что на сегодняшний день некоторые производители защитных экранов для веб-приложений (web application firewall, WAF) постепенно отходят от сигнатурного анализа, перенаправляя свои усилия в сторону обновлений сигнатур. WAF должен иметь широкую базу данных сигнатур нелегитимного трафика и действий, применимых для всех типов веб-приложений, для формирования модели защиты, гарантирующей достаточный уровень безопасности. Рассмотренную методику анализа и прогнозирования сетевого трафика, основанную на аддитивной модели временных рядов, в последующем можно интегрировать в такие сложные по своему строению WAF. При этом основой будет являться не прогнозирование действий злоумышленника и авторизованного пользователя, а создание модели защиты, которая может базироваться не только на URL (uniform resource locator единый указатель ресурса, который позволяет определить местонахождение ресурса в сети), но и на параметрах и файлах cookie. После создания модели защиты возникает вопрос тестирования, а именно – анализа проходящего трафика, с целью предотвращения использования как известных, так и неизвестных уязвимостей.

2. Классификация угроз информационной безопасности облачных сервисов

Рассмотрим классификацию угроз информационной безопасности в облачных сервисах. В *таблице 1* перечислены наиболее популярные виды угроз (на основе работы [6]). Для каждого из видов угроз указана возможная цель злоумышленника при реа-

Таблица 1.

Виды угроз информационной безопасности облачных сервисов, цели злоумышленника и методы защиты от угроз

№	Угроза информационной безопасности облачных сервисов	Цель злоумышленника	Контрмеры
1.	Кража данных	Получить доступ к базе данных (например, к адресам электронной почты пользователей)	Децентрализация базы данных и шифрование данных с использованием SSL–сертификата
2.	Потеря данных	Модификация или уничтожение информации базы данных	Резервное копирование данных и ограничение прав доступа
3.	Кража аккаунтов / взлом услуг	Модификация или уничтожение информации базы данных	Двухфакторная аутентификация
4.	Незащищенные интерфейсы и API	Полный доступ к информации базы данных	Аутентификация, управление доступом, шифрование
5.	DDOS–атаки	Отказ в обслуживании облачного сервиса санкционированным пользователям	Управление доступом
6.	Злонамеренный инсайдер	Доступ к информации базы данных	Управление доступом
7.	Использование облачных ресурсов хакерами	Доступ к вычислительным ресурсам облачной инфраструктуры	Ограничение вычислительных мощностей системы

лизации угрозы соответствующего вида. Также для каждого из видов угроз указаны методы защиты информации. Ни один из методов не гарантирует защиту от всех видов угроз, поэтому на практике полностью избежать таких угроз не удается. Статистические показатели, описывающие каждую из реализаций угроз, могут быть протоколироваться в системе и использоваться для дальнейшего анализа при построении новой системы защиты.

3. Моделирование сетевой атаки

Для анализа сетевого трафика, поступающего на узлы выстроенной администратором сети, специалисты по информационной безопасности используют специальные программные средства, установленные на узлах сети. В исследовании использовалось программное обеспечение Wireshark (версии 2.6.1). Этот инструмент позволяет захватить и проанализировать сетевой трафик на различных общепринятых сетевых протоколах (TCP, UDP, HTTP и т.д.).

В работах [8, 9] представлены данные о сетевом трафике, описывающем реализацию DDOS-атак. Однако этих данных оказалось недостаточно, поэтому было принято решение о получении собственных данных путем имитационного моделирования сетевой атаки, основываясь на алгоритмах, представленных в работе [10]. Для этого потребовались два узла сконфигурированной администратором сети, один из которых использовался в качестве устройства жертвы, а второй – в качестве устройства злоумышленника. Устройствами жертвы и злоумышленника послужили виртуальные машины, установленные на одном персональном компьютере. На устройстве жертвы было установлено программное обеспечение Wireshark, а на устройстве злоумышленника – программа LOIC (программа с открытым исходным кодом, предназначенная для осуществления DDOS-атак)[1], воспроизводящая поток неблагоприятного для узла жертвы трафика.

В нашем исследовании предполагается, что злоумышленник проводит сетевые атаки периодически со всевозможными начальными конфигурациями вредоносной программы, при этом не исключается санкционированное использование сети жертвой. Сетевой поток (количество сетевых пакетов в секунду), проходящий через сетевой узел жертвы, представлен на *рисунке 1.*

Как видно из рисунка, из общего сетевого потока нельзя выделить поток, относящийся к конкретному пользователю, поэтому рекомендуется рассматривать сетевые потоки отдельных пользователей.

Для удобства анализа можно выполнить фильтрацию сетевого трафика, проходящего через узел жертвы, выделив только те пакеты, которые поступают от узла злоумышленника. Сетевой поток, приходящий с узла злоумышленника, изображен на *рисунке 2.* Он представляет собой периодически возрастающий по величине на несколько порядков физический процесс, который характеризует действия отдельного пользователя.

[1] https://www.darknet.org.uk/2017/10/loic-download-low-orbit-ion-cannon-ddos-booter/

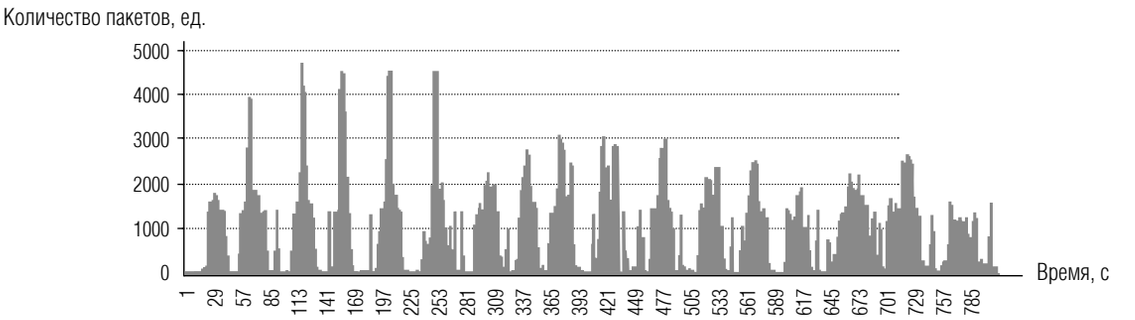

Рис. 1. Количество сетевых пакетов, проходящих через узел жертвы, в секунду

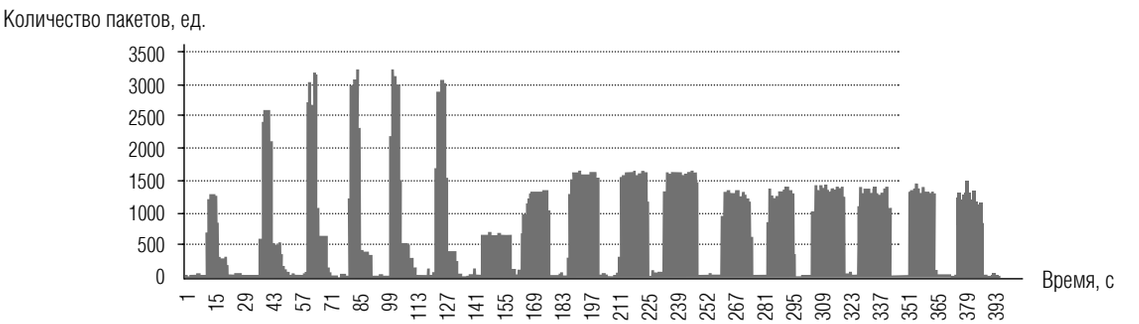

Рис. 2. Количество сетевых пакетов, приходящих от узла злоумышленника, в секунду

Обнаружить адрес узла злоумышленника, осуществляющего сетевую атаку, можно путем анализа параметров, описывающих приходящий сетевой трафик, например, плотности распределения количества пакетов по их размеру в битах. В *таблице 2* представлены данные, полученные путем моделирования сетевых атак и моделирования благоприятного сетевого трафика с помощью программы Wireshark.

Результаты процесса моделирования показывают, что при неблагоприятном сетевом трафике более 92% общего числа пакетов имеют размер 40–79 бит. В то же время при благоприятном сетевом трафике доля пакетов этого размера составляет около 39%, при этом более 42% пакетов имею размер 1280–2559 бит, а около 11% – размер 640–1279 бит. Подозрительным также можно считать передачу трафика с большой интенсивностью (число пакетов в единицу времени) или какой-либо другой неестественный характер сетевого трафика. В качестве выборки для анализа принимаются данные о количестве сетевых пакетов, приходящих от узла злоумышленника, в секунду (*рисунок 2*).

4. Прогнозирование сетевых атак на основе анализа временного ряда

В качестве метода статистического анализа выбран метод, основанный на анализе временных рядов. По данным ежегодного отчета Cisco по кибербезопасности за 2018 год [11] для обнаружения действий злоумышленника 39% организаций делают ставку на автоматизацию, а остальные используют машинное самообучение (искусственный интеллект) [12–15].

Таблица 2.

Процентное соотношение количества группы пакетов по размеру от общего количества пакетов при благоприятном сетевом трафике и при сетевой атаке

№	Размер пакета	Процентное соотношение количества группы пакетов по размеру от общего количества пакетов	
		при благоприятном сетевом трафике	при сетевой атаке
1.	0–19	0,00%	0,00%
2.	20–39	0,00%	0,00%
3.	40–79	39,06%	92,79%
4.	80–159	3,81%	0,48%
5.	160–319	0,93%	3,30%
6.	320–639	1,35%	3,25%
7.	640–1279	11,05%	0,16%
8.	1280–2559	42,90%	0,02%
9.	2560–5119	0,82%	0,00%
10.	5120 и более	0,08%	0,00%

Задача прогнозирования сетевых атак решалась на основе аддитивной модели временного ряда. Эта модель предполагает, что каждый уровень временного ряда (F) может быть представлен как сумма трендовой (T), сезонной (S) и случайной (E) компонент:

$$F = T + S + E. \quad (1)$$

Для определения трендовой компоненты была применена линейная регрессия, имеющая вид:

$$y = a \cdot x + b, \quad (2)$$

где y – значение тренда;

x – лаг;

a и b – коэффициенты регрессии.

В формуле (2) коэффициенты a и b определяются предыдущими значениями исходной выборки по формулам:

$$b = \frac{\sum (x - \overline{x})(y - \overline{y})}{\sum (x - \overline{x})^2}, \quad (3)$$

$$a = y - b \cdot x, \quad (4)$$

где $\overline{x}$ – среднее значение лага;

$\overline{y}$ – среднее значение исходной выборки.

На *рисунке 3* представлены исходные данные о количестве сетевых пакетов, приходящих от узла злоумышленника вместе с линией тренда, полученной по формуле (2), где $a = 0{,}973$, $b = 615{,}87$. Линия тренда носит возрастающий характер, связанный с увеличением интенсивности сетевого потока.

Далее необходимо определить сезонную компоненту, которая носит периодический характер, определяемый по автокорреляционной функции (АКФ). *На рисунке 4* изображен график зависимости значений автокорреляционной функции от номера лага. Пунктирной линией обозначается уровень «белого шума» (граница статистической значимости коэффициентов корреляции – ошибка автокорреляционной функции). Значения автокорреляционной функции были рассчитаны до тридцатого лага.

Анализ значений автокорреляционной функции показал, что исходные данные обладают периодичностью изменений значений. Наблюдается высокая положительная корреляционная зависимость в лагах под номерами 22 и 23. Следовательно, при определении сезонной компоненты аддитивной модели прогнозирования периодичность будет составлять примерно в 23 лага. Таким образом, длительность одного сезона составляет $N = 23$ лага (номер лага принимает значения $n = 1, 2, ..., N$), где за один лаг принята одна секунда.

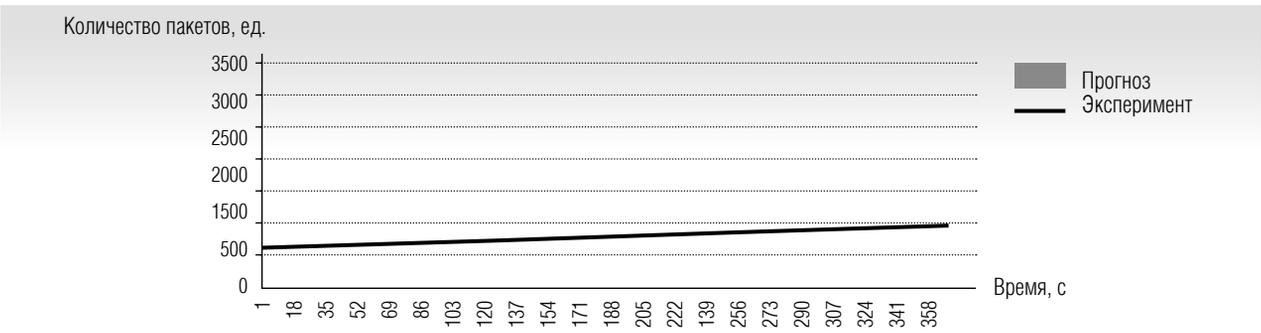

Рис. 3. Количество сетевых пакетов, приходящих от узла злоумышленника, в секунду, с учетом трендовой компоненты

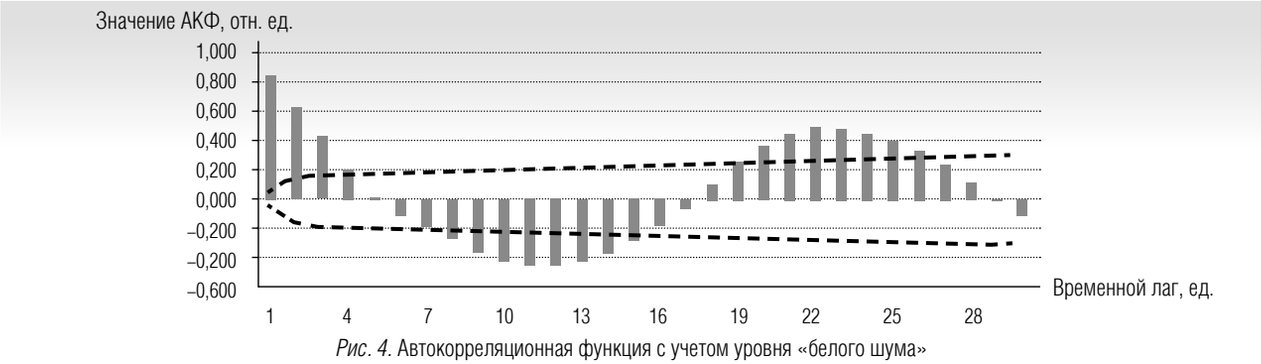

Рис. 4. Автокорреляционная функция с учетом уровня «белого шума»

Значения сезонной компоненты S_n определяются как средние значения разности между действующим значением F_n и трендовой компонентой T_n, рассчитанные для каждого номера лага n:

$$S_n = \sum_{k=1}^{K} \frac{(F_n)_k - (T_n)_k}{K}, \quad (5)$$

где k – номер сезона;

K – количество сезонов.

Тогда общее количество лагов во всем временном ряде равно $M = N \cdot K$.

Получив значения трендовой (2) и сезонной (5) компонент, можно построить прогнозные значения F по формуле (1) (в данной модели случайная компонента не учитывается) [16]. На *рисунке 5* изображен график значений тестовой выборки F_n и прогнозных значений F. Различия между графиками F и F_n можно оценить с помощью средней абсолютной ошибки в процентах (mean absolute percentage error, MAPE).

Для описания погрешности модели прогнозирования, используемой в работе, такая оценка не применяется, поскольку в действующих значениях тестовой выборки присутствуют значения, близкие к единице. В нашем случае в качестве оценки погрешности используется среднеквадратическая ошибка (root mean square error, RMSE), значение которой равно 353. Эта оценка определяется по следующей формуле:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum (y - \hat{y})^2}, \quad (6)$$

где N – объем исходной выборки;

y – прогнозное значение,

$\hat{y}$ – действующее значение.

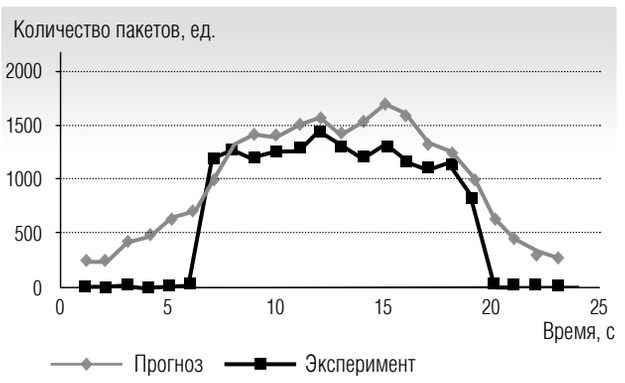

Рис. 5. Действующие значения тестовой выборки и прогнозные значения количества сетевых пакетов в секунду

Таблица 3.

Оценка трендовой и сезонной компонент

№	Действующие значения, количество пакетов в секунду	Оценка трендовой компоненты, количество пакетов в секунду	Оценка сезонной компоненты, количество пакетов в секунду
1.	65	661	–596
2.	21	662	–641
3.	9	663	–654
4.	18	663	–645
5.	1088	664	424
6.	1398	665	733
7.	1301	666	635
8.	1363	667	696
9.	1343	668	675
10.	1375	669	706
11.	1283	670	613
12.	1378	671	707
13.	1387	672	715
14.	1304	673	631
15.	1276	674	602
16.	1302	675	627
17.	1295	676	619
18.	1380	677	703
19.	1391	678	713
20.	1062	679	383
21.	15	679	–664
22.	23	680	–657
23.	11	681	–670
24.	10	682	–672
25.	19	683	–664
26.	24	684	–660
27.	13	685	–672
28.	36	686	–650
29.	36	687	–651
30.	1313	688	625
31.	1342	689	653
32.	1360	690	670
33.	1439	691	748
34.	1380	692	688
35.	1290	693	597
36.	1384	694	690
37.	1329	695	634
38.	1306	695	611
39.	1315	696	619
40.	1296	697	599
41.	1309	698	611
42.	1298	699	599
43.	93	700	–607
44.	37	701	–664
45.	21	702	–681
46.	9	703	–694

Полученное значение свидетельствует о том, что модель прогнозирования не оптимальна. Для более точного прогноза в качестве компонент аддитивной модели прогнозирования (трендовой и сезонной) выбраны компоненты предыдущих (ближайших по времени) сезонов, представленные в *таблице 3* и на *рисунке 6*.

Рис. 6. Количество сетевых пакетов, приходящих от узла злоумышленника, в секунду, с учетом трендовой компоненты ближайших двух сезонов исходной выборки

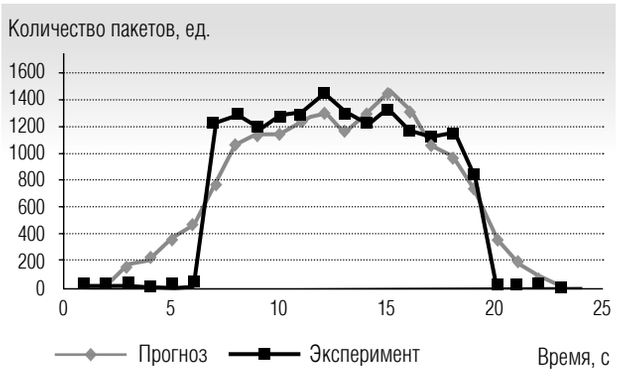

Рис. 7. Действующие значения тестовой выборки и прогнозные значения количества сетевых пакетов в секунду, с внесением изменений

Амплитуда и длительность этих сезонов наиболее близка к последующему сезону, что может повысить качество прогнозирования.

На *рисунке 7* представлены графики последующих действующих значений тестовой выборки и прогнозных значений с внесением изменений в компонентах аддитивной модели временных рядов.

В этом случае оценка погрешности модели прогнозирования RMSE составляет 201, что существенно лучше предыдущего значения оценки RMSE, равной 353. Из этого можно сделать вывод, что модель прогнозирования сетевых атак заметно эффективнее, если она строится не на всей выборке значений, а на основе нескольких относительно недавних экспериментальных данных. Для предоставления относительной оценки погрешности модели прогнозирования можно рассчитать процентное соотношение оценки RMSE и максимального значения тестовой выборки. Здесь вместо среднего значения тестовой выборки выбрано максимальное, поскольку в рассматриваемой выборке присутствует большое количество значений, численно близких к единице. Это приводит к относительно малому среднему значению и не позволяет корректно определить оценку относительной погрешности (MAPE). В итоге процентное отношение оценки RMSE к максимальному значению тестовой выборки составляет 13%.

Таким образом, построенная модель прогнозирования неблагоприятного сетевого трафика обладает относительно низкой погрешностью, что позволяет использовать ее в задачах обнаружения сетевых атак.

В случае необходимости построенную модель прогнозирования DDOS-атак можно применить в исследованиях общей динамики числа DDOS-атак, производимых во всем интернет-пространстве [17]. Взяв в качестве эмпирических предыдущие значения числа совершаемых DDOS-атак за временные кварталы 2017 и 2018 годов, можно определить прогнозные значения числа совершаемых DDOS-атак за первое полугодие 2019 года.

Анализ данных, представленных на *рисунке 8*, показывает, что в поведении динамики числа совершаемых DDOS-атак наблюдаются две периодичности – с периодами порядка 60 и 7 дней. По всей видимости, максимумы активности (15.02.2019, 10.04.2019 и 5.06.2019) с большим периодом находятся между относительно длинными праздничными днями (мартовские, майские и июньские праздники). Короткие периодичные всплески, по-видимому, определяются с активностью в определенные дни недели. Таким образом, достаточно простая модель прогнозирования позволяет обнаружить связь между периодичностью DDOS-атак и особенностями календаря 2019 года.

Также следует отметить, что эффективность исследуемой модели прогнозирования повысится, если DDOS-атаки по статистическим показателям практически идентичны. Если каждая из реализаций DDOS-атак не похожа друг на друга по статистическим показателям, то обнаружить и спрогнозировать действия злоумышленника будет сложнее.

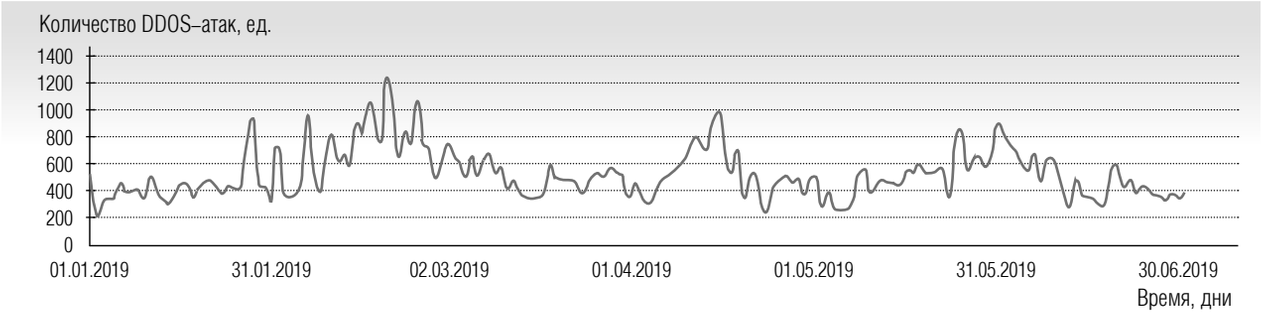

Рис. 8. Прогнозные значения количества DDOS–атак

Заключение

В статье изложены результаты анализа сетевого трафика в задаче прогнозирования угроз информационной безопасности облачных сервисов. Представлена статистика об угрозах информационной безопасности средств хранения и передачи информации, подтверждающая актуальность применения современных методик защиты информации. Подобные методики для анализа угроз компьютерной безопасности предполагают использование специальных программно-аппаратных средств. В нашем случае рассмотрена вредоносная программа, воспроизводящая сетевую атаку, и программа, предназначенная для получения и обработки необходимых для исследования эмпирических показателей. Смоделирован процесс сетевой атаки (DDOS-атаки) и зафиксированы необходимые показатели в файлах, удобных для анализа и дальнейшей обработки. Из множества моделей прогнозирования выбрана аддитивная модель временных рядов. Результаты применения этой модели показали, что, зная характер изменения статистических показателей всевозможных реализаций DDOS-атак, можно выделять, а затем и прогнозировать действия злоумышленника, совершающего такого рода сетевые атаки. Показана эффективность рассмотренной модели прогнозирования путем сравнения полученных прогнозных значений и последующих действующих значений. Получена количественная оценка погрешности построенной модели прогнозирования в виде оценки RMSE, которая составляет 201. Результаты исследований показывают, что статистическая методика может быть применена в средствах защиты облачных сервисов от различных сетевых атак для анализа сетевого трафика. ■

Литература

1. Максимов К.В. Планирование деятельности ИТ-компании в условиях неопределенности с учетом использования облачных сервисов // Прикладная информатика. 2018. Т. 13. № 1 (73). С. 25–31.
2. Тумбинская М.В. Обеспечение защиты от нежелательной информации в социальных сетях // Вестник Мордовского университета. 2017. Т. 27. № 2. С. 264–288.
3. Weins K. Cloud computing trends: 2017 state of the cloud survey. [Электронный ресурс]: https://www.rightscale.com/blog/cloud-industry-insights/cloud-computing-trends-2017-state-cloud-survey (дата обращения 10.10.2018).
4. Крупин А. Идем в облака: российская премьера IaaS-платформы Flexible Computing Express. [Электронный ресурс]: https://servernews.ru/813983?k292300 (дата обращения 10.10.2018).
5. Tumbinskaya M.V. Process of distribution of undesirable information in social networks // Business Informatics. 2017. No 3. P. 65–76.
6. Угрозы безопасности в облаке. [Электронный ресурс]: http://tadviser.ru/a/170054 (дата обращения 10.10.2018).
7. Baranov P.A., Beybutov E.R. Securing information resources using web application firewalls // Business Informatics. 2015. No 4 (34). P. 71–78.
8. Garcia S., Grill M., Stiborek J., Zunino A. An empirical comparison of botnet detectionmethods // Computers and Security. 2014. No 45. P. 100–123.
9. Косенко М.Ю., Мельников А.В. Вопросы обеспечения защиты информационных систем от ботнет атак // Вопросы кибербезопасности. 2016. № 4 (17). С. 20–28.
10. Гнеушев В.А., Кравец А.Г., Козунова С.С., Бабенко А.А. Моделирование сетевых атак злоумышленников в корпоративной информационной системе // Промышленные АСУ и контроллеры. 2017. № 6. С. 51–60.
11. Ежегодный отчет Cisco по кибербезопасности за 2018 г. [Электронный ресурс]: https://www.cisco.com/c/ru_ru/about/press/press-releases/2018/03-12.html (дата обращения 10.10.2018).
12. Glushenko S.A. An adaptive neuro-fuzzy inference system for assessment of risks to an organization's information security // Business Informatics. 2017. No 1 (39). P. 68–77.
13. Картиев С.Б., Курейчик В.М. Алгоритм классификации, основанный на принципах случайного леса, для решения задачи прогнозирования // Программные продукты и системы. 2016. № 2. С. 11–15.

14. Afanas'ev A.P., Dzyuba S.M., Emelyanova I.I. Horner's Scheme for investigation of solutions of differential equations with polynomial right-hand side // Business Informatics. 2017. No 2 (40). P. 33–39.
15. Tomilin A., Tumbinskaya M., Tregubov V., Smolevitskaya M. The BESM-6 virtualization project // 2017 Fourth International Conference on Computer Technology in Russia and in the Former Soviet Union (SoRuCom 2017). Moscow, 3–5 October 2017. P. 241–245.
16. Певцова Т.А., Рябухина Е.А., Гущина О.А. Вычисление индекса сезонности // Вестник Мордовского университета. 2015. Т. 25. № 4. С. 18–36.
17. Купреев О., Бадовская Е., Гутников А. DDoS-атаки в третьем квартале 2018 года. [Электронный ресурс]: https://securelist.ru/ddos-report-in-q3-2018/92512/ (дата обращения 12.11.2018).

Об авторах

Тумбинская Марина Владимировна
кандидат технических наук;
доцент кафедры систем информационной безопасности,
Казанский национальный исследовательский технический университет им. А.Н. Туполева, 420111, г. Казань, ул. К. Маркса, д. 10;
E-mail: tumbinskaya@inbox.ru

Баянов Булат Ильмирович
студент, Казанский национальный исследовательский технический университет им. А.Н. Туполева,
420111, г. Казань, ул. К. Маркса, д. 10;
E-mail: bayanov_bulat@mail.ru

Рахимов Руслан Жамилович
студент, Казанский национальный исследовательский технический университет им. А.Н. Туполева,
420111, г. Казань, ул. К. Маркса, д. 10;
E-mail: rahimov96@mail.ru

Кормильцев Никита Вячеславович
студент, Казанский национальный исследовательский технический университет им. А.Н. Туполева,
420111, г. Казань, ул. К. Маркса, д. 10;
E-mail: kormiltcev@hotmail.com

Уваров Александр Дмитриевич
студент, Казанский национальный исследовательский технический университет им. А.Н. Туполева,
420111, г. Казань, ул. К. Маркса, д. 10;
E-mail: obg-96@mail.ru

Analysis and forecast of undesirable cloud services traffic

Marina V. Tumbinskaya
E-mail: tumbinskaya@inbox.ru

Bulat I. Bayanov
E-mail: bayanov_bulat@mail.ru

Ruslan Zh. Rakhimov
E-mail: rahimov96@mail.ru

Nikita V. Kormiltcev
E-mail: kormiltcev@hotmail.com

Alexander D. Uvarov
E-mail: obg-96@mail.ru

Kazan National Research Technical University named after A.N. Tupolev
Address: 10, Karl Marx Street, Kazan 420111 Russia

Abstract

These days one of the main problems that must be solved to ensure information security in cloud services for corporations as well as for individual clients is to correctly identify and predict hacking in the network traffic. This paper presents statistics on information security threats, provides classification of information security threats for cloud services, identifies hackers' goals, and proposes countermeasures.

A vital task is to develop an effective method that could be used to protect cloud services from various network threats, as well as to analyze the network traffic. For these purposes, we chose a method based on an additive time series model, which allows us to predict the undesirable network traffic. To test this method, we obtained quantitative parameters for the undesirable traffic by simulating a network attack and collecting empirical data that describe this process. We used special software that simulates a network attack, and software that records and processes all the empirical data needed for the research.

Using the data obtained, we analyzed the efficiency of the method based on the additive time series model. We demonstrated that this method is also applicable for research into the general dynamics of the number of network attacks in cyberspace. This method also allows us to reveal how the dynamics of the number of hacker network attacks depends on season, date, or time. The results show that, based on data describing the network traffic, one can identify and predict the undesirable hacker threats.

Key words: forecasting; DDOS attack; cloud services; network traffic; modeling; additive time series model; autocorrelation function; error estimation.

Citation: Tumbinskaya M.V., Bayanov B.I., Rakhimov R.Zh., Kormiltcev N.V., Uvarov A.D. (2019) Analysis and forecast of undesirable cloud services traffic. *Business Informatics*, vol. 13, no 1, pp. 71–81

DOI: 10.17323/1998-0663.2019.1.71.81

References

1. Maksimov K.V. (2018) Planning of activities in the IT-company in conditions of uncertainty taking into account the use of cloud services *Applied Informatics*, vol. 13, no 1, pp. 25–31 (in Russian).
2. Tumbinskaya M.V. (2017) Providing protection from targeted information in social networks. *Mordovia University Bulletin*, vol. 27, no 2, pp. 264–288 (in Russian).
3. Weins K. (2017) *Cloud computing trends: 2017 state of the cloud survey*. Available at: https://www.rightscale.com/blog/cloud-industry-insights/cloud-computing-trends-2017-state-cloud-survey (accessed 10 October 2018).
4. Krupin A. (2014) *We go into the clouds: Russian premiere of the IaaS-platform Flexible Computing Express.* Available at: https://servernews.ru/813983?k292300 (accessed 10 October 2018) (in Russian).
5. Tumbinskaya M.V. (2017) Process of distribution of undesirable information in social networks. *Business Informatics*, no 3, pp. 65–76.
6. Tadviser (2018) *Security threats in the cloud*. Available at: http://tadviser.ru/a/170054 (accessed 10 October 2018) (in Russian).
7. Baranov P.A., Beybutov E.R. (2015) Securing information resources using web application firewalls. *Business Informatics*, no 4, pp. 71–78.
8. Garcia S., Grill M., Stiborek J., Zunino A. (2014) An empirical comparison of botnet detectionmethods. *Computers and Security*, no 45, pp. 100–123.
9. Kosenko M.Yu., Melnikov A.V. (2016) Issues of protecting business information systems from botnets attacks. *Voprosy Kiberbezopasnosti*, no 4, pp. 20–28 (in Russian).
10. Gneushev V.A., Kravets A.G., Kozunova S.S., Babenko A.A. (2017) Modeling network attacks of attackers in a corporate information system. *Industrial Automatic Control Systems and Controllers*, no 6, pp. 51–60 (in Russian).
11. Cisco (2018) *Cisco annual report on cybersecurity for 2018.* Available at: https://www.cisco.com/c/ru_ru/about/press/press-releases/2018/03-12.html (accessed 10 October 2018) (in Russian).
12. Glushenko S.A. (2017) An adaptive neuro-fuzzy inference system for assessment of risks to an organization's information security. *Business Informatics*, no 1, pp. 68–77.
13. Kartiev S.B., Kureichik V.M. (2016) Classification algorithm based on random forest principles for forecasting problem. *Software & Systems (Programmnye produkty i sistemy)*, no 2, pp. 11–15 (in Russian).
14. Afanas'ev A.P., Dzyuba S.M., Emelyanova I.I. (2017) Horner's Scheme for investigation of solutions of differential equations with polynomial right-hand side. *Business Informatics*, no 2, pp. 33–39.
15. Tomilin A., Tumbinskaya M., Tregubov V., Smolevitskaya M. (2017) The BESM-6 virtualization project. Proceedings of the *2017 Fourth International Conference on Computer Technology in Russia and in the Former Soviet Union (SoRuCom 2017). Moscow, 3–5 October 2017*, pp. 241–245.
16. Pevtsova T.A., Ryabukhina E.A., Gushchina O.A. (2015) Calculation of seasonality index. *Mordovia University Bulletin*, vol. 25, no 4, pp. 18–36 (in Russian).
17. Kupreev O., Badovskaya E., Gutnikov A. (2018) *DDoS attacks in the third quarter 2018.* Available at: https://securelist.ru/ddos-report-in-q3-2018/92512/ (accessed 12 November 2018) (in Russian).

About the authors

Marina V. Tumbinskaya
Cand. Sci. (Tech.);
Associate Professor, Department of Information Protection Systems,
Kazan National Research Technical University named after A.N. Tupolev, 10, Karl Marx Street, Kazan 420111, Russia;
E-mail: tumbinskaya@inbox.ru

Bulat I. Bayanov
Student, Kazan National Research Technical University named after A.N. Tupolev, 10, Karl Marx Street, Kazan 420111, Russia;
E-mail: bayanov_bulat@mail.ru

Ruslan Zh. Rakhimov
Student, Kazan National Research Technical University named after A.N. Tupolev, 10, Karl Marx Street, Kazan 420111, Russia;
E-mail: rahimov96@mail.ru

Nikita V. Kormiltcev
Student, Kazan National Research Technical University named after A.N. Tupolev, 10, Karl Marx Street, Kazan 420111, Russia;
E-mail: kormiltcev@hotmail.com

Alexander D. Uvarov
Student, Kazan National Research Technical University named after A.N. Tupolev, 10, Karl Marx Street, Kazan 420111, Russia;
E-mail: obg-96@mail.ru