РЕЗУЛЬТАТ ИССЛЕДОВАНИЯ АЛГОРИТМОВ ВЫЯВЛЕНИЯ СЕТЕВЫХ АНОМАЛИЙ

***Микова Софья Юрьевна**, студентка кафедры информационной безопасности ФГАОУ ВО «Волгоградский государственный университет», г. Волгоград*
E-mail: mikova.s@yandex.ru
***Оладько Владлена Сергеевна**, кандидат технических наук, доцент кафедры информационной безопасности ФГАОУ ВО «Волгоградский государственный университет», кафедра информационной безопасности г. Волгоград*
E-mail: oladko.vs@yandex.ru

В статье рассмотрена проблема обнаружения сетевых аномалий. Исследованы алгоритмы: на основе дискретного вейвлет–преобразования с применением статистических критериев и обнаружения аномалий Бродского–Дарховского. Предложены критерии, позволяющие проанализировать результаты работы каждого алгоритма и оценить его точность. Приведено формализованное описание процедуры выбора наиболее точного алгоритма выявления сетевых аномалий. В результате проведения экспериментов и оценки критериев был найден алгоритм, имеющий наименьшее количество ошибок 1 и 2 рода и наибольшее количество правильно обнаруженных аномалий.

***Ключевые слова**: атака, информационная система, злоумышленник, ошибки первого рода, ошибки второго рода, размер окна, точность.*

THE RESULT OF THE RESEARCH OF ALGORITHMS TO IDENTIFY NETWORK ANOMALIES

***Sof'ya Mikova**, student Volgograd State University, Department of Information security, Volgograd*
E-mail: mikova.s@yandex.ru
***Vladlena Oladko**, Ph.D, Associate Professor of Information Security, Volgograd State University, Volgograd*
E-mail: oladko.vs@yandex.ru

The article deals with the problem of detection of network anomalies. We investigated algorithms: the algorithm based on discrete wavelet transform using statistical criteria and anomaly detection algorithm Brodsky–Darhovsky. We found criteria to analyze the performance of each algorithm. Powered rule of evaluation criteria for identifying the most accurate algorithm to identify network anomalies. As a result of 3 experimentation and evaluation criteria was found algorithm having the least amount of errors of 1 and 2 kind and the largest number of correctly detected anomalies.

***Keywords**: attacks, the information system, the attacker, false negative, false positive, the window size.*

Введение

Проблема обнаружения сетевых аномалий актуальна в настоящее время, так как их действие может привести к утечке или искажению данных, обрабатываемых в информационных системах организаций и в технологических системах предприятия. Часто аномалия в сети - это один из признаков атаки злоумышленника. В соответствии с [1] основными причинами возникновения аномалий в сети являются:

– действия злоумышленников;
– действия и ошибки некомпетентных пользователей;
– неисправность аппаратного обеспечения;
– дефекты и ошибки программного обеспечения.

Существует много алгоритмов обнаружения аномалий в сети, основными из которых являются:

1) алгоритм на основе дискретного вейвлет-преобразования;

2) алгоритм Бродского-Дарховского;

3)алгоритм на основе сумме квадратов вейвлет-коэффициентов;

4) алгоритм на основе максимума квадратов вейвлет-коэффициентов.

Из них наиболее простыми в реализации являются: алгоритм на основе дискретного вейвлет-преобразования с применением статистических критериев (ДВП) и алгоритм обнаружения аномалий Бродского-Дарховского (БД). [4,5]

А поскольку алгоритмы обнаружения аномалий часто используются при диагностике атак на корпоративные сети и информационные системы предприятий, то одним из важных показателей качестве алгоритма будет являться его точность.

В соответствии с [3] для описания точности работы алгоритмов могут использоваться следующие показатели:

1)ошибки первого рода - E1;

2) ошибки второго рода - E2 ;

3) размеры окон – W1 и W2;

4) количество правильно обнаруженных аномалий – S.

Формализация процедуры оценки точности алгоритмов сетевых аномалий

Формализованная процедура оценки точности и выбора лучшего алгоритма может быть описана следующей последовательностью шагов.

1) Для нормализованной ранговой оценки значений W1, W2, S, E1, E2 каждого алгоритма вводится множество критериев оценки алгоритмов $K_i \in K | K \in [0,3], i = 1..5$.

2) Каждому критерию оценки $K_i \in K$ присваивается ранговое значение в соответствии с правилами, описанными формулами 1-4.

Критерии $K_{1,2}$- описывают размеры окон W1, W2 в алгоритмах обнаружения сетевых аномалий. При чем размер окна W2 применяется только для работы алгоритма дискретного вейвлет-преобразования с применением статистических критериев. [3]

$$K_{1,2}=\begin{cases} 3, \text{если } W_{1,2} \leqslant 10; \\ 2, \text{если } 10 < W_{1,2} \leq 19 \\ 1, \text{если } 19 < W_{1,2} < 30 \\ W_{1,2} \geqslant 30, \ \text{rang}K_{1,2} = 0. \end{cases} \quad (1)$$

Критерий K_3 позволяет оценить количество верно идентифицированных алгоритмом сетевых аномалий.

$$K_3=\begin{cases} 3, \text{если } \frac{P}{N} = 1 \\ 2, \text{если } 3 < \frac{P}{N} < 1 \\ 1, \text{если } 0 < \frac{P}{N} \leqslant 0,3 \\ 0, \text{если } \frac{P}{N} = 0 \end{cases}, \quad (2)$$

где P- количество правильно обнаруженных аномалий; N- общее число аномалий.

Критерий K_4, позволяет оценить количество ошибок первого рода алгоритма при обнаружении им аномалий.

$$K_4=\begin{cases} 3, \text{если } \frac{L}{N} = 0 \\ 2, \text{если } 0 < \frac{L}{N} < 0,5 \\ 1, \text{если } 0,5 \leqslant \frac{L}{N} < 1 \\ 0, \text{если } \frac{L}{N} = 1 \end{cases}, \quad (3)$$

где L- количество ложных тревог, N- общее число аномалий.

Критерий K_5, позволяет оценить количество ошибок второго рода, которые появляются в процессе работы алгоритма.

$$K_5=\begin{cases} 3, \text{если } \frac{H}{N} = 0 \\ 2, \text{если } 0 < \frac{H}{N} < 0,5 \\ 1, \text{если } 0,5 \leqslant \frac{H}{N} < 1 \\ 0, \text{если } \frac{H}{N} = 1 \end{cases} \quad (4)$$

3) Для каждого анализируемого алгоритма j={ДВП, БД} по формуле 5 вычисляется комплексный показатель K_{jsum} . Более точным будет считаться тот алгоритм обнаружения сетевых аномалий, который имеет наибольшую комплексную оценку.

$$K_{jsum} = \sum_{i=1}^{5} K_i \quad (5)$$

4) Для получения более достоверной оценки рекомендуется провести m оценок алгоритмов и для полученного ряда значений $K_{jsum}^L = \{K_{jsum}^1, \dots, K_{jsum}^m\}$ каждого алгоритма рассчитать математическое ожидание $M[K_{jsum}^L]$. В этом случае, как и в шаге 3, наиболее точным признается тот алгоритм, математическое ожидание которого имеет максимальную оценку.

Экспериментальные исследования алгоритмов обнаружения сетевых аномалий Бродского-Дарховского и дискретного вейвлет-преобразования с применением статистических критериев

Для проведения исследований точности алгоритмов был разработан программный комплекс, с помощью которого было проведено множество экспериментов, направленных:

на оценку влияния зависимостей входных параметров алгоритмов БД и ДВП на точность обнаружения аномалий;

на выбор наиболее точного в своей работе алгоритма обнаружения сетевых аномалий.

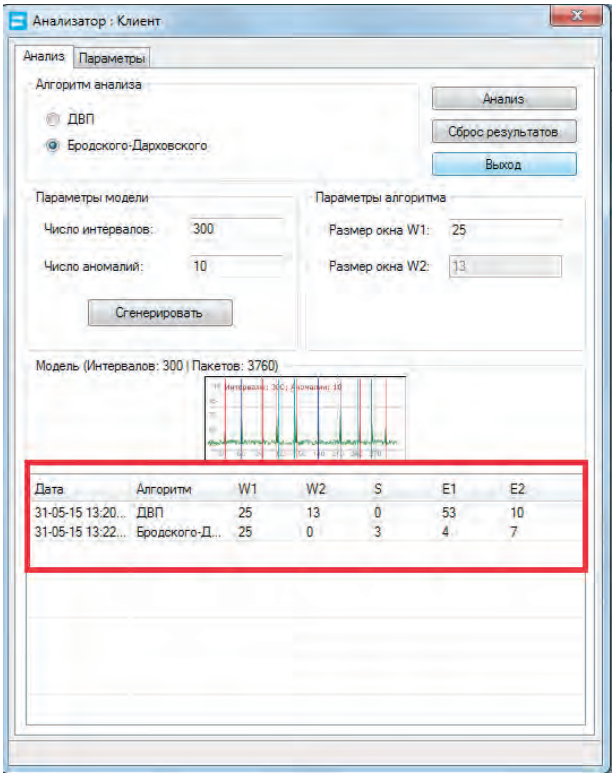

Рис. 1. *Результат работы алгоритмов.*

Экспериментальные исследования алгоритмов обнаружения сетевых аномалий ДВП и БД состояли из трех экспериментов, входные данные которых, одинаковые для каждого из двух исследуемых алгоритмов представлены в таблице 1.

Таблица 1*. Входные данные экспериментов*

Номер эксперимента	1	2	4
Входные данные эксперимента	I= 500 N= 4 W_1=30 W_2=10	I= 300 N= 10 W_1=25 W_2=13	I= 450 N= 15 W_1=10 W_2=5

где W_1, W_2 -размеры окон, I- количество интервалов, N- число аномалий.

Каждый эксперимент включал в себя исследование работы обоих алгоритмов. Пример проведения одного из этапов эксперимента с помощью разработанной программы представлен на рисунке 1.

В результате проведения эксперимента для разных входных данных были получены значения, представлены в таблице 2:

Таблица 2*. Результаты работы алгоритмов.*

№ этапа эксперимента	Алгоритмы	W1	W2	S	E1	E2
1 (I= 500; N= 4)	ДВП	30	10	1	0,13	0
	БД	30	0	0,25	0,75	0,75
2 (I= 300; N= 10)	ДВП	25	13	0	5,3	1
	БД	25	0	0,3	0,4	0,7
3 (I= 450; N= 15)	ДВП	10	5	0	0,6	0,8
	БД	10	0	0,4	0,5	0,4

Таблица 3. *Критериальная оценка точности алгоритмов обнаружения сетевых аномалий*

Алгоритм	Эксперимент	K_1	K_2	K_3	K_4	K_5	K_{jsum}
ДВП	1	0	3	3	2	3	11
	2	1	2	0	0	0	3
	3	3	3	0	1	1	8
БД	1	0	3	1	1	1	6
	2	1	3	1	2	1	8
	3	3	3	2	1	2	11

В соответствии с результатами работы алгоритмов была проведена оценка каждого критерия (формулы 1-5), результаты которой представлены в таблице 3 и графиков на рисунке 2.

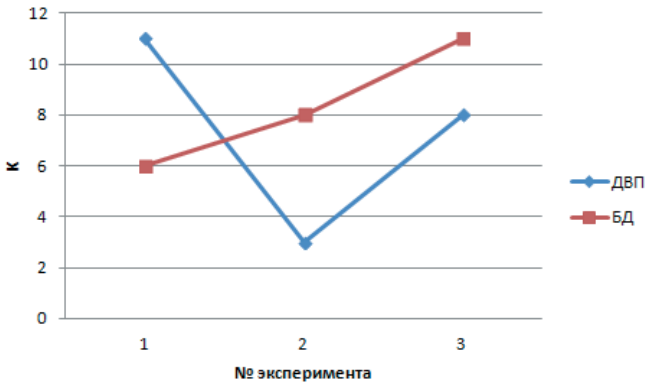

Рис. 2. *Результаты оценки критериев работы алгоритмов.*

Вывод

Из анализа результатов работы алгоритмов следует, что наиболее точным в обнаружении аномалий является алгоритм Бродского-Дарховского поскольку математическое ожидание его обобщённых оценок $M\left[K_{БДsum}^{L=3}\right] = 8{,}33$,на 12% превышает математическое ожидание алгоритма на основе дискретного вейвлет-преобразования с применением статистических критериев $M\left[K_{ДВПsum}^{L=3}\right] = 7{,}33$.Также при его использовании обнаруживается меньше ошибок 1-ого и 2-ого рода, чем при использовании алгоритма на основе дискретного вейвлет-преобразования с применением статистических критериев. Кроме того, алгоритм Бродского-Дарховского имеет наибольшее количество правильно обнаруженных аномалий.

Литература:

1. Жидков И.В., Кадушкин И.В. О признаках потенциально опасных событий в информационных системах//Вопросы кибербезопасности., -2014.-№1(2).-С. 40-48.
2. Басараб М.А., Строганов И.С. Обнаружение аномалий в информационных процессах на основе мультифрактального анализа // Вопросы кибербезопасности., -2014.-№4(7).-С. 30-40.
3. Микова С.Ю. Оладько В.С. Нестеренко М.А., Кузнецов И.А. Критерии оценки качества алгоритмов обнаружения сетевых аномалий// Международный научно-исследовательский журнал, -2015 - №4 (35) –С. 87-88.
4. Шелухин О.И., Филинова А.С. Обнаружение сетевых аномальных выбросов трафика методом разладки Бродского-Дарховского // T-Comm - Телекоммуникации и Транспорт., -2013.-№10 , том 7-С. 116-118.
5. Шелухин О.И., Панкрушин А.П. Оценка достоверности обнаружения аномалий сетевого трафика методами дискретного вейвлет-преобразования // T-Comm - Телекоммуникации и Транспорт., -2013.-№10 , том 7-С. 110-113.

References:

1. Zhidkov I.V. , Kadushkin I.V. O priznakakh potentsial'no opasnykh sobytiy v informatsionnykh sistemakh//Voprosy kiberbezopasnosti , -2014.- №1 (2) .- S . 40-48
2. Basarab M.A. , Stroganov I.S. Obnaruzheniye anomaliy v informatsionnykh protsessakh na osnove mul'tifraktal'nogo analiza // Voprosy kiberbezopasnosti . , -2014. - №4 (7) .- S . 30-40 .
3. Mikova S.YU. Olad'ko V.S. Nesterenko M.A. , Kuznetsov I.A. Kriterii otsenki kachestva algoritmov obnaruzheniya setevykh anomaliy // Mezhdunarodnyy nauchno - issledovatel'skiy zhurnal , -2015 - №4 (35) -S . 87-88 .
4. Shelukhin OI Filinova AS Detection of abnormal network traffic emissions by discord Brodsky-Darhovsky // T-Comm - Telecommunications and Transport., -2013.-№10 (7)-S. 116-118.
5. Shelukhin OI Pankrushina AP Evaluation of reliability of network traffic anomaly detection method of discrete wavelet transform // T-Comm - Telecommunications and Transport., -2013.-№10(7)-S. 110-113.