Discrete & Continuous Models
& Applied Computational Science
ISSN 2658-7149 (online), 2658-4670 (print)

2021, 29 (3) 230–241
http://journals.rudn.ru/miph

Research article

UDC 519.872
DOI: 10.22363/2658-4670-2021-29-3-230-241

Evaluation of firewall performance when ranging a filtration rule set

Anatoly Y. Botvinko[1], **Konstantin E. Samouylov**[1,2]

[1] *Peoples' Friendship University of Russia (RUDN University)*
6, Miklukho-Maklaya St., Moscow, 117198, Russian Federation
[2] *Research Center "Computer Science and Control" of the Russian Academy of Sciences*
44-2, Vavilov St., Moscow, 119333, Russian Federation

(received: August 4, 2021; accepted: September 9, 2021)

This article is a continuation of a number of works devoted to evaluation of probabilistic-temporal characteristics of firewalls when ranging a filtration rule set. This work considers a problem of the decrease in the information flow filtering efficiency. The problem emerged due to the use of a sequential scheme for checking the compliance of packets with the rules, as well as due to heterogeneity and variability of network traffic. The order of rules is non-optimal, and this, in the high-dimensional list, significantly influences the firewall performance and also may cause a considerable time delay and variation in values of packet service time, which is essentially important for the stable functioning of multimedia protocols. One of the ways to prevent decrease in the performance is to range a rule set according to the characteristics of the incoming information flows. In this work, the problems to be solved are: determination and analysis of an average filtering time for the traffic of main transmitting networks; and assessing the effectiveness of ranging the rules. A method for ranging a filtration rule set is proposed, and a queuing system with a complex request service discipline is built. A certain order is used to describe how requests are processed in the system. This order includes the execution of operations with incoming packets and the logical structure of filtration rule set. These are the elements of information flow processing in the firewall. Such level of detailing is not complete, but it is sufficient for creating a model. The QS characteristics are obtained with the help of simulation modelling methods in the Simulink environment of the matrix computing system MATLAB. Based on the analysis of the results obtained, we made conclusions about the possibility of increasing the firewall performance by ranging the filtration rules for those traffic scripts that are close to real ones.

Key words and phrases: firewall, ranging the filtration rules, network traffic, phase service, simulation model, queuing system

1. Introduction

In order to ensure information security of automated systems (AS) that have connections to external untrusted resources, we have to pay attention to the possibility of threats such as violation of confidentiality, integrity and availability of information. A required condition to prevent the threats aimed on violating AS's normal operation is using the firewall technologies [1]–[3].

The main firewall technology is network traffic filtration according to a certain rule set. It is executed at the points of the connection of the AS under protection to external uncontrolled systems and is implemented by using special hardware or software complexes, i.e., firewalls. The firewall filtration rule set is a list of conditions according to which the further transmission of network traffic packets is allowed or denied. The parameters, attributes and characteristics of network traffic flows are usually used to set filtering conditions [4].

The important fact is that the network traffic filtration brings additional time delays during data transmission. High values of the delays during packet filtration can cause packet losses, denials for session initiation and failures in AS's normal work [5], [6].

In works [7]–[13], a great influence of the rule set size and the order of filtration rules in the set on the firewall performance is noted. The influence can be explained by the sequential scheme used to check the packet compliance with the set rules. The maximum decrease in the performance happens while checking the compliance of attributes of packets under filtration with the conditions at the end of the high-dimensional rule set. Defining a rule set that correctly realizes the security policy, but is ineffective in terms of performance, can be considered an error in firewall configuring.

We should also consider that real network traffic has heterogeneity caused by various non-parameterizable factors. This can lead to a decrease in the effectiveness of the static filtration rule set configured initially. One of the ways to prevent the decrease in the performance caused by traffic heterogeneity is to range the rule set according to the incoming traffic characteristics.

Therefore, the task of ranging a rule set in accordance with the characteristics of information flows is not only actual and in demand. This is especially important for the firewalls that ensure information security for the AS with a complex network architecture and large volumes of network traffic. The main goal of this work is to develop a model for evaluating the firewall performance when ranging the filtration rule set.

This paper has the following structure. A method for ranging the filtration rule set is proposed in section 2. In section 3, a model for ranging the rules in the form of a queuing system (QS) with a phase-type service discipline is developed [14]. The results of simulation modelling and firewall performance evaluation for the network traffic script that is close to real are presented in section 4. The Conclusion contains the main aspects of our study.

2. Ranging a filtration rule set for a firewall

By ranging the filtration rule set we mean putting the rules in descending order by their weights in accordance with the evaluation of the characteristics of information flows. We consider that traffic filtration is executed at the network and transmission levels of the standard model for the open system interaction (OSI). According to the generally accepted classification [1]–[3], such firewalls relate to the type of packet filters.

Ranging is executed at discrete moments of time $t_k^- = t_k - 0$ (see the figure 1).

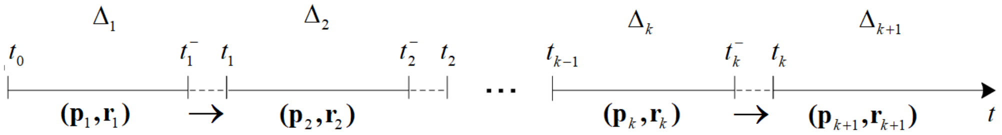
Δ_1
Δ_2
Δ_k
Δ_{k+1}
t_0
t_1^-
t_1
t_2^-
t_2
t_{k-1}
t_k^-
t_k
$(\mathbf{p}_1,\mathbf{r}_1)$
$(\mathbf{p}_2,\mathbf{r}_2)$
$(\mathbf{p}_k,\mathbf{r}_k)$
$(\mathbf{p}_{k+1},\mathbf{r}_{k+1})$
t

Figure 1. Ranging the filtration rule set

The packets received during the time $\Delta_k = t_k - t_{k-1}$, $k \geqslant 1$ are combined into a package, and a group of q packages generates a redundancy errors in the rule set under data segment. It is assumed that there are no inconsistency and consideration [7]–[10]. We also consider that there is a certain minimum number of rules M under which ranging can provide a significant effect. The logical structure of the filtration rules is a linear list of conditions.

Let us introduce the following designations: M is the number of filtration rules in the set r_i^k — the rule in i-th position in the k-th set of filtration rules on the interval $[t_{k-1}, t_k)$, $\mathbf{r}_k = (r_1^k, \dots, r_M^k)$ k-th set of filtration rules on the interval $[t_{k-1}, t_k)$, p_i^k — the weight of r_i^k rule, $\mathbf{p}_k = (p_1^k, \dots, p_M^k)$ — the rule weight vector on the interval $[t_{k-1}, t_k)$.

In this work, a value of the average number of packets, the attributes of which match the conditions of the r_i^k rule on the interval Δ_k is used as a weight p_i^{k+1}.

The nonparametric method of local approximation (MLA) is used to evaluate the average number of requests [15]–[18]. The same method is used for the analysis of other characteristics investigated in this work.

A method for ranging the rules is proposed in the next section of this paper.

3. The model for ranging the filtration rule set

The complexity and variety of the firewall functioning do not allow to create a model reflecting all the regularities and features that are characteristic for various manufacturers, such as Cisco Systems, Juniper Networks, etc. Therefore, the model describes only the main regularities and factors of the firewall functioning that are of interest for our tasks.

For all firewall types in the process of network traffic filtration, the following stages can be distinguished [10]:

— initial packet processing, i.e., operations with a packet when it enters the receiving path;

— checking the filtration rule set;

— completion of packet processing, i.e., operations with a packet when it is transmitted to the output path and then to the physical medium.

During the initial processing of the packet received the firewall network interface controller (NIC) decodes the sequence of electrical or optical signals, checks the correctness of information delivered and writes the packet into the NIC input buffer memory. Then the packet is transmitted to a program buffer located in RAM for operations executed by the central process.

As a next step, if computing resources are available, the filtration of the incoming packet is executed in accordance with the filtration rule set. The compliance of the received packet parameters with the filtration rules is checked in sequence. Only one packet can be checked at a time. The other packets received are kept in the buffer. Their service is executed according to the order of their entrance to the buffer (FCFS, First-Come, First-Served).

If the packet parameters match the filtration rules, the firewall transmits the packet to the NIC output buffer. If the packet parameters do not match the permissive rules, the firewall rejects the packet. The packet processing is considered complete when it is encoded and transmitted to the physical medium.

Let us present the firewall model as a queuing system (QS) with a $B_k(t)$ distribution function (DF) for the request service duration, which depends on the order of the filtration rules on time interval $[t_{k-1}, t_k)$. A request flow $\Lambda(t)$, corresponding to the packet flow incoming the firewall, enters the QS. We consider the incoming packets as the service requests for the QS.

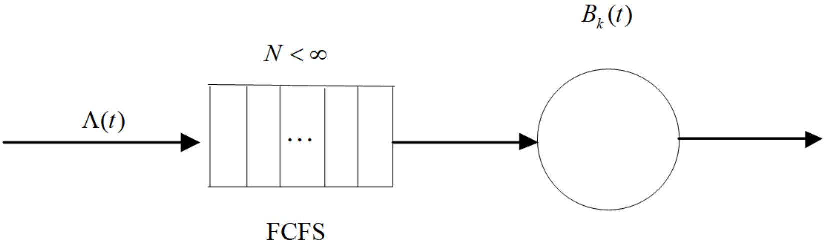

Figure 2. The scheme of the firewall QS with a complex request service discipline

The $B_k(t) = B(\mathbf{r}_k, \mu_0, \mu, t)$ distribution function (DF) is a function of phase type, its parameters are shown in the figure 3, from which it is clear that the $B_k(t)$ DF corresponds to the Cox distribution [19].

The request service time at zero phase corresponds to the total time of packet initial processing and the time of transmission along the output path. The request service time at the m-th phase $m > 1$ corresponds to the time of packet filtration the by the m rule. It is assumed that the filtration time for each rule is the same and equal to τ.

The scheme of the request service process in the firewall model is presented in the figure 3.

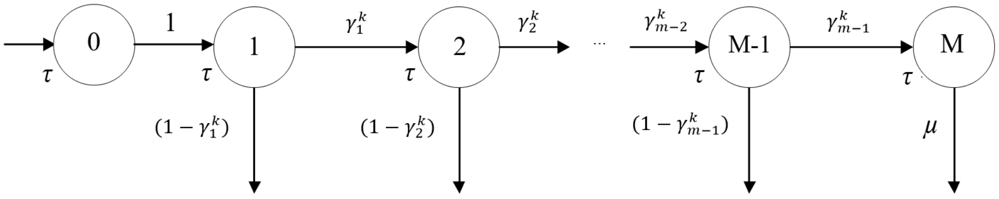

Figure 3. The request service process

The $1 - \gamma_i^k$ value corresponds to the probability of completing the request service at the i-th phase. That is the case when the packet attributes do not correspond to the r_i^k rule. Therefore, the DF of the request service time in the QS on the interval $[t_{k-1}, t_k)$ is as follows:

$$B_k(t) = \gamma_1^k E(1, \mu_0) + \sum_{i=1}^{M} \gamma_1^k E(i, \mu), \tag{1}$$

where $E(1, \mu_0)$ is the Erlang distribution of the i-th order.

The task of analyzing the QS (shown in Fig. 2) characteristics can be solved with the help of the simulation modelling method, the results of which are presented in the next section. It should be noted that in case of a Poisson incoming flow and exponential filtering time, the QS has an analytical solution [14].

4. Evaluation of the firewall performance when ranging the rules

Firewall is a network node processing large volume of incoming and outgoing traffic. Therefore, the average packet filtering time is usually used as the major performance indicator [3], [7]. In this work, to evaluate the firewall performance, we use ΔU_S, i.e., a value equal to the difference between U_1 — the average filtering time in the first data segment (without rule ranging) and U_S — the average filtering time in the S-th data segment (after the rules ranged).

The initial data used for the implementing the simulation model of the process of network traffic filtration are shown in the table 1.

Table 1

Initial data

[rules]	μ_0^{-1} [ms]	μ^{-1} [ms]	$\Delta_k, k = 1, ..., 25$ [ms]	q [packages]	s [segments]
100	$2.7 \cdot 10^{-3}$	$5 \cdot 10^{-5}$	1000	5	5

The number of packet types is M. The values of request service intensities — μ_0 and μ — have been taken from the work [10], which is about the analysis of the firewall performance under the Poisson incoming flow of requests.

To provide the numerical analysis of the QS (see the figure 1), a simulation model (SM) is built in the Simulink simulation environment of the MATLAB matrix computing system with the use of SimEvents discrete state library. The scheme of the model is presented in the figure 4.

The request flow in the SM is determined in the Traffic Generation subsystem. A request collector is realized by the FIFO Queue block, and the request service process is executed by the Single Server blocks (the QS service device) and the `function_f` subsystem (the calculation of the request service time in accordance with the rule set and request type).

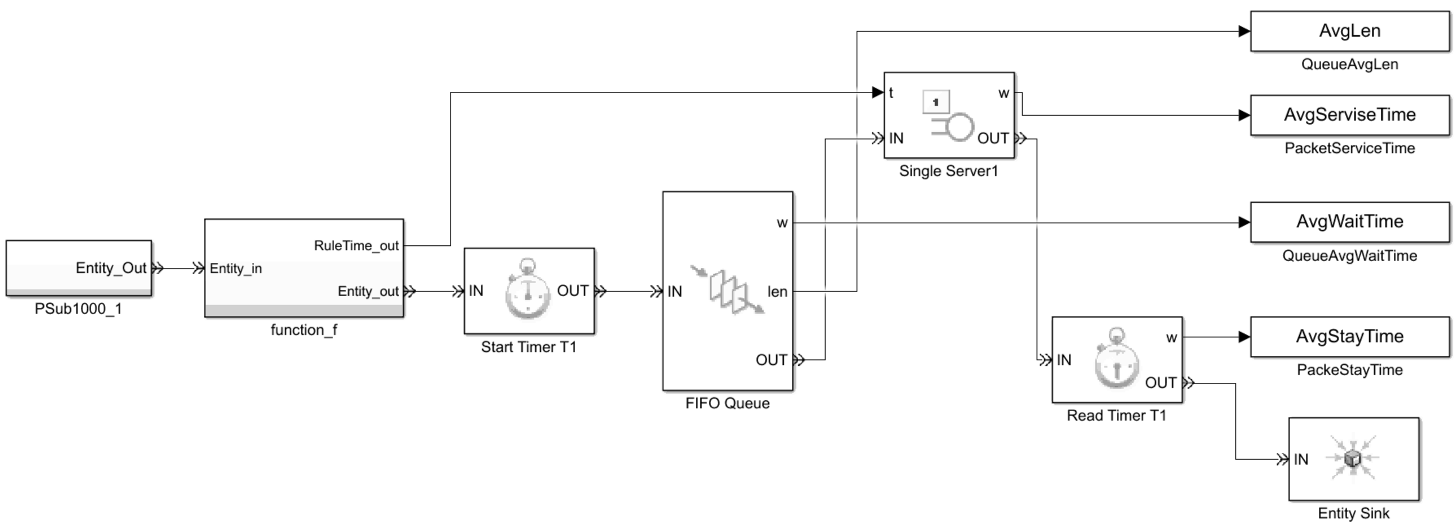

Figure 4. The scheme of the simulation model in the Simulink environment

The statistical data accumulation for evaluating the performance indicators is executed with the help of the statistics collection options of the SimEvents blocks and data recording structures such as `PacketServiceTime`, `QueueAvgWaitTime`, `PackeStayTime`, `QueueAvgLen`. The following times are fixed at this stage: the packet service time, the waiting time for the packet in the queue, the average time of the packet residence in the system, and the average length of the packet queue.

To define the incoming flow of requests, data from the WIDE academic core network in Japan have been used. Traffic records are contained in the MAWI Group Traffic Archive traffic repository by 01/10/2019. For each packet type, using the Wireshark tool for network traffic capture and analysis, the values of the time intervals between packets for the TCP, UDP and ICMP protocols have been extracted. The data massive obtained has been exported to MATLAB to set the intervals between the moments of request generation in the Traffic Generation subsystem using Time-Based Entity Generator blocks. The request types corresponding to the traffic packet types are determined in the Traffic Generation subsystem by the SetPacketAtt blocks. An example of the request flow obtained for packets of r^1_{81} and r^1_{29} types is presented in figures 5–6.

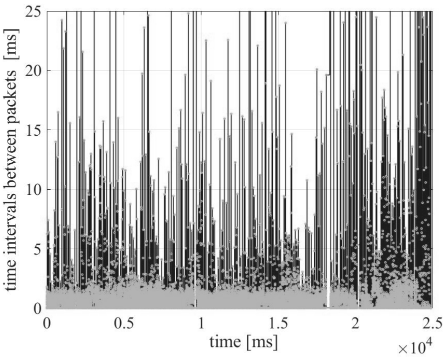

Figure 5. The packet flow of 81st type

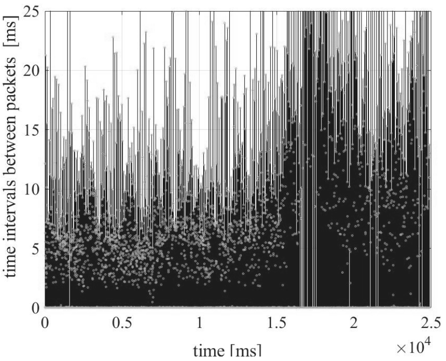

Figure 6. The packet flow of 29th type

The following actions are implemented in M-files of the MATLAB system: determination of the initial data for simulation modelling (see the table 1) and the initial rule set, calculation of the performance indicators, execution of functions for calculating weight, rule set ranging and other algorithms and SM variables.

The process of ranging the r^1_{81} and r^1_{29} filtration rules in accordance with the evaluation of the information flow characteristics is illustrated by figures 7–8. The figures show that:

— the r^1_{81} rule, when ranging, takes the 7th place in the set (average). This can be explained by the short time interval between the packet income (4 ms) and the small value of the time dispersion between the income of the packets;

— the r^1_{29} rule is characterized by moving to the middle of the set. It happens due to the increase in the time interval between the income of the packets.

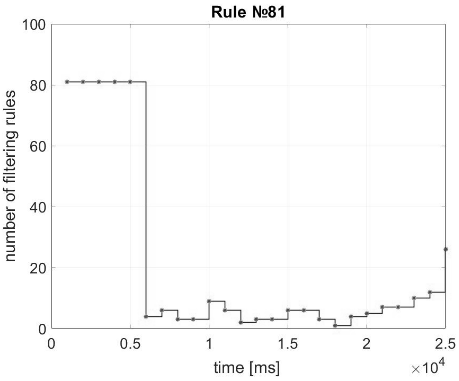

Figure 7. Ranging the r^1_{81} rule

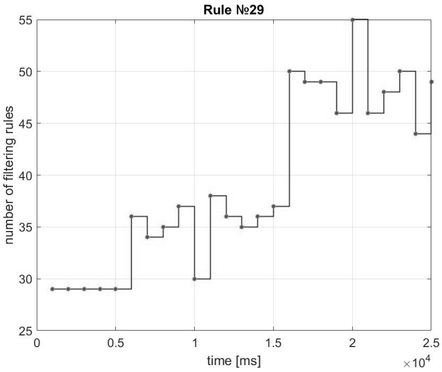

Figure 8. Ranging the r^1_{29} rule

Simulation modelling has demonstrated that the average packet filtration time for all time intervals $[t_{k-1}, t_k) \in T$, $k > 5$ on which ranging has been executed, has a decrease compared to the average time on the intervals $[t_{k-1}, t_k) \in T$, $k = 1, \ldots, 5$.

For the first interval $[t_0, t_1)$, where there is no set ranging, and for the last interval $[t_{24}, t_{25})$, where the set is ranged, we can present the graph of the average packet filtration time (see the figure 9).

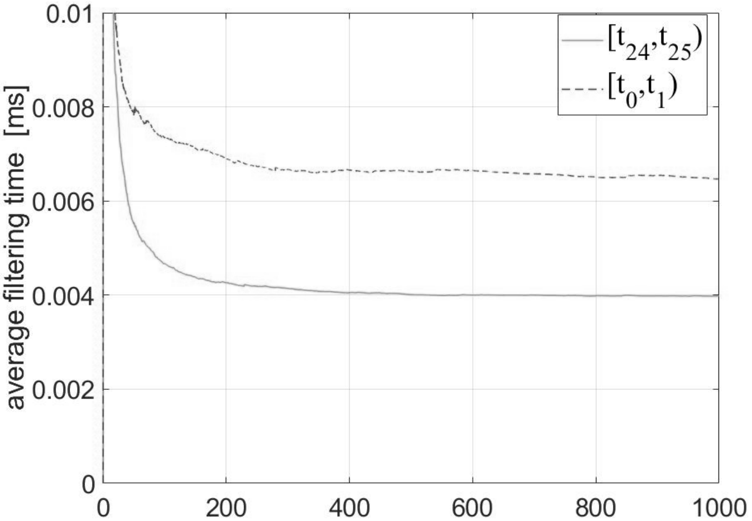

Figure 9. Average packet filtration time for first and the last intervals

As can be seen from the figure, the value of the average packet filtration time on the interval $[t_0, t_1)$ is larger than the average time on the interval $[t_{24}, t_{25})$ by about 2.5 [s]. The results of the firewall performance evaluation for all segments obtained during the simulation modelling are presented in the table 2.

Table 2

The firewall performance

s	U_S [s]	ΔU_S [s]	ΔU_S [%]
1	6.233	-	-
2	4.937	1.296	20.785
3	4.660	1.573	25.229
4	4.989	1.244	19.960
5	4.406	1.827	29.304

5. Conclusion

The created QS with a complex request service discipline and the simulation methods allowed us to obtain the firewall performance estimates when ranging a rule set. These estimates demonstrate that, for the traffic of the main transmission networks, ranging has increased the firewall performance by 20–29% compared to traffic filtering without ranging. So, the results obtained indicate the possibility of increasing the firewall performance for traffic scripts that are close to real ones. These results also confirm the assumptions made in work [20] about the advisability of ranging.

The authors plan to study the influence of the ranging interval and MLA parameters on the firewall performance in further works. They also plan to develop criteria for the need of re-ranging the set depending on changes in the firewall performance indicators, as well as recommendations for ranging the filtration rule sets.

References

[1] S. V. Lebed, *Firewall protection. Theory and practice of external perimeter protection [Mezhsetevoye ekranirovaniye. Teoriya i praktika zashchity vneshnego perimetra]*. Moscow: BMSTU, Bauman Moscow State Technical University Publ., 2002, p. 304, in Russian.

[2] O. R. Laponina, *The foundation of network security [Osnovy setevoy bezopasnosti]*. Moscow: Publishing house of the national Open University «INTUIT», 2014, p. 377, in Russian.

[3] K. V. Ivanov and P. I. Tutubalin, *Markov models of protection of automated control systems for special purposes [Markovskie modeli zashhity' avtomatizirovanny'x sistem upravleniya special'nogo naznacheniya]*. Kazan: Publishing house of GBU Republican center for monitoring the quality of education Publ., 2012, p. 216, in Russian.

[4] "Governing document. Computer aids. Firewall. Protection against unauthorized access to information. Indicators of security against unauthorized access to information [Rukovodyashhij dokument. Sredstva vy'chislitel'noj texniki. Mezhsetevy'e e'krany'. Zashhita ot nesankcionirovannogo dostupa k informacii. Pokazateli zashhishhennosti ot nesankcionirovannogo dostupa k informacii] approved by the decision of the Chairman of the State Technical Commission under the President of the Russian Federation dated July 25, 1997," in Russian.

[5] H. Hamed, A. El-Atawy, and E. Al-Shaer, "On dynamic optimization of packet matching in high-speed firewalls," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 10, pp. 1817–1830, 2006. DOI: 10.1109/JSAC.2006.877140.

[6] R. Mohan, A. Yazidi, B. Feng, and J. Oommen, "On optimizing firewall performance in dynamic networks by invoking a novel swapping window-based paradigm," *International Journal of Communication Systems*, vol. 31, no. 15, e3773, 2018. DOI: 10.1002/dac.3773.

[7] E. Al Shaer, *Automated firewall analytics: Design, configuration and optimization*. Springer International Publishing, 2014, p. 132. DOI: 10.1007/978-3-319-10371-6.

[8] R. Mohan, A. Yazidi, B. Feng, and B. J. Oommen, "Dynamic ordering of firewall rules using a novel swapping window-based paradigm," in *Proceedings 6th International Conference on Communication and Network, ICCNS 2016*, Singapore: ACM Proceedings, 2016, pp. 11–20. DOI: 10.1145/3017971.3017975.

[9] Z. Trabelsi, S. Zeidan, M. M. Masud, and K. Ghoudi, "Statistical dynamic splay tree filters towards multilevel firewall packet filtering enhancement," *Computers & Security*, vol. 53, pp. 109–131, 2015. DOI: 10.1016/j.cose.2015.05.010.

[10] K. Salah, K. Elbadawi, and R. Boutaba, "Performance modeling and analysis of network firewalls," *IEEE Transactions on Network and Service Management*, vol. 9, no. 1, pp. 12–21, 2012. DOI: 10.1109/TNSM.2011.122011.110151.

[11] C. Diekmann, L. Hupel, J. Michaelis, M. Haslbeck, and G. Carle, "Verified iptables firewall analysis and verification," *Journal of Automated Reasoning*, vol. 61, no. 1–4, pp. 191–242, 2018. DOI: 10.1007/s10817-017-9445-1.

[12] S. Khummanee, "The semantics loss tracker of firewall rules," *Advances in Intelligent Systems and Computing*, vol. 769, pp. 220–231, 2018. DOI: 10.1007/978-3-319-93692-5_22.

[13] V. Clincy and H. Shahriar, "Detection of anomaly in firewall rule-sets," *Advances in Intelligent Systems and Computing*, vol. 842, pp. 422–431, 2018. DOI: 10.1007/978-3-319-98776-7_46.

[14] P. P. Bocharov and A. V. Pechenkin, *Queuing theory [Teoriya massovogo obsluzhivaniya]*. Moscow: Publishing RUDN, 1995, p. 529, in Russian.

[15] V. Y. Katkovnik, *Non-parametric data identification and smoothing: local approximation method [Neparametricheskaya identifikatsiya i sglazhivaniye dannykh: metod lokal'noy approksimatsii]*. Moscow: The science. Main editorial office of physical and mathematical literature Publ., 1985, in Russian.

[16] J. M. Bravo, T. Alamo, M. Vasallo, and M. E. Gegúndez, "A general framework for predictors based on bounding techniques and local approximation," *IEEE Transactions on Automatic Control*, vol. 62, no. 7, pp. 3430–3435, 2017. DOI: 10.1109/TAC.2016.2612538.

[17] H. Al-Shuka, "On local approximation-based adaptive control with applications to robotic manipulators and biped robots," *International Journal of Dynamics and Control*, vol. 6, no. 1, pp. 339–353, 2018. DOI: 10.1007/s40435-016-0302-6.

[18] D. E. Plotnikov, T. S. Miklashevich, and S. A. Bartalev, "Using local polynomial approximation within moving window for remote sensing data time-series smoothing and data gaps recovery [Vosstanovleniye vremennykh ryadov dannykh distantsionnykh izmereniy metodom polinomialnoy approksimatsii v skolzyashchem okne peremennogo razmera]," *Modern problems of remote sensing of the Earth from space of the Russian Academy of Sciences*, vol. 11, no. 2, pp. 103–110, 2014, in Russian.

[19] D. R. Cox, "A use of complex probabilities in the theory of stochastic processes," *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 51, no. 2, pp. 313–319, 1955. DOI: 10.1017/S0305004100030231.

[20] A. Y. Botvinko and K. E. Samouylov, "Adaptive ranking of the firewall rule set using local approximation [Adaptivnoye ranzhirovaniye nabora pravil mezhsetevogo ekrana metodom lokal'noy approksimatsii]," in *Distributed Computer and Communication Networks: Control, Computation, Communications*, in Russian, 2018, pp. 334–341.

For citation:

A. Y. Botvinko, K. E. Samouylov, Evaluation of firewall performance when ranging a filtration rule set, Discrete and Continuous Models and Applied Computational Science 29 (3) (2021) 230–241. DOI: 10.22363/2658-4670-2021-29-3-230-241.

Information about the authors:

Botvinko, Anatoly Y. — postgraduate of Department of Applied Probability and Informatics (e-mail: botviay@sci.pfu.edu.ru, ORCID: https://orcid.org/0000-0003-1412-981X, Scopus Author ID: 57222085424)

Samouylov, Konstantin E. — Doctor of Technical Sciences, Professor, Head of Department of Applied Probability and Informatics (e-mail: samuylov-ke@rudn.ru, ORCID: https://orcid.org/000-0002-6368-9680, ResearcherID: E-9966-2014, Scopus Author ID: 14009785000)

УДК 519.872
DOI: 10.22363/2658-4670-2021-29-3-230-241

Оценка производительности межсетевого экрана при ранжировании набора правил фильтрации

А. Ю. Ботвинко[1], **К. Е. Самуйлов**[1, 2]

[1] *Российский университет дружбы народов*
ул. Миклухо-Маклая, д. 6, Москва, 117198, Россия
[2] *Федеральный исследовательский центр «Информатика и управление» РАН*
ул. Вавилова, д. 44, корп. 2, Москва, 119333, Россия

Данная статья является продолжением ряда работ, посвящённых оценке вероятностно-временных характеристик межсетевых экранов при ранжировании набора правил фильтрации. В публикации рассматривается проблема снижения эффективности фильтрации информационных потоков. Проблема возникла из-за использования последовательной схемы проверки соответствия пакетов правилам, а также из-за неоднородности и изменчивости сетевого трафика. Порядок правил неоптимален, и это в многомерном списке существенно влияет на производительность межсетевого экрана, а также может вызывать значительную временную задержку и вариации в значениях времени обслуживания пакетов, что существенно важно для стабильной работы мультимедийных протоколов. Один из способов предотвратить снижение производительности — это ранжировать набор правил в соответствии с характеристиками входящих информационных потоков. В исследовании решаются следующие задачи: определение и анализ среднего времени фильтрации трафика основных передающих сетей; оценка эффективности ранжирования правил. Предложен метод ранжирования набора правил фильтрации и построена система массового обслуживания со сложной дисциплиной обслуживания запросов. Определённый порядок используется для описания того, как запросы обрабатываются в системе, и включает в себя выполнение операций с входящими пакетами и логическую структуру набора правил фильтрации. Таковы элементы обработки информационного потока в межсетевом экране. Подобный уровень детализации не полный, но его достаточно для создания модели. Характеристики СМО получены с помощью методов имитационного моделирования в среде Simulink матричной вычислительной системы MATLAB. На основании анализа полученных результатов были сделаны выводы о возможности повышения производительности межсетевого экрана за счёт ранжирования правил фильтрации для тех скриптов трафика, которые близки к реальным.

Ключевые слова: межсетевой экран, ранжирование правил фильтрации, сетевой трафик, фазовое обслуживание, имитационная модель, система массового обслуживания