International Journal of Applied Engineering Research ISSN 0973-4562 Volume 13, Number 8 (2018) pp. 5647-5654

DDoS Attack Detection Based on Network Traffic Phase Coordinates Analysis

Andrey Evgenievich Krasnov
Chief Researcher, State Institute of Information Technologies and Telecommunications (SIIT&T "Informika"), Moscow, Russia, Doctor of Physical and Mathematical Sciences
Researcher ID: C-3673-2017, Orcid id: 0000-0002-4075-4427

Evgeniy Nikolaevich Nadezhdin
Chief Researcher, State Institute of Information Technologies and Telecommunications (SIIT&T "Informika"), Moscow, Russia, Doctor of Technical Sciences
Researcher ID: C-3673-2017, Orcid id: 0000-0002-5861-5847

Vladimir Sergeevich Galayev
Senior Researcher, State Institute of Information Technologies and Telecommunications (SIIT&T "Informika"), Moscow, Russia, Candidate of Physical and Mathematical Sciences
Researcher ID: Q-5283-2017, Orcid id: 0000-0001-7948-4543

Evgenia Andreevna Zykova
Researcher, State Institute of Information Technologies and Telecommunications (SIIT&T "Informika"), Moscow, Russia, Candidate of Physical and Mathematical Sciences,
Researcher ID: C-9882-2015, Orcid id: 0000-0001-5992-7888

Dmitrii Nikolaevich Nikol'skii
Leading Researcher, State Institute of Information Technologies and Telecommunications (SIIT&T "Informika"), Moscow, Russia,
Candidate of Physical and Mathematical Sciences,
Researcher ID: I-7308-2013, Orcid id: 0000-0002-3596-1560

Dmitrii Sergeevich Repin
Deputy Director, State Institute of Information Technologies and Telecommunications (SIIT&T "Informika"), Moscow, Russia, Candidate of Technical Sciences,
Researcher ID: Q-8207-2017, Orcid id: 0000-0003-3232-4803

Abstract

In the work a novel approach for network traffic anomaly detection is proposed. It describes network traffic with phase coordinates consisted of generalized coordinates of traffic workload parameters and generalized velocities determined by non-recursive filtering of generalized coordinates. The information about traffic states is stored in the form of phase portraits — two-dimensional frequency distributions of traffic phase coordinates or their hash functions. Wald's sequential analysis and the Bayesian decision theory with given levels of significance is applied to discriminate phase portraits of different traffic state. Experimental studies proved the efficiency of the developed approach on model traffic states and demonstrated its applicability for the DDoS attack detection on real traffic flows.

Keywords: traffic anomaly detection, DDoS attacks, network traffic, phase coordinates, phase portrait, sequential analysis, Bayesian decision theory.

INTRODUCTION

Major organizations, and particularly network providers, keep large network infrastructures with high volumes of network traffic. For system maintenance they often monitor values of a set of specific traffic parameters such as total traffic volume, incoming and outcoming traffic volume, number of sent packets for certain network protocols, *etc*. The information is used to characterize normal network traffic state and to detect and quantitatively describe deviations from it (anomalous traffic states). Anomalous traffic states could occur as a result of legitimate users behavior (e.g. massive watching of viral video or popular sales promotion), as well as cyberattacks.

Distributed Denial of Service (DDoS) is one of the most critical classes of cyberattacks which aims to disrupt availability of information resources. DDoS attacks grow more popular year by year: they are continuously evolving and become easier and cheaper to conduct. They can lead to the failure of individual hosts and network services, as well as root DNS servers, which will cause a reduced functionality or complete breakdown of corporate information network. One of the urgent tasks in the field of information security today is

the development of special systems for DDoS attacks detection and methods for selection of adequate protection mechanisms that would be most effective under specific conditions.

The problem of network attack identification can be considered as two nearly independent tasks: fast and reliable *detection* of the fact that the system is under attack and precise *classification* of the attack type among known classes with given significance levels. Systematic analysis of traffic states, normal and anomalous, could be used in solving of both tasks. Continuous processing of traffic parameters and wide collection of different system states snapshots would improve the detection efficiency and classification accuracy giving necessary information for cyberattack identification systems.

The main attention in this work is paid to the problem of DDoS attacks identification based on the author's technique of statistical traffic states description by phase coordinates and phase portraits and use of Wald's sequential analysis and Bayesian decision theory with given significance levels.

The aim of the research is to develop the technique of phase coordinates and phase portraits formation for description of network traffic states and to demonstrate its applicability for DDoS attacks identification.

In order to achieve the aim we set the following objectives:

- specify secondary traffic workload characteristics and transform them into phase coordinates;
- form a set of phase portraits for different traffic states;
- conduct experimental studies for DDoS attacks identification using model and real traffic states.

RELATED WORKS

In recent years there were published a number of studies for traffic anomaly detection based on deviations from templates of normal system states. Different statistical and machine learning techniques were applied to the problem including principal component analysis [1, 2], wavelets [3], histogram-based modelling [4], support vector machines [5], correlation analysis [6]. Most of them have shown good performance on detection of intrusion, infected hosts and network attacks.

Data from headers of incoming traffic flow packets (control information) are widely used for DDoS attacks identification. Saied *et al*. applied on artificial neural network algorithm for detection of known and unknown TCP, UDP, ICMP DDoS attacks based on specific characteristic features of the packets control information [7]. In work [8] it was shown that flags in the TCP header could serve for reliable discrimination of normal and malicious traffic and detection of SYN flood DDoS attacks.

In addition to control information workload parameters can be used (information about the sizes and number of packets in a flow) [9-11]. Sizes, number of incoming packets and time intervals between them compose primary traffic workload characteristics. For more in-depth analysis secondary characteristics are used, they are formed as logical [12], entropy [13], correlation [14] and structural [15] functions of primary characteristics.

There are various methods for analysis of network traffic secondary characteristics. Of particular interest are works with formation of probability density distributions and their comparison by different metrics [16, 17], including sequential analysis metrics [18]. However, the efficiency of any approach is mainly achieved by proper selection of secondary characteristics, as specificity of probability density distributions for attacks types is highly dependent on them. In work [12] the specificity is provided by joint use of packet sizes and time intervals between their arrivals. If only one of the characteristics is used, the corresponding probability density distribution would be degenerate — one distribution would correspond to different combinations of characteristics regardless of the data packets order. In this connection, we would like to extend the approach described in [12] to cases in which different secondary characteristics are used in the absence of information about time intervals between data packets. It is possible if we consider network traffic flows as dynamical systems with corresponding generalized coordinates and velocities.

RESULTS

Transformation of Secondary Traffic Workload Characteristics into Phase Coordinates

Let us observe different values of primary traffic characteristics sequences $S_n(k)$ $(n = 1,2,\dots,N;\ k = 1,2,\dots,K)$ and select values of traffic workload parameters from the corresponding fields of data packet headers, where N — number of consecutive values, and K — number of traffic workload parameters taken from different data flows (packet length, octets, number of packets).

Next, we will form different values of secondary characteristics sequences $x_n(k)$ $(n = 1,2,\dots,N;\ k = 1,2,\dots,K)$ from the selected workload parameters. For instance, it is possible to form these values in proportion to:

a) values of $S_n(k)$;

b) normalized values of $S_n(k)/S_n^{max}(k)$;

c) increments $\Delta S_{n,p}(k) = S_n(k) - S_{n-p}(k)$ over the interval $p \in [1,2,\dots]$;

d) relative increments $\Delta S_{n,p}(k)/\left[S_n(k) + S_{n-p}(k)\right] = \left[S_n(k) - S_{n-p}(k)\right]/\left[S_n(k) + S_{n-p}(k)\right]$ over the interval $p \in [1,2,\dots]$;

e) averaged values over a certain sliding interval m

$$\left(S_n(k)\right)_m = \frac{\sum S_i(k)}{m};$$

f) averaged values for non-overlapping intervals with duration m

$\langle S_n(k)\rangle_m = \frac{\sum S_i(k)}{m}$, where $n = 2m, 3m, \dots, nm, \dots$;

g) values of $S_n(k)/S_n(l)$ for $k \neq l$;

h) values of entropy [13], correlation [14] and structural [15] dependencies.

It is possible to combine values of characteristics a)–h) forming their sums. Thus, when summing values from a), c), and e), secondary characteristics are similar to signals of a PID controller.

Along with secondary characteristics sequences $x_n(k)$ let us introduce the sequences $y_n(k)$ $(n = 1,2,\dots,N;\ k = 1,2,\dots,K)$ of conjugate secondary traffic characteristics as a result of passing $x_n(k)$ through a filter with a finite impulse response with coefficients h_n:

$$y_n(k) = h_1\big(x_{n+1}(k) - x_{n-1}(k)\big) + h_1\big(x_{n+3}(k) - x_{n-3}(k)\big) + \dots + h_L\big(x_{n+L}(k) - x_{n-L}(k)\big) \quad (1)$$

The filter coefficients will be fitted so as

$$h_{l-2} < h_l, \forall l = 3,5,\dots,L \quad (2)$$

and sequences $x_n(k)$, $y_n(k)$ are orthogonal:

$$\sum_n x_n(k) y_n(k) = 0 \quad (3)$$

As a case in point, one of the solutions for (3) satisfying the constraint (2) is the Hilbert filter with coefficients $h_l: 1, 1/3, 1/5, \dots, 1/L$.

Conjugate secondary traffic characteristics $y_n(k)$ describe traffic dynamics since they take into account the decreasing contribution of the secondary characteristics $x_n(k)$ derivatives of L-th orders. In this case, for the linear independence of secondary and conjugate secondary characteristics in (1), central differences ought to be used so that each $y_n(k)$ does not depend on $x_n(k), \forall n = 1,2,\dots,N$. Thus, the introduced set of secondary traffic characteristics allows to consider the functioning of the network data transfer channel as the L-th order dynamical system with phase coordinates (generalized coordinates and velocities) X and Y.

Formation of Traffic Phase Portraits. Reasonable Statistics Choice

In accordance with the approach of Ludwig Boltzmann, Henri Poincare and Willard Gibbs [19], the state of a dynamical system with K degrees of freedom and generalized coordinates $X(k)$ and generalized velocities $Y(k)$ $k = (0,1,\dots,K)$ is described by a family of phase portraits that do not contain time in the explicit form, but characterize changes in system states. That is its dynamics in the form of phase trajectories in the corresponding phase spaces $\{x,y\}_k$. Then, phase portraits of any traffic in r-th state (r = 0, 1, …, R) will have the form of two-dimensional scatter diagrams:

$$W[x(k), y(k)|r] = \sum_n^N N_n(k|r)\delta[x(k) - x_n(k|r)]\delta[y(k) - y_k(k|r)] \quad (4)$$

where $N_n(k|r)$ is the number of points on the k-th two-dimensional scatter diagram with coordinates $\big(x_n(k|r), y_n(k|r)\big)$, and $\delta[\dots]$ is the Dirac delta-function.

To describe a phase portrait, we can take various statistics depending on the dynamics of the generalized coordinate X. So if all sequences x_n $(n = 1,2,\dots,N)$ of the random variable X are distributed according to the normal law and are not statistically interrelated, then as the statistics it is suggested to take average values X, Y and the covariance matrix $\sum$ of the column vectors $z_n = (x_n, y)'$. Then, based on Fisher's criteria for any traffic it is possible to judge whether the values of its secondary characteristics $X(k)$ and $Y(k)$ are related to the corresponding given distributions $W[x(k), y(k)|r]$ describing r-th traffic state $(r = 0,1,\dots,R)$. Values of autocorrelation can be calculated to determine the degree of statistical coupling for values of sequence x_n.

If all sequences x_n, y_n $(n = 1,2,\dots,N)$ of a variable $Z = (X, Y)$ have noise distributed with some normalized probability density $f[(z(k) - \bar{z}_n)'A(z(k) - \bar{z}_n)]$, where A is a real symmetric and positively defined matrix of weights, but there is a significant interrelation in x_n, then it is advisable to take the fuzzy phase portraits themselves as the main statistics:

$$W[x(k), y(k)|r] = \sum_n^N N_n(k|r) f\big[\big(z(k) - \bar{z}_n(k|r)\big)'A\big(z(k) - \bar{z}_n(k|r)\big)\big]. \quad (5)$$

In this case, to judge whether the values of the secondary traffic characteristics $X(k)$ and $Y(k)$ are related to a corresponding distribution $W[x(k), y(k)|r]$ from (5) it is possible to apply metrics proposed in [16-18], directly taking into account values from (5).

The descriptions of statistical interrelation between characteristics of network traffic in the r-th state $(r = 0,1,\dots,R)$ are formed on the basis of their phase portraits $W[x(k), y(k)|r]$ for each k-th characteristic $X(k)$ $k = (0,1,\dots,K)$ of data packets in the form:

$$W[x, y|r] = \prod_{k=1}^K W[x(k), y(k)|r]. \quad (6)$$

Expression (6) is valid if secondary characteristics of network traffic are independent. Statistical relationship of dependent secondary characteristics was investigated in the work [15] on the basis of their cross-correlation dependencies.

Experimental Studies

Anomaly Detection

In the first experiment phase portraits (4) were used to describe traffic states. Figure 1 shows an example of sequences x_n for normal and anomalous (TCP SYN flood DDoS attack) traffic states produced by the same filter. Herein as secondary characteristics averaged values for non-overlapping intervals with duration m (case f)) were taken. Phase portraits $W[x, y|0]$ and $W[x, y|1]$ for both traffic states are represented on Fig. 2.

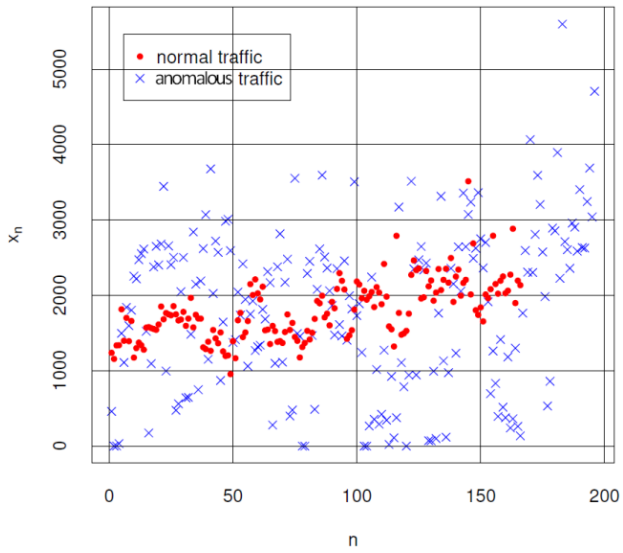
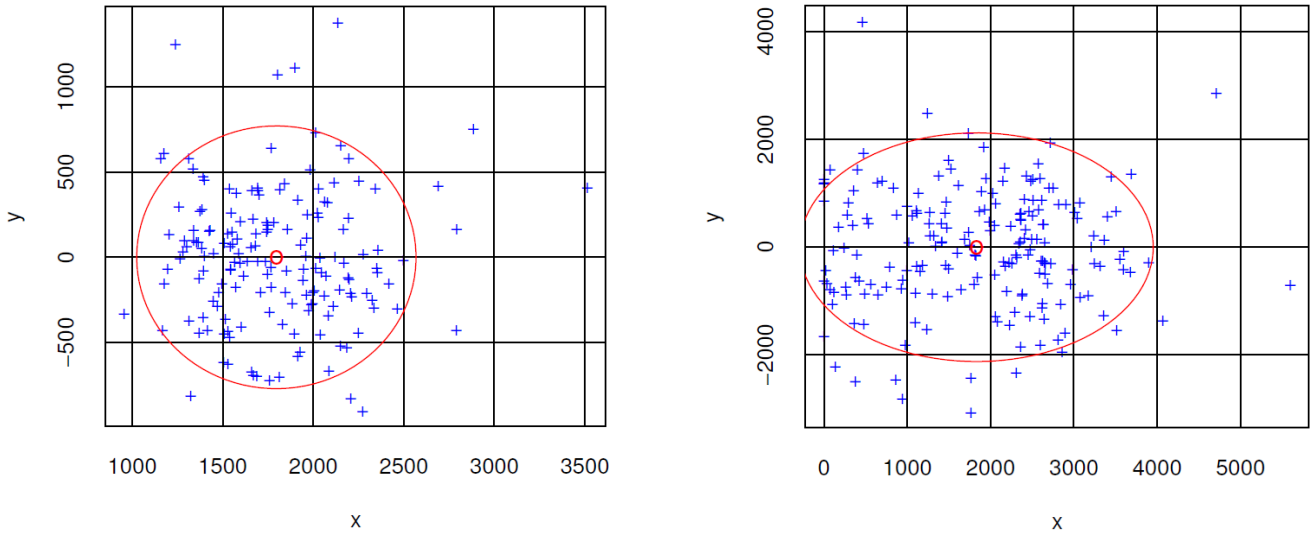


Figure 1. Sequences x_n for normal and anomalous TCP SYN flood DDoS attack traffic states produced by the same filter.



(a) (b)

Figure 2. Phase portraits for normal $W[x, y|0]$ (a) and anomalous $W[x, y|1]$ (b) traffic states.

Assuming that sequences x_n $(n = 1,2, \dots, N)$ are distributed according to the normal law and are not statistically interrelated, the following phase portrait statistics of normal and anomalous traffic states were obtained.

For $[x, y \vee 0]$: $x = 1796, y = 0, \sum = 1.55 * 10^5 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

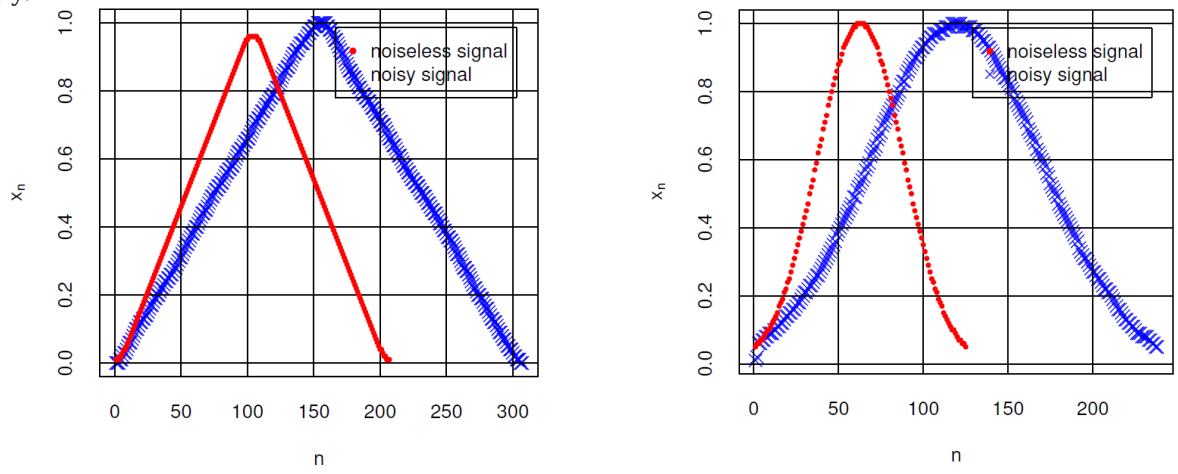
For $[x, y \vee 1]$: $x = 1831, y = 0, \sum = 1.17 * 10^6 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

Based on these statistics and using Fisher's method χ^2 Mahalanobis distances were calculated. They were used to determine boundaries of the confidence intervals for reliable assignment (95%) of experimental samples to normal or anomalous traffic states. The boundaries of the confidence intervals are shown on Fig. 2 by ellipses with corresponding Mahalanobis distances $D_1 = 772$ and $D_2 = 2122$. As can be seen from the figure, the hit of most points in the corresponding confidence intervals guarantees the detection of an attack with a probability of at least 95%.

Discrimination of Model Signals. Hash Functions

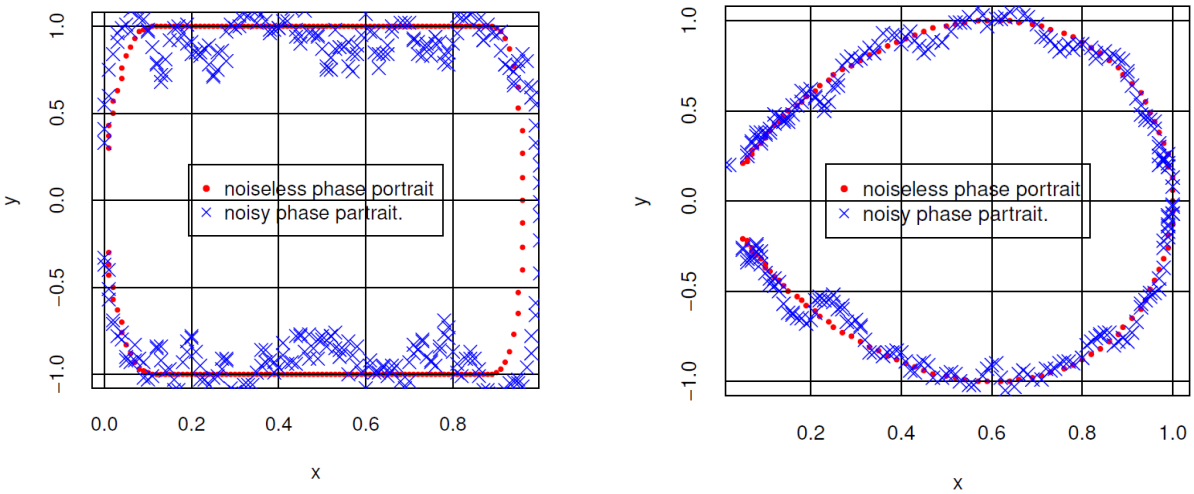
The second experiment involves the discrimination of sequences x_n $(n = 1,2, \dots, N)$ with interrelated (correlated) values. Values from case b) were used as secondary traffic characteristics. We used two model types of signals: normalized triangular (T) signal and normalized Gaussian (G) signal. To estimate influence of interference we also took in consideration both signals affected by noise (standard deviation equal to 0.05) and with scaling. The studied signals

and their phase portraits are shown on Fig. 3 and Fig. 4, respectively.



(a) (b)

Figure 3. Sequences x_n forming triangular signal with and without noise and scaling (a) and Gaussian signal with and without noise and scaling (b).



(a) (b)

Figure 4. Phase portraits $W[x, y|T_1]$ and $W[x, y|T_2]$ for triangular signal with and without noise and scaling (a) and phase portraits $W[x, y|G_1]$ and $W[x, y|G_2]$for Gaussian signal with and without noise and scaling (b).

The signals were discriminated by comparing their phase portraits formed by formula (5). The probability density $f[(z(k) - \bar{z}_n)'A(z(k) - \bar{z}_n)]$ was estimated on the basis of a similarity measure analogical to the potential function 1 / (1 + square of the distance to the nearest neighbor) [20]. Then, obtained values $W[x, y|r]$ $(r = T, G)$ were taken as likelihood, and for likelihood ratios, judgments were made about the presence of noisy and scaled signals to their noiseless versions [21]. In this case, two hypotheses were considered: H_T — a noisy signal refers to a triangular signal; H_G — a noisy signal refers to a Gaussian signal. In the experiment, a reliable discrimination of the sequences was obtained for the probabilities of errors of the first and second type lying in the range 0.01-0.05 (with standard deviation for noise up to 10%).

In order to avoid time-consuming analysis of two-dimensional frequency distributions we advise to transform phase coordinates with hash functions: $\pm X + Y = \mathrm{IF}(Y > 0, X + Y, -X + Y)$ and $X/Y = \mathrm{IF}(Y \neq 0, X/Y, "")$.

Figures 5 and 6 illustrates one-dimensional probability density distributions for hash function $\pm X + Y$ and X/Y for three sequences $(r = 1, 2, 3)$ corresponding to rectangular, triangular, and Gaussian signals and probability density $w(\pm X + Y\,|\mathrm{Noise})$ for normal noise with a standard deviation of 10%. It is clearly seen that probability densities $w(\pm X + Y\,|r)$ and $w(X/Y\,|r)$ differ significantly from each other.

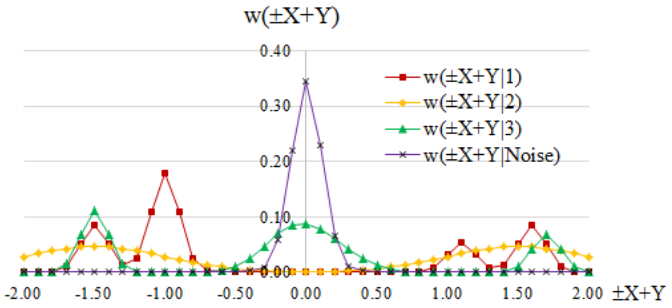


Figure 5. One-dimensional probability density distributions for $\pm X + Y$ hash-functions. Signals $r = 1, 2, 3$ are corresponding to rectangular, triangular, and Gaussian signals.

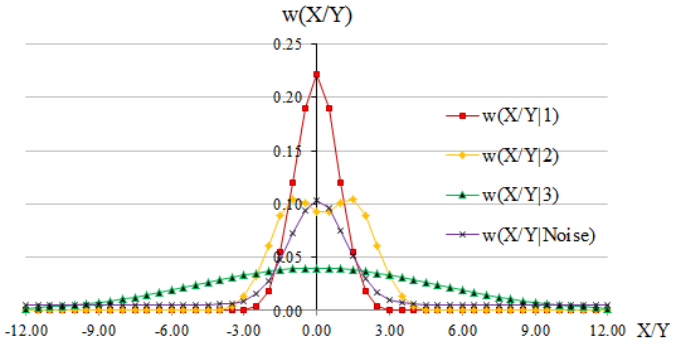


Figure 6. One-dimensional probability density distributions for X/Y hash-function. Signals $r = 1, 2, 3$ are corresponding to rectangular, triangular, and Gaussian signals.

In experiments the hash function $\pm X + Y$ makes it possible to detect all three sequences with errors of the first and second type equal to 0.05 (with a standard deviation of 10%) using two consecutive observations. While for the hash function X/Y it requires at least eight consecutive observations of signal samples x_n, y_n.

Discrimination of DDoS Attacks in Real Traffic Traces

In the third experiment real traffic traces were observed:

- normal traffic with web-surfing;
- anomalous traffic with HTTP flood DDoS attack;
- anomalous traffic generated by Slowloris DDoS tool.

All data were collected by NetFlow protocol. As a secondary characteristic of X, we used the ratio g) which indicates the number of octets per packet.

We investigated the possibility of detecting attacks using the hash function Load-hash $= X + Y = \text{IF}(Y > 0, X + Y, -X + Y)$ formed on the basis of traffic phase coordinates. On Fig. 7 an example of one-dimensional probability density distributions $w(\text{Load} - \text{hash}|r)$ for hash functions for normal traffic state and two attacks $(r = \{\text{Normal}, \text{HTTP} - \text{flood}, \text{Slowloris}\})$. It is clearly seen from the figure that the probability density $w(\text{Load} - \text{hash})$ for hash values differs significantly from each other.

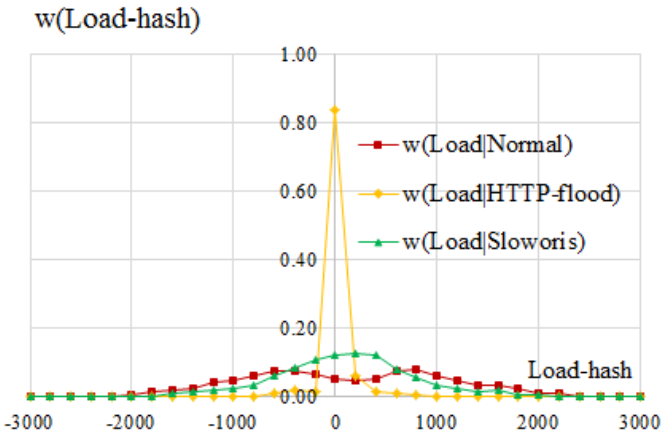


Figure 7. One-dimensional probability density distributions $w(\text{Load} - \text{hash}|r)$ for hash functions of normal traffic state and two network attacks.

Similarly to work [22] the possibility to use address-based statistics of network packets headers was studied. Figure 8 shows an example of one-dimensional probability density distributions $w(\text{Adr} - \text{hash}|r)$ of address hash functions for the normal traffic state and two types of attacks $(r = \{\text{Normal}, \text{HTTP} - \text{flood}, \text{Slowloris}\})$. In this case, the address hash function was constructed as the sum of changes to all address fields of traffic packets headers within specified sliding interval. It can be seen from the figure that probability densities $w(\text{Adr} - \text{hash})$ of address hash values functions also differ significantly from each other.

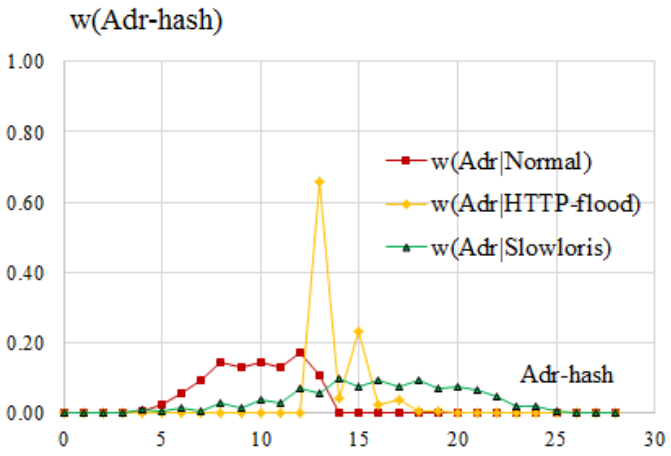


Figure 8. One-dimensional probability density distributions $w(\text{Adr} - \text{hash}|r)$ for hash functions of normal traffic state and two network attacks.

Detection of attacks and identification of their types were built on the basis of Wald sequential analysis and Bayesian decision theory. In this case $w(\text{Adr} - \text{hash}|r)$ were used as *a priori* probabilities, and $w(\text{Adr} - \text{hash}|r)$ — as conditional probabilities. In the sequential analysis of Wald the following likelihood ratios were used:

$$\frac{w(\text{Adr} - \text{hash}_1|r)\times w(\text{Load} - \text{hash}_1|r)\times\ldots\times w(\text{Adr} - \text{hash}_{n^*}|r)\times w(\text{Load} - \text{hash}_{n^*}|r)}{w(\text{Adr} - \text{hash}_1|0)\times w(\text{Load} - \text{hash}_1|0)\times\ldots\times w(\text{Adr} - \text{hash}_{n^*}|0)\times w(\text{Load} - \text{hash}_{n^*}|0)}, \quad (7)$$

where $n*$ — number of consecutive observations (packets or aggregates of packets in traffic) sufficient for given levels of significance (errors of the first and second types). Bayesian *a posteriori* probabilities of attack types were calculated according to a known scheme:

$$W(r|\{\text{Adr} - \text{hash}, \text{Load} - \text{hash}\}_{n^*}) =$$

$$= \frac{w(\text{Adr} - \text{hash}_1|r)\times w(\text{Load} - \text{hash}_1|r)\times\ldots\times w(\text{Adr} - \text{hash}_{n^*}|r)\times w(\text{Load} - \text{hash}_{n^*}|r)}{\sum_{r=1} w(\text{Adr} - \text{hash}_1|r)\times w(\text{Load} - \text{hash}_1|r)\times\ldots\times w(\text{Adr} - \text{hash}_{n^*}|r)\times w(\text{Load} - \text{hash}_{n^*}|r)}. \quad (8)$$

DISCUSSION

Statistical descriptions of network channel traffic states are obtained, which allow forming strategies for both detecting attacks and identifying their types. These strategies were built on the basis of Wald sequential analysis and Bayesian decision theory. In both cases, the main indicators of traffic states are their phase coordinates, generalized coordinates, allowing to describe dynamics of traffic loads of the network channel, and the generalized velocities — changes in the dynamics of these loads. It is shown that traffic phase coordinates can describe both independent and interrelated dynamics of network traffic.

It was demonstrated in experimental studies that compression of traffic state descriptions by different hash functions of its phase coordinates and one-dimensional distributions of the probability densities can be successfully applied for anomalous traffic detection. The presence of anomalies is clearly visible on graphic representations and could be distinguished by any machine learning approach. Therefore, it is arguable that the developed technique allows detecting anomalies in traffic for further classification. Based on traffic phase coordinates and their hash functions with given identification probability of 95% it is possible to classify types of attacks with levels of errors of the first and second types lying in the range 0.01-0.05. It was demonstrated using Wald sequential analysis and Bayesian decision theory.

There is an open question of constructing a hash function that would provide both high classification accuracy and a high speed for traffic processing algorithms. In further research it is planned to apply other methods for hash functions formation and verify their applicability.

CONCLUSIONS

The technique for transformation of network traffic workload parameters into phase coordinates is developed. It describes dynamics of various network channel loads by generalized coordinates and define changes in the dynamics of these loads by generalized velocities. On the basis of phase coordinates, descriptions of various traffic states are introduced in the form of phase portraits. Experimental studies have shown the principal possibility of using the developed technique in DDoS attacks identification systems. It can successfully complement methods widely used in applied network security systems. The main advantages of the technique are the flexibility of the approach to the description of traffic states and the ability to detect deviations from normal traffic states by automatic setting of decision thresholds for given detection probabilities and errors based on sequential analysis and Bayesian decision theory. This is especially important for network traffic of extremely large volumes, when attacks with more than 100 Gbit / s volume are possible.

ACKNOWLEDGMENTS

The work was supported by the Ministry of Education and Science of Russia by lot code 2017-14-579-0002 on the topic: "The development of effective algorithms for detection network attacks based on identifying of deviations in the traffic of extremely large volumes arriving at the border routers of the data network and creating a sample of software complex for detection and prevention of information security threats aimed at denial of service". The agreement No. 14.578.21.0261 on granting a subsidy at September, 26, 2017, a unique identifier of the work (project) is RFMEFI57817X0261.

REFERENCES

[1] Brauckhoff, D., Salamatian, K., and May, M., 2009, "Applying PCA for traffic anomaly detection: Problems and solutions," In INFOCOM 2009, IEEE, pp. 2866-2870.

[2] Liu, Y., Zhang, L., and Guan, Y., 2010, "Sketch-based streaming PCA algorithm for network-wide traffic anomaly detection," In Distributed Computing Systems (ICDCS), 2010 IEEE 30th International Conference, pp. 807-816.

[3] Lu, W., and Ghorbani, A. A., 2009, "Network anomaly detection based on wavelet analysis," EURASIP J. Adv. Signal Process, pp. 4–10.

[4] Kind, A., Stoecklin, M. P., and Dimitropoulos, X., 2009, "Histogram-based traffic anomaly detection," IEEE Trans. on Netw. Serv. Manag, 6(2), pp. 110–121.

[5] Catania, C. A., Bromberg, F., and Garino, C. G., 2012, "An autonomous labeling approach to support vector machines algorithms for network traffic anomaly detection," Expert Syst. Appl. 39(2), pp. 1822–1829.

[6] Chen, C. M., and Lin, H. C., 2015, "Detecting botnet by anomalous traffic," J. Inf. Secur. Appl. 21, pp. 42–51.

[7] Saied, A., Overill, R. E., and Radzik, T., 2016, "Detection of known and unknown DDoS attacks using Artificial Neural Networks," Neurocomputing, 172, pp.385-393.

[8] Jin, S., and Yeung, D. S., 2004, "A covariance analysis model for DDoS attack detection," In Communications, IEEE International Conference, 4, pp.1882-1886.

[9] Ranjan, S., Swaminathan, R., Uysal, M., and Knightly, E. W., 2006, "DDoS-Resilient Scheduling to Counter Application Layer Attacks Under Imperfect Detection," In INFOCOM, pp. 23-29.

[10] Zhou, L., Liao, M., Yuan, C., and Zhang, H., 2017, "Low-Rate DDoS Attack Detection Using Expectation of Packet Size," Secur. Commun. Netw., Article ID 3691629, 14 pages

[11] Tritilanunt, S., Sivakorn, S., Juengjincharoen, C., and Siripornpisan, A., 2010, "Entropy-based input-output traffic mode detection scheme for dos/ddos attacks," ISCIT, 2010 International Symposium on IEEE, pp. 804-809.

[12] Crotti, M., Dusi, M., Gringoli, F., and Salgarelli, L., 2007, "Traffic classification through simple statistical fingerprinting," ACM SIGCOMM Computer Communication Review, 37(1), pp. 5-16.

[13] Reddy, V. S., Rao, K. D., and Lakshmi, P. S., 2012, "Efficient Detection of Ddos Attacks by Entropy Variation," IOSR Journal of Computer Engineering (IOSRJCE), 7(1), pp. 13-18.

[14] Yu, S., Zhou, W., Jia, W., Guo, S., Xiang, Y., and Tang, F., 2012, "Discriminating DDoS attacks from flash crowds using flow correlation coefficient," IEEE Transactions on Parallel and Distributed Systems, 23(6), pp. 1073-1080.

[15] Galayev, V.S., Krasnov, A.E., Nikol'skii, D.N., and Repin, D.S., 2017, "The space of structural features for increasing the effectiveness of algorithms for detecting network attacks, based on the detection of deviations in traffic of extremely large volumes", IJAER, 12, pp. 10781-10790.

[16] Xiang, Y., Li, K., and Zhou, W., 2011, "Low-Rate DDoS Attacks Detection and Traceback by Using New Information Metrics," IEEE Transactions on Information Forensics and Security, 6(2), pp. 426-437.

[17] Li, K., Zhou, W., Yu, S., and Dai, B., 2009, "Effective DDoS attacks detection using generalized entropy metric," In International Conference on Algorithms and Architectures for Parallel Processing, pp. 266-280.

[18] Thatte, G., Mitra, U., and Heidemann, J., 2008, "Detection of low-rate attacks in computer networks," In INFOCOM Workshops 2008, IEEE, pp. 1-6.

[19] Nolte, D. D., 2010, "The tangled tale of phase space," Phys. today, 63(4), pp. 33-38.

[20] Tou, J. T., and Gonzalez, R. C., 1974, "Pattern recognition principles," Addison-Wesley Publishing Company, P.377.

[21] Wald, A., 1947, "Sequential analysis," J. Wiley & Sons, Incorporated, P.212.

[22] Feinstein, L., Schnackenberg, D., Balupari, R., and Kindred, D., 2003, "Statistical approaches to DDoS attack detection and response," In DARPA Information Survivability Conference and Exposition Proceedings, 1, pp. 303-314.