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

2023, 31 (4) 345–358
http://journals.rudn.ru/miph

Research article

UDC 519.872:519.217
PACS 07.05.Tp, 02.60.Pn, 02.70.Bf
DOI: 10.22363/2658-4670-2023-31-4-345-358
EDN: DYDLCY

Evaluation of firewall performance metrics with ranging the rules for Poisson incoming packet flow and exponential filtering time

Anatoly Yu. Botvinko, Konstantin E. Samouylov

RUDN University,
6 Miklukho-Maklaya St., Moscow, 117198, Russian Federation

(received: October 20, 2023; revised: December 1, 2023; accepted: December 29, 2023)

Abstract. The given article is a continuation of a number of works devoted to the development of models and methods for ranging the filtration rules to prevent a decrease in the firewall performance caused by the use of a sequential scheme for checking packet compliance with the rules, as well as by the heterogeneity and variability of network traffic. The article includes a description of a firewall mathematical model given in the form of a complex system and a queuing system with a phase-type discipline for request servicing, which formalizes the network traffic filtering process with the functionality of ranging the rules. The purpose of modeling is to obtain estimates for major firewall performance metrics for various network traffic behavior scenarios, as well as to evaluate an increase in the firewall performance due to ranging a filtration rule set. Calculation of estimates for the firewall (FW) performance metrics was made using the analytical method for a Poisson request flow. Based on the analysis of the modeling results, conclusions were drawn on the effectiveness of ranging the filtration rules in order to improve the firewall performance for traffic scenarios that are close to real ones.

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

1. Introduction

Sustainable operation of information infrastructure, including for special-purpose automated systems (AS), the uninterrupted functioning of which is critical for ensuring the security and defense capability of any state, given the avalanche-like growth in the volume of information flows in public networks, high heterogeneity and variability of network traffic parameters, widespread use of multimedia protocols (that are quite sensitive to the length of data transmission delay), as well as a significant increase in the number of various computer attacks, requires firewalls to provide really high performance. A firewall is a local or functional distributing tool that provides control over

the incoming and/or outgoing information in the automated system, and ensures the protection for the AS by filtering the information, i.e., providing analysis of the information by the criteria set and making a decision on its distribution [1].

One of the major factors affecting the search time for filtration rules, and therefore the FW performance, is the order in which the filtration rules are arranged in sets that are linear lists of large dimensions. This is due to the fact that the search time for any rule corresponding to the data under filtration, is in proportion to the number of checked rules. And the filtering time for information flow that meets the conditions contained at the end of a large dimension set, will be much longer than the time required to filter data that meet the conditions contained at the beginning of the rule set [1, 2].

The papers [1, 3, 4] published earlier by the authors, describe the developed method for optimizing a filtration rule set (method for ranging the rules). An increase in the efficiency of traffic filtration is achieved by periodically ranging the filtration rules in descending order of their weights, obtained in accordance with the estimates of the parameters of the filtered information flows. A particularity of the developed approach is the use of the non-parametric method of local approximation (MLA) [4, 5] to evaluate the parameters of filtered information flows. In the ranging process for a rule set, the current characteristics and dynamics of changes in the parameters of information flows are considered. At the same time, there is no need to select a parametric model that is acceptable for all evaluated parameters of information flows. The implementation of MLA has provided the adaptability of the method, as well as a high response speed for changes in the parameters of filtered information flows thanks to the specifics of MLA estimates.

In earlier works [1–3], the effectiveness of using methods for optimizing a filtration rule set was evaluated with the help of a simulation modeling method. To evaluate the effectiveness of methods for set optimization, the presented paper proposes an analytical solution, as well as an algorithm for calculating the probabilistic and temporal characteristics of the queuing system (QS) that formalizes the network traffic filtering process.

2. Firewall model with ranging the filtration rules

As the FW model we chose the previously developed one [2, 3] that reflects the basic patterns and factors of the FW functioning when processing the network traffic. The traffic processing in this model includes two stages that are an initial processing stage and a stage of checking the packet filtration rules.

At the initial processing stage, a packet, which is transmitted in the communication network, is received by the network interface card of the FW. After decoding a sequence of electrical or optical signals and verifying the correctness of the delivered information, the packet is written to the input buffer memory of the network interface controller (NIC). After that, the packet is transferred to a common software buffer allocated in RAM for further extraction and processing by the central processing unit. In the proposed model, the process of decoding a sequence of signals and the process of receiving and transferring the packet from the buffer memory of the network card to the random access memory (RAM) of the FW are considered

as a single initial processing of packets with a given service intensity. It's considered that the packet immediately arrives at the RAM of the firewall. The waiting time of the packet in the buffer memory of the network card isn't taken into account, as well as the packet losses due to distortions in information transmission.

At the stage of checking the packet filtration rules, the central processing unit (CPU), if computing resources are available, provides the sequential check of the compliance of the incoming packet parameters with the conditions of the filtration rules. The remaining incoming packets await the start of servicing in the buffer in the order of their arrival (FCFS, First-Come, First-Served). A similar approach was used in papers [6, 7].

If the packet parameters comply with the conditions of the filtration rules, the packet processing is completed with encoding and transferring the packet to the physical medium. The packets that don't comply with the filtration rules are discarded.

The following operations aren't considered within the traffic filtering process: fragmentation and reassembly of transmitted packets, reassembly of fragmented packets, network address translation, and packet routing. The level of detail of this scheme doesn't include the architecture and operating algorithms of individual components of the microprocessor system, as well as the interface lines between them, commands and control signals. Therefore, the scheme isn't complete, but is sufficient to develop a mathematical model, neglecting (to make things simpler) unimportant secondary factors.

Only permissive rules are considered as filtration ones. The logical structure of the rule set is a linear list. When the first match of a packet to a rule is found, the packet is considered to be processed by the FW. Packets that don't comply with the rules are rejected. One set of filtration rules, implementing the default deny policy, is used. Additional rule sets aren't considered.

Therefore, the traffic filtering process includes the initial processing of the packet and checking if the packet complies with the filtration rules. The packet checking time is considered a random variable distributed according to an exponential law for the initial processing and checking the filtration rules.

The time required to calculate weights and to range the rule set is considered negligible.

Ranging the filtration rule set is considered as ordering the rules in descending order of their weights in accordance with estimates of the parameters of information flows. We suppose that traffic is filtered at the network and transport levels of the reference model of interaction of open OSI systems.

A model with ranging the filtration rules is a complex stochastic system, and, to build it, an aggregative approach was used, which represents the system as an aggregate that has input and output signals. To demonstrate the operation of subsystems, approaches to describing systems adopted in queuing theory, were used [8].

Let's represent the FW model in the form of a system $M(k) = \{Z, L, T, \Phi, G, X, Y\}$ [3], the moment of transition of which from one state to another one is shown in figure 1:

1. $Z = \{\mathbf{z}_0, \ldots, \mathbf{z}_k, \ldots\}$ is a set of system states; $L = (\mu_0, \mu, N, C)$ is set of system parameters; T is a time interval of system operation. Changes in system states occur at discrete time points $t_k^- = t_k - 0$, $t_k \in T$, $k \geqslant 1$; Φ is a system state transition operator, G is a system output

operator; $X = \{\mathbf{x}_0, \ldots, \mathbf{x}_k, \ldots\}$ is a set of input signals entering the system; $Y = \{\mathbf{y}_1, \ldots, \mathbf{y}_k, \ldots\}$ is a set of system output signals.

2. μ_0 is the intensity of the service during initial processing of the packet by the network card of the FW; μ is the intensity of the service during the checking stage (whether a packet complies with one filtration rule in a set); C is the system storage capacity; N is the number of rules in the filtration set.
3. $\mathbf{z}_k = (\mathbf{r}_k, \mathbf{d}_k) \in Z$ is the state of the system on the time interval $[t_{k-1}, t_k)$, where $\mathbf{r}_k = \left(r_1^k, \ldots, r_N^k\right)$ is a rule set, in which the r_i^k component is a rule in the i-th position in the set; $\mathbf{d}_k = \left(d_1^k, \ldots, d_N^k\right)$ is a vector of packet servicing times, in which d_i^k corresponds to the processing time for i-type packets on the interval $[t_{k-1}, t_k)$.
4. $\mathbf{x}_k = \left(x_1^k, \ldots, x_N^k\right) \in X$ is the input signal; x_i^k, $i = 1, \ldots, N$ is random value that characterizes the number of i-type packets corresponding to the r_i^k rule, entering the system on the interval $[t_{k-1}, t_k)$.
5. $\mathbf{p}_k = \left(p_1^k, \ldots, p_N^k\right)$ is a vector of rule weights, set in accordance with MLA estimates of the parameters of information flows; the component p_i^k, $i = 1, \ldots, N$ corresponds to the weight of the rule that takes the i-th position on the interval.
6. $\mathbf{y}_k = (q_k, v_k, w_k, u_k) \in Y$ is an output signal, the components of which correspond to the estimates of performance metrics on the interval $[t_{k-1}, t_k)$; q_k corresponds to the average number of packets in a drive, v_k corresponds to the average time needed to service the packet in the system, w_k corresponds to the average waiting time before the start of servicing the packet in the system, and u_k corresponds to the average residence time of the packet in the system.
7. $\Phi\left(L, \mathbf{z}_k, \mathbf{x}_k\right) = \mathbf{z}_{k+1}$. At the time point t_{k+1}^-, this operator calculates the weights of the $\mathbf{p}_k$ rules in accordance with the estimates of the parameters of the $\widehat{\mathbf{x}}_k$ information flows. It also determines the state of the system $\mathbf{z}_{k+1} = (\mathbf{r}_{k+1}, \mathbf{d}_{k+1})$, where the $\mathbf{r}_{k+1}$ vector is calculated by ranging the $\mathbf{r}_k$ rule set according to the $\mathbf{p}_k$ weights, and the $\mathbf{d}_{k+1}$ vector is calculated according to the resulting set $\mathbf{r}_{k+1}$ and intensities μ_0, μ.
8. $G\left(L, \mathbf{z}_k, \mathbf{x}_k\right) = y_{k+1}$. At the time point t_{k+1}^-, this operator determines the performance metrics on the time interval $[t_{k-1}, t_k)$.

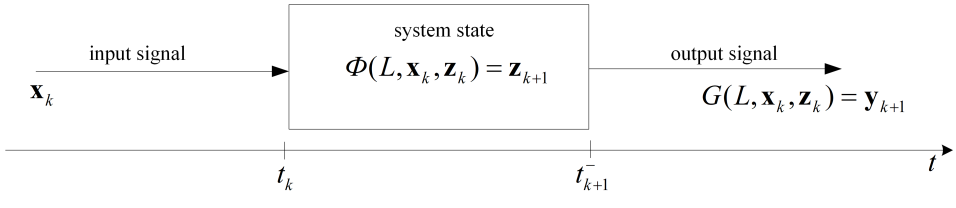

Figure 1. Scheme of the FW model with ranging the filtration rules

Considering the operation of $M(k)$ on the interval $[t_{k-1}, t_k)$, let's present the aggregate in the form of a single-line QS with a storage of limited capacity C,

heterogeneous Poisson incoming flow, and a request service with a distribution function (DF) $B_k(t)$ for the duration of servicing phase-type requests, which depends on the order of filtration rules. The QS receives a request flow that is a superposition of N independent Poisson flows of requests. According to the Basharin-Kendall classification, this QS is designated as $\mathrm{M}_N/\mathrm{PH}/1/\mathrm{C}$ [8].

Representing the system in the form of the QS makes it possible to calculate the performance metrics using an analytical approach for the Poisson flow of requests. Hereinafter, the packets entering the model will be considered as requests, and the QS reflecting the operation of the system on the interval $[t_{k-1}, t_k)$, $k \geqslant 1$ will be designated as $M(k)$.

3. Algorithm for calculating the performance metrics for the exponential distribution of service time

To calculate the performance metrics, let's consider the operation of $M(k)$ on the interval $[t_{k-1}, t_k)$, $k \geqslant 1$. On this interval, the filtration of one batch of packets takes place, there is no ranging for the rule set, and the state vector $\mathbf{z}_k = (\mathbf{r}_k, \mathbf{d}_k)$ remains unchanged.

Let's imagine that a system receives a Poisson flow of requests with intensity $\lambda_{\cdot}(k) = \sum_{i=1}^{N} \lambda_i^k$, which is the sum of independent Poisson flows of requests of N different types. Let's define a rule set $\mathbf{r}_k = \left(r_1^k, \ldots, r_N^k\right)$ in such a way that the i-type request with the λ_i^k intensity will correspond to the r_1^k filtration rule for all $i = 1, 2, \ldots, N$.

Let's define the process of servicing the requests as a sequential transition of phases, starting with the first one. Servicing the request in the 0-th phase corresponds to the stage of the initial processing of FW packets, servicing the request from the 1-st to the N-th phases corresponds to checking the filtration rules. Only one request can be served at a time. If there is no free space in the drive, then the incoming request exits the system without being serviced. If the request matches the rule, its service in the system gets completed; otherwise, the request is transferred to the next phase. After the N-th phase, the service gets completed.

Let's consider the case in which the request service times in each phase are independent of each other and distributed exponentially. The process of request servicing is schematically shown in figure 2.

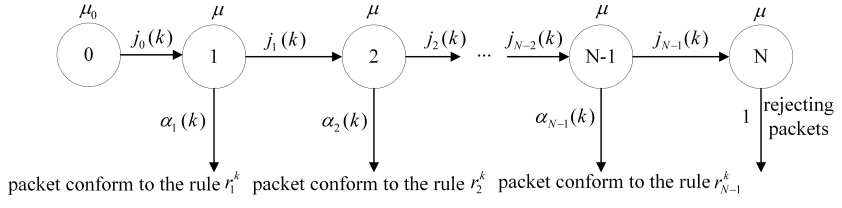

Figure 2. Phase representation of the request servicing process in the firewall model

According to lemma 3 from [8], the probability of a request transition from the i-th phase to the next $i+1$-th phase for the k-th time interval can be

given as follows:

$$j_i(k) = \begin{cases} 1, & i = 0, \\ 1 - \lambda_i(k)/\lambda_{\cdot}(k), & i \in 1, \ldots, N-1. \end{cases} \tag{1}$$

Let's designate $\alpha_i(k) = 1 - j_i(k)$ as the probability of completing the request service in the i-th phase for the k-th time interval. The distribution function (DF) for the time of servicing the request in the QS $M(k)$ will be as follows:

$$B_k(t) = \begin{cases} j_0(k)E(1,\mu_0) + \sum\limits_{i=1}^{N-1} j_i(k)E(i,\mu), & k \in 1, \ldots, N-1, \\ E(N,\mu), & k = N, \end{cases} \tag{2}$$

where $E(i,\mu_0)$ is the Erlang distribution of the i-th order. To analyze the probabilistic and temporal characteristics of the QS, the algorithmic approach proposed by P. Bocharov in [8] was implemented. His main idea is to obtain a solution to the system of global balance equations for the Markov process that describes the system, and to find the parameters of the QS in the form of matrix-recurrent formulas. The use of such an approach makes it possible to effectively calculate the QS parameters.

Let's define a random process (RP) $\{\eta(t), t \geqslant 0\}$ on the set of states $X = \{(0)\ (h,i),\ h = 1, \ldots, C+1,\ i = 0, \ldots, N\}$. The states of the X set have the following meaning. If at some point in time $\eta(t) = 0$, then there are no requests in the system. If $\eta(t) = (h,i)$, then there are h requests in the system, and the serviced request is in the i-th phase. The RP build in such a way is a homogeneous Markov process (MP).

All states of the RP are interconnected, their number is finite and equal to $(C+1)(N+1)+1$. Therefore, according to the ergodic theorem for the MP with a finite set of states [8], the RP $\eta(t)$ is ergodic.

Using a matrix representation, below we give the DF $B_k(t)$ of the request service time. Hereinafter, for brevity, the index k of the time interval of the $M(k)$ aggregate is omitted.

$$\begin{gathered} B_k(t) = 1 - \boldsymbol{\beta}^{\mathrm{T}} e^{\mathrm{M}t}\mathbf{1}, \quad t > 0, \\ \boldsymbol{\beta}^{\mathrm{T}}\mathbf{1} = \mathbf{1}, \end{gathered} \tag{3}$$

where the pair of $\left(\boldsymbol{\beta}^{\mathrm{T}}, \mathrm{M}\right)$ is a PH representation of order $N+1$; $\boldsymbol{\beta}^{\mathrm{T}} = (\beta_0, \ldots, \beta_N)$ is the vector of probabilities of sending a request for service to phases $0, 1, \ldots, N$ at the time point t_k, the component β_i, $i \in 0, \ldots, N$ corresponds to the probability of starting the request service in the i-th phase at the time point t_k; and M is an infinitesimal matrix that has the following form:

$$\mathbf{M} = \begin{bmatrix} -\mu_0 & \mu_0 & 0 & 0 & 0 & 0 \\ 0 & -\mu & j_1\mu & 0 & 0 & 0 \\ 0 & 0 & -\mu & \ddots & 0 & 0 \\ 0 & 0 & 0 & \ddots & j_{N-2}\mu & 0 \\ 0 & 0 & 0 & 0 & -\mu & j_{N-1}\mu \\ 0 & 0 & 0 & 0 & 0 & -\mu \end{bmatrix}. \tag{4}$$

Vector $\boldsymbol{\beta}$, in accordance with the fact that the request service in the considered QS always starts from the zero phase, has the following form:

$$\boldsymbol{\beta}^{\mathrm{T}} = (1, 0 \ldots 0). \tag{5}$$

Let's use the following designations for the probabilities of $\{\eta(t), t \geqslant 0\}$ process states:

1. $p_0 = P\{\eta(t) = 0,\, t \geqslant 0\}$ is the stationary probability of the absence of requests in the system.
2. $p_{i,j} = P\{\eta(t) = (i, j),\, t \geqslant 0\}$ is the stationary probability of servicing the request in the j-th phase and presence of i requests within the system.
3. $\mathbf{p}_h^{\mathrm{T}} = (p_{h,0}, p_{h,1}, \ldots, p_{h,N})$, $h = 1, \ldots, C+1$ is a vector of stationary probabilities.
4. p_h is the stationary probability of presence of h requests within the system.

So, the system of global balance equations in matrix form (for the QS), which describes the process of the FW operation and takes filtration into account, is as follows:

$$\begin{cases} -\lambda_{\cdot} p_0 + \mathbf{p}_1^{\mathrm{T}}\boldsymbol{\mu} = 0, \\ \mathbf{p}_1^{\mathrm{T}}(-\lambda_{\cdot}\mathbf{I} + \mathbf{M}) + \lambda_{\cdot}\boldsymbol{\beta}^{\mathrm{T}} p_0 + \mathbf{p}_2^{\mathrm{T}}\boldsymbol{\mu}\boldsymbol{\beta}^{\mathrm{T}} = \mathbf{0}^{\mathrm{T}}, \\ \mathbf{p}_h^{\mathrm{T}}(-\lambda_{\cdot}\mathbf{I} + \mathbf{M}) + \lambda_{\cdot}\mathbf{p}_{h-1}^{\mathrm{T}} + \mathbf{p}_{h+1}^{\mathrm{T}}\boldsymbol{\mu}\boldsymbol{\beta}^{\mathrm{T}} = \mathbf{0}^{\mathrm{T}}, \quad h = 2, 3 \ldots, C, \\ \mathbf{p}_{C+1}^{\mathrm{T}}\mathbf{M} + \lambda_{\cdot}\mathbf{p}_C^{\mathrm{T}} = \mathbf{0}^{\mathrm{T}}, \end{cases} \tag{6}$$

where $\boldsymbol{\mu}^T = (0, \alpha_1\mu, \ldots, \alpha_{N-1}\mu, \mu)$ is the vector of intensities for the completion of request servicing in the QS of $N+1$ dimension; and $\mathbf{I}$ is the identity matrix of $N+1$ dimension/degree.

Similarly to [8], for the convenience of solving the system of equations (SoE), let's introduce an extra vector $\tilde{p}_h$:

$$\tilde{p}_h^T = \frac{\mathbf{p}_h^T}{p_0}. \tag{7}$$

The solution of the SoE (6) allows us to calculate the performance metrics for the stationary operating mode of the QS (see Algorithm 1).

Algorithm 1. Algorithm for calculating efficiency metrics for exponential distribution of service time Algorithm for calculating the performance metrics for the exponential distribution of service time.

Step 1
Calculate matrices and vectors:

$$\mathbf{D} = -\lambda_{\cdot}\mathbf{I} + \mathbf{M}, \tag{8}$$

$$\beta(\lambda) = 1 + \lambda_{\cdot}\boldsymbol{\beta}^T\mathbf{D}^{-1}\mathbf{1}, \tag{9}$$

$$\widetilde{\mathbf{M}}^{-1} = \mathbf{D}^{-1} - \frac{\lambda_{\cdot}\mathbf{D}^{-1}\mathbf{1}\boldsymbol{\beta}^T\mathbf{D}^{-1}}{\beta(\lambda)}, \tag{10}$$

$$\mathbf{W} = -\lambda_{\cdot}\widetilde{\mathbf{M}}^{-1}, \tag{11}$$

$$\mathbf{W}_R = -\lambda_*\mathbf{M}^{-1}, \tag{12}$$

$$\boldsymbol{\omega}_0^T = \frac{-\lambda_{\cdot}\boldsymbol{\beta}^T\,\mathbf{D}^{-1}}{\beta(\lambda_{\cdot})}. \tag{13}$$

Step 2
Calculate the probabilities $\tilde{p}_h$:

$$\tilde{p}_h^T = \begin{cases} \boldsymbol{\omega}_0^T\mathbf{W}^{h-1}, & h = 1, 2, \ldots, C, \\ \boldsymbol{\omega}_0^T\mathbf{W}^{C-1}\mathbf{W}_R, & h = C+1. \end{cases} \tag{14}$$

Step 3
Calculate p_0, using the normalization conditions for the system of global balance equations and formula (7):

$$p_0 = \left(1 + \sum_{h=1}^{C+1}\tilde{p}_h\right)^{-1}. \tag{15}$$

Step 4
Calculate vector:

$$\mathbf{p}_h^T = p_0\tilde{p}_h^T, \quad h = 1, 2, \ldots, C+1. \tag{16}$$

The calculated vector $\mathbf{p}_h^T$ is the desired matrix-geometric solution for the system of global balance equations (6).

Step 5
Calculate stationary probabilities p_h:

$$p_h = \mathbf{p}_h^T\mathbf{1}, \quad h = 1, 2, \ldots, C+1. \tag{17}$$

Step 6
Using p_h and formulas (18)–(23), calculate the performance metrics for the QS stationary operating mode.

Average number of requests in the QS:

$$l = \sum_{h=1}^{C+1} hp_h. \tag{18}$$

Average queue length:

$$q = \sum_{h=2}^{C+1} (h-1)p_h. \tag{19}$$

Probability of losing the requests:

$$\pi = \vec{p}_{C+1}^{T}\vec{1}. \tag{20}$$

Average residence time:

$$u = l/\lambda_{\cdot}(1-\pi). \tag{21}$$

Average waiting time for service:

$$w = q/\lambda_{\cdot}(1-\pi). \tag{22}$$

Average service time:

$$v = u - w. \tag{23}$$

4. Evaluation of firewall performance metrics

The initial data selected for calculating the performance metrics are as follows: system storage capacity $C = 10$; max number of filtration rules in a set $N = 1500$; initial packet processing time $\mu_0^{-1} = 2.7 \cdot 10^{-5}$ [ms]; time for checking one rule $\mu^{-1} = 5 \cdot 10^{-5}$ [ms]. The packet processing times were taken from papers [6, 7]. The incoming flow is the sum of N independent Poisson flows of requests. The request service time is exponential. To calculate the performance metrics, a program code has been developed in the MATLAB system language.

To check the correctness of the model, the following graphs were constructed:

1. Graphs illustrating the dependence of the performance metrics on the total request flow. The intensities of requests entering the system increase at time points $t_k \in T, k \geqslant 1$; the initial value of the total flow intensity is $\lambda_{\cdot} = 25$ [ms^{-1}]. The number of filtration rules is constant $N = 100$.
2. Graphs illustrating the dependence of the performance metrics on the number of filtration rules. The number of rules increases at time points t_k. The total intensity of the request flow remains constant $\lambda_{\cdot} = 50$ [ms^{-1}].

The values of the performance metrics depend on the types and intensities of the corresponding incoming packets, the rule set and structural parameters of the system. Obviously, the maximum values (given the same system parameters) will be observed when the packets match the last filtration rule, and the minimum ones will be obtained when the packets match the first rule.

That's why, when constructing graphs illustrating the performance metrics, we considered the following cases:

1. Graphs of the performance metrics constructed for the case when incoming requests comply only with the first rule in the rule set, and the time for servicing the request is the least possible.

2. Graphs of the performance metrics constructed for the case when incoming requests comply with the last rule in the rule set, and the time for servicing the request is the largest possible.
3. Graphs of the performance metrics constructed for the case when incoming requests comply with a random rule.

Further in the description, despite the fact that in all three cases the service time is random, for convenience, the graphs will be called as follows: graph of the first rule, graph of the last rule, graph of the random rule, respectively.

The average service time (see figures 3–4) depends on the types of incoming packets and the order of filtration rules and doesn't depend on the total intensity of received requests and the operating mode of the QS. Therefore, when the total request flow increases, the average service time remains constant. Meanwhile, if the number of filtration rules increases, the average service time grows linearly.

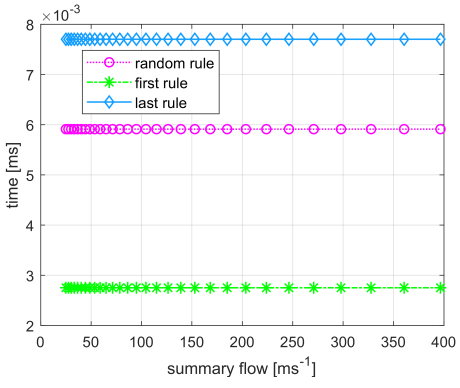

Figure 3. Average service time (constant number of rules)

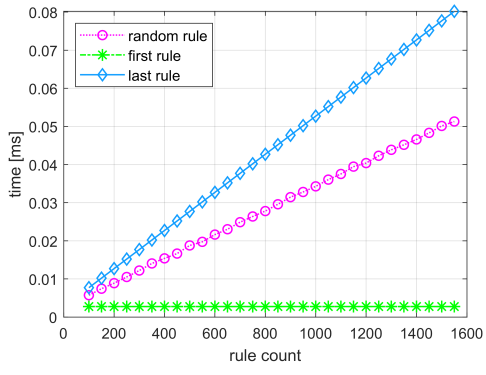

Figure 4. Average service time (constant total intensity of requests received)

Thus, the efficiency of reducing the average service time when ranging the filtration rule set for constant values of the probability of receiving of each type of request will grow with the increase of the rule set and won't depend on the total flow intensity.

The figure 5 shows the dynamics of changes in the average residence time of requests with an increase in the total intensity of the incoming request flow. The average residence time of requests increases with a growth in the flow intensity. For values of the total request flow [ms^{-1}], the type of dependence for the graph of the random rule will change, which corresponds to functioning of the QS in overload mode. Such an overload in the system doesn't result in an unlimited increase in the QS parameters due to the limited storage capacity. The value of the average residence time tends to 0.056 [ms]. At the same time, the difference between the values of the average residence time on the graph of the random rule and the graph of the last rule indicates the possibility of obtaining a significant reduction in the average residence time when ranging the filtration rule set.

For larger rule sets, the difference between the values of the average residence time of requests for the graph with a minimum service time and the graph with a random service time increases with the number of rules (see figure 6).

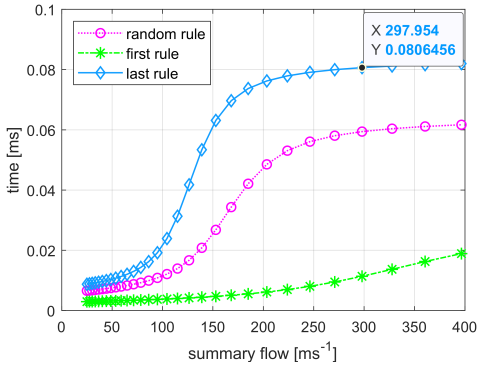

Figure 5. Average residence time in the system (constant number of rules)

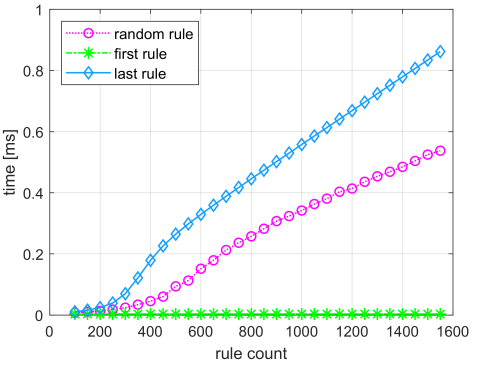

Figure 6. Average residence time in the system (constant total intensity of requests received)

This complies with the peculiarities of the functioning of the FW, since the search time for a rule that matches the filtered data is proportional to the number of checked rules, and indicates the advisability of ranging the filtration rule set. Overload in the system doesn't lead to an unlimited increase in the QS parameters.

Thus, the value of the decrease in the average residence time when ranging the filtration rule set will grow with an increase of the set itself. Also, it won't depend on the intensity of the total request flow starting from the moment when the system gets overloaded.

Figures 7–8 show graphs of changes in the average queue length. Increase of the rule set results in faster filling of the storage, which corresponds to the logic of the FW, because checking a rule set of a large dimension takes more time.

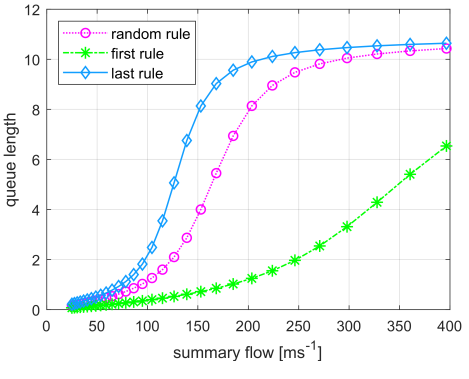

Figure 7. Average queue length (constant number of rules)

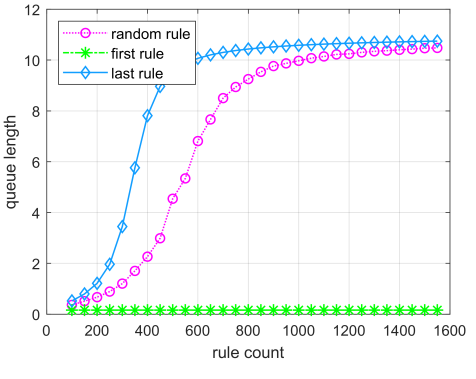

Figure 8. Average queue length (constant total intensity of requests received)

When the QS gets significantly overloaded, the average queue length will be equal to the storage capacity. And the average queue length in the graph of the random rule will approach the values given in the graph of the last rule (see figure 7).

Therefore, the efficiency of ranging the rule set (in order to reduce the queue length) will grow with increasing the load on the system until the system operates in the mode of losing the requests.

5. Conclusion

The obtained estimates of the firewall performance metrics allow us to draw a conclusion on the adequacy of the built analytical model of the FW. We can also draw a conclusion about the possibility to increase the firewall performance by implementing the method for ranging the rule set.

Acknowledgments

This paper has been supported by the RUDN University Strategic Academic Leadership Program.

References

[1] A. Y. Botvinko and K. E. Samouylov, "Evaluation of firewall performance when ranging a filtration rule set," *Discrete and Continuous Models and Applied Computational Science*, vol. 29, no. 3, pp. 230–241, 2013. DOI: 10.22363/2658-4670-2021-29-3-230-241.

[2] A. Y. Botvinko and K. E. Samouylov, "Firewall simulator development for performance evaluation of ranging a filtration rules set," *Distributed Computer and Communication Networks: Control, Computation, Communications. DCCN 2022. Lecture Notes in Computer Science. Lecture Notes in Computer Science*, vol. 13766, no. 3, pp. 221–229, 2022. DOI: 10.1007/978-3-031-23207-7_15.

[3] A. Y. Botvinko and K. E. Samouylov, "Firewall simulation model with filtering rules ranking," *Distributed Computer and Communication Networks: Control, Computation, Communications. DCCN 2020. Communications in Computer and Information Science*, vol. 1337, pp. 533–545, 2020. DOI: 10.1007/978-3-030-66242-4_42.

[4] V. Katkovnik, *Non-parametric data identification and smoothing: local approximation method [Neparametricheskaya identifikaciya i sglazhivanie danny'x: metod lokal'noj approksimacii]*. The science. Main editorial office of physical and mathematical literature Publ., 1985, 336 pp., in Russian.

[5] W. Hardle, *Applied nonparametric regression*. Cambridge: Cambridge university press, 1990, 349 pp.

[6] M. Cheminod, L. Durante, L. Seno, and A. Valenzano, "Performance evaluation and modeling of an industrial application-layer firewall," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 5, pp. 2159–2170, 2018. DOI: 10.1109/TII.2018.2802903.

[7] 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, 2011. DOI: 10.1109/TNSM.2011.122011.110151.

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

For citation:
A. Y. Botvinko, K. E. Samouylov, Evaluation of firewall performance metrics with ranging the rules for Poisson incoming packet flow and exponential filtering time, Discrete and Continuous Models and Applied Computational Science 31 (4) (2023) 345–358. DOI: 10.22363/2658-4670-2023-31-4-345-358.

Information about the authors:

Botvinko, Anatoly Yu. — Candidate of Physical and Mathematical Sciences, assistant professor of Department of Probability Theory and Cyber Security of Peoples' Friendship University of Russia named after Patrice Lumumba (RUDN University) (e-mail: botvinko-ayu@rudn.ru, ORCID: https://orcid.org/0000-0003-1412-981X)

Samouylov, Konstantin E. — Professor, Doctor of Technical Sciences, Head of the Department of Probability Theory and Cyber Security of Peoples' Friendship University of Russia named after Patrice Lumumba (RUDN University) (e-mail: samuylov-ke@rudn.ru, ORCID: https://orcid.org/0000-0002-6368-9680)

УДК 519.872:519.217
PACS 07.05.Tp, 02.60.Pn, 02.70.Bf
DOI: 10.22363/2658-4670-2023-31-4-345-358
EDN: DYDLCY

Оценка показателей эффективности межсетевого экрана с ранжированием правил для пуассоновского входящего потока пакетов и экспоненциального времени фильтрации

А. Ю. Ботвинко, К. Е. Самуйлов

Российский университет дружбы народов,
ул. Миклухо-Маклая, д. 6, Москва, 117198, Российская Федерация

Аннотация. Данная статья является развитием ряда работ по разработке моделей и методов ранжирования правил фильтрации для предотвращения снижения производительности межсетевого экрана, обусловленной использованием последовательной схемы проверки соответствия пакетов правилам, неоднородностью и изменчивостью сетевого трафика. В статье приведено описание математической модели межсетевого экрана в виде сложной системы и системы массового обслуживания с дисциплиной обслуживания заявок фазового типа, формализующей процесс фильтрации сетевого трафика с функциональной возможностью ранжирования правил. Целью моделирования является получение оценок основных показателей эффективности межсетевого экрана для различных сценариев поведения сетевого трафика, а также оценка повышения производительности за счёт ранжирования набора правил фильтрации. Вычисление оценок показателей эффективности МЭ проводится аналитическим способом для пуассоновского потока заявок. На основании анализа результатов моделирования сделаны выводы об эффективности ранжирования правил фильтрации для повышения производительности межсетевых экранов для сценариев трафика, близких к реальным.

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