Investigation of time for conducting a successful DDoS attacks in IIoT network

Aleksandar Valentinov Hristov and Valentin Panchev Hristov
Department of Computer Systems, Faculty of Computer Systems and Technologies
Technical University of Sofia
8 Kliment Ohridski blvd., 1000 Sofia, Bulgaria
{ahristov, v_hristov}@tu-sofia.bg

Abstract – **The purpose of this paper is to present the conduction of a natural experiment and to statistically justify the time distribution laws for conducting DDoS network attacks in IIoT. In order to achieve the purpose, the following tasks have to be solved: creating a conceptual model, determination of sufficient number of iterations, conducting a natural experiment, summarizing the raw data, assuming if the resulting distribution is a known theoretical distribution.**

Keywords – **DDoS network attacks, Goodness of fit test, natural experiment, Kolmogrov and chi-square tests.**

I. INTRODUCTION

Obtaining accurate quantitative estimates of the time needed for conducting successful network attacks in Automated Systems operating in secure Industrial Internet of Things (IIoT), Vehicle-to-everything [4] and etc. networks leads to the need of developing an analytical model for assessing the probability of network attacks in IIoT. The development of such a model implies justification of the law of time distribution for the implementation of network attacks. Obtaining accurate quantitative estimates for the time of successful implementation of cyberattacks has great importance for information security. The analysis of the literature devoted to this problem shows that traditionally the time distribution law for the implementation of network cyberattacks is chosen to be exponential [1,4] when developing analytical models. At the same time, the process of conducting a cyberattack is considered either independent of the system that is used for information security and/or it is assumed that in conflict interaction (intruder - protection system in IIoT network) the initial capabilities of the parties are equal and the conflict actions begin simultaneously [3, 4], which in practice happens extremely rarely [5-7].

The purpose of this paper is to present the conduction of a natural experiment and to statistically justify the time distribution laws for conducting Distributed Denial of Services - DDoS network attacks in IIoT. In order to achieve the purpose, the following tasks have to be solved:

1. Creating a conceptual model of the experimental study of the time needed for successful implementation of a specific DDoS network attack, based on the dynamics of the main stages of the attack and the interaction between the intruder and the IIoT network.
2. Determination of the number of iterations of conducting the experiment that have to be done in order to achieve sufficient accuracy.
3. Conducting a natural experiment and presenting its results for the time needed for conducting successful DDoS attacks in IIoT network.
4. Summarizing the raw data in the form of an appropriate frequency histogram, and determination of the statistical properties and the associated empirical Probability Density Function - PDF.
5. Assuming if the resulting empirical distribution is sampled from a known theoretical distribution using the goodness of fit test.

II. RESEARCH METHOD

The research method that is chosen for this paper is natural experiment describing the dynamics of a DDoS network attack in an IIoT network.

To conduct a natural experiment describing the dynamics of DDoS network attack in IIoT, a test setting had to be created. To conduct this experiment, a test equipment was deployed in the laboratory "Information Security in Industrial Systems" of Faculty of Computer Systems and Technologies including the following:

- 2 workstations with the following characteristics: Intel Core i7 7th generation processor with a clock frequency of 2.7 GHz, 12 GB RAM, 500 GB disk storage (SSD), 64-bit operating system (OS) Windows 10. Outside the IIoT network, Kali Linux OS is installed on a separate personal computer (PC) in order to conduct the DDoS network attacks.
- 2 Programmable Logic Controllers (PLCs) Siemens-1512 and Siemens-1510;
- Scalance XB005 Switch that is used for creating the topology depicted in fig. 1 (unmanaged Industrial Ethernet Switch with 5 RJ45 ports that support 10 and 100 Mbit/s broadband speeds; 1 console port; assembly: 35 mm top hat DIN rail mounting);
- Engineers' workstation (WS) with pre-installed PLC programming software (SIMATIC Step 7 Basic V17 executable in Windows 10 Professional 64-bit OS);
- Human Machine Interface (HMI) Siemens KTP700 Basic.

The software emulating computer numerical control machine is implemented on PLC1 and PLC2. To create a DDoS effect, an intruder personal computer is connected to the industrial network through a switch as it is shown on Fig. 1.

The practical component of the step-by-step DDoS attack was implemented by Bash scripts. These scripts were run from the intruder's computer running Kali Linux OS and using pre-installed software developed for this type of attack (UDP requests generation program). The installation and configuration of the intruder's computer were based on the recommendations of the developers of the virtual machine with the Kali Linux OS.

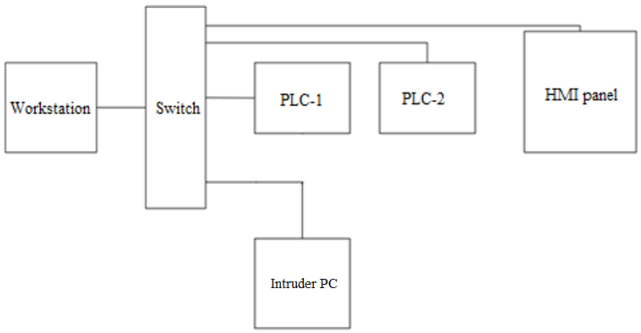

Fig. 1. Experimental setting

For half of the experiments, the intruder is connected directly to the switch and to a PLC. For the second half of the experiments, the DDoS attacks were generated through the workstation as it is shown on Fig. 1 (internal DDoS). Sending tons of UDP requests to the PLC1 is chosen for simulating the DDoS flooding.

III. Experimental results

In order to conduct the experiment, it is assumed that:
- the network is configured so that there is a connectivity between the intruder and attacked device;
- the intruder is familiar with the topology of the network and its vulnerabilities;
- the software used for flooding with UDP requests is running.

Several experimental launches indicated that between the 7th and the 10th second after starting the DoS exposure, the PLC connection with the tracking module was lost. Complete termination of the communication took place after about 70 to 80 seconds. As a result of the natural experiment, those values for the times of successful DDoS attacks were obtained.

To determine the number of iterations of conducting the experiment that have to be done in order to achieve sufficient accuracy it is used Chebyshev's inequality [4], by using the following:

$$P\{|E - E'| > q\sigma\} < 1/q^2 \qquad (1)$$

, where q is any positive constant and σ is the standard deviation.
Whence:

$$(1 - P_o) = 1/(q^2 N) \qquad (2)$$

, or

$$N = q^2/(1 - P_o) \qquad (3)$$

Assuming that the constant q = 1 and the accuracy P_o = 0.95, the number of iterations of conducting the experiment is calculated as:
$n = 1 / (1-0.95) = 20$.

In the preceding section, a natural experiment was presented and more specifically conducting DDoS attack and the empirical time for successful attack is sampled. The determination of the attack time, its distribution function and more specifically properties of the density function are obtained from the raw data that was collected about the time for successful DDoS attack (random variable).

Below it is shown the transformation of sampled to a probability density function as well as results are obtained from conducting goodness of fit test for a specific theoretical distribution (e.g. exponential, and normal).

An initial evaluation for goodness of fit test of the data for a DDoS attack is achieved by comparing the empirical distribution function with the distribution function of the assumed theoretical exponential and normal distributions. If the deviations for the empirical and the theoretical (exponential and normal) functions are not excessive, it may be concluded that the sample is drawn from the exponential or normal distribution. Further, the corresponding hypothesis has to be assumed or discarded trough by applying the goodness of fit test.

For implementation of goodness of fit test, the most widely used tests are:
- the chi-square test;
- the Kolmogrov test.

As a result of the natural experiment, the corresponding times for the successful DDoS attack were obtained. The raw data from the natural experiment, i.e. n=20 values, are given in Table 1.

TABLE 1. TIMES FOR CONDUCTING SUCCESSFUL DDOS ATTACK

79	77	78	79	77
78	79	77	79	77
84	79	70	77	82
85	81	70	73	83

Corresponding Histogram (Fig 2) is constructed conveniently using Excel spreadsheet trough values of relative frequency in the bins.

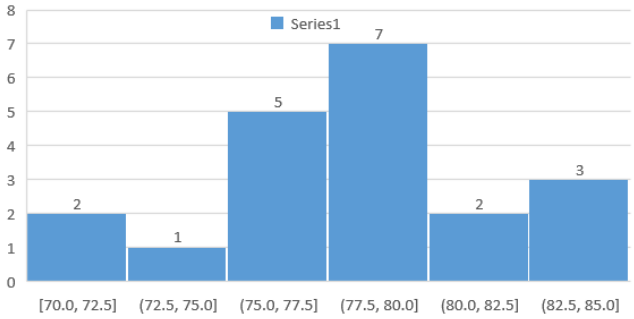

Fig. 2. Histogram of times for conducting DDoS attack

From the obtained values, the mean value and standard deviation can be estimated for the empirical distribution, i.e. $\bar{x}$ = E(x_i) = 78.2, and σ = √D= 3,955 (D=E(x_i^2) - (E(xi))^2).

A. Goodness of fit test for exponential distribution

Here is used the Kolmogrov-Smirnov test (Table 2).

The range of times for conducting DDoS attack (data) from minimum to maximum is divided into six bins: [70,72.5], (72.5,75], (75,77.5], (77.5,80], (80,82.5], (82.5,85]. The histogram shown on Fig 2 is constructed trough values of relative frequency in these bins.

The steps of goodness of fit test implementation for exponential distribution are summarized in Table 2. As it can be seen the first column contains the ranges of the bins, the second column contains the number of occurrences n_i, followed by the average value of the bin. The fourth column of the table contains relative (n = 20) frequencies $F^*(x_i^*)$, which is calculated by equation (4):

$$F^*(x) = n_i / n \ . \qquad (4)$$

Using theoretical frequencies for exponential distribution, i. e. $F(x) = 1 - e\exp(-\lambda x)$, the corresponding values in column $F(x_i^*)$ are filled. In the last column is given by modulo the difference between the empirical and theoretical frequencies:

$$|F^*(x) - F(x)| \qquad (5)$$

Next step is to determine observation value λ_{obs} of Kolmogorov's statistics, i.e.:

$$\lambda_{obs} = max\ (|F^*(x) - F(x)|)\sqrt{k}. \qquad (6)$$

The observation value is filled in the last row of Table 2 (λ_{obs}=1,371).

Finally, this value, λ_{obs}, is compared with the critical value λ_ε (0,05) = 1,358 (See in [5] tabulated values with critical points for Kolmogorov's test).

Because the value λ_{obs}=1,371 is greater than the critical value λ_ε (0,05) = 1,358, the hypothesis for exponential probability density function has to be discarded.

B. Goodness of fit test for normal distribution

Below the chi-square test is used for goodness of fit test for normal distribution. The steps of goodness of fit test implementation for normal distribution are summarized in Table 3. As it was mentioned above, if the deviations for the empirical and the normal density functions are not excessive, it may be concluded that the sample is drawn from the normal distribution.

As it can be seen the first column contains the ranges of the bins, the second column contains the number of occurrences n_i. Values for the next three columns are calculated using equations in (7) and (8):

$$n_i' = n\int_{Xi-1}^{Xi} f\ (t\ )dt = \frac{nh}{\sigma}\ \varphi(u_i) \qquad (7)$$

, where n is number of observations (n = 20), h is length of the bins, (h =2.5) and

$$u_i = \frac{x_{ц} - \overline{x_в}}{\sigma_в},\ \varphi(u_i) = \frac{1}{\sqrt{2\pi}} e^{-u_i^2/2} \qquad (8)$$

Given n_i' and n_i for bin i of the histogram, a measure of the deviation between the empirical and observed frequencies is computed trough an Excel formulae =NORM.S.DIST(C2,FALSE).

It is well known, that x^2 is asymptotically a chi-square probability density function with k degrees of freedom, where k is the number of parameters estimated from the experimental data (Table 3) and used for goodness of fit test for normal distribution.

The hypothesis that the observed sample is drawn from the normal distribution f(t) is accepted if $x_{obs}^2 < x^2$, where x^2 is the chi-square value for corresponding k degrees of freedom and the assumed level of significance.

The steps of goodness of fit test implementation and the calculations for normal distribution are summarized in Table 3.

TABLE 2. GOODNESS OF FIT TEST FOR EXPONENTIAL DISTRIBUTION

x_i, c	n_i	x_i^*	$x_i^* n_i$	$F^*(x_i^*)$	$F(x_i^*)$	$\lvert F^*(x_i^*) - F(x_i^*)\rvert$
70 – 72.5	2	71.25	142.5	0.1000	0.5970	0.4970
72.5 - 75	1	73.75	73.75	0.0500	0.6096	0.5596
75 – 77.5	5	76.25	381.25	0.2500	0.6219	0.3719
77.5 - 80	7	78.75	551.25	0.3500	0.6338	0.2838
80 – 82.5	2	81.25	162.5	0.1000	0.6453	0.5453
82.5 - 85	3	83.75	251.25	0.1500	0.6564	0.5064
					λ_{obs}=1.371	

TABLE 3. GOODNESS OF FIT TEST FOR NORMAL DISTRIBUTION

x_i, c	n_i	u_i	$\phi(u_i)$	n'_i	$n_i - n'_i$	$(n_i - n'_i)^2/n_i$
70 – 72.5	2	-1.757	0.085	1.077	0.923	0.791
72.5 - 75	1	-1.125	0.212	2.678	-1.678	1.052
75 – 77.5	5	-0.493	0.353	4.466	0.534	0.064
77.5 - 80	7	0.139	0.395	4.995	2.005	0.805
80 – 82.5	2	0.771	0.296	3.746	-1.746	0.814
82.5 - 85	3	1.403	0.149	1.884	1.116	0.661
					x_{obs}^2 = 4.186	

In this test two parameters are estimated from the observed data - mean time and standard deviation, and the number of bins is 6 so the degrees of freedom is $k = 6 - 1 - 2 = 3$. Using Excel formulae =CHISQ.INV.RT(0.05,3), for level of significance a = 0.05, the critical value is:
x_{crit}^2 (0,05; 3) = 7.814.
Because the critical value $x_{crit}^2 = 7.814$ is greater than $x_{obs}^2 = 4.186$, we may accept the hypothesis for normal distribution.

IV. Conclusion

In this paper is discussed the conduction of a natural experiment for successful DDoS network attacks in IIoT. The following tasks have been solved: creating a conceptual model, determination of adequate number of iterations, conducting a natural experiment, summarizing the raw data. Goodness of fit tests for exponential distribution and for normal distribution have been conducted and the hypothesis that the sample is drawn from the exponential distribution has been discarded while it can be assumed the resulting distribution is sampled from normal distribution.

Acknowledgment

The authors of the present paper are thankful to Research and Development Sector at the Technical University of Sofia for the financial support.

References

[1] Bokova, O. I, at. al. Stages and procedures for forming a method to assess reliability of the information security systems in automated systems and main areas of its implementation in the normative-technical documentation. Computational Science and Mechanics: Current Problems, Institute of Physics Publishing, 2020. P. 012022. DOI: https://doi.org/10.1088/1742-6596/1479/1/012022)

[2] Mikhailov R.L. Dynamic model of information conflict of special purpose information and telecommunications systems. Control, communication and security systems. 2020. No. 3. P. 238–251. DOI: https://doi.org/10.24411/2410-9916-2020-10309 (in Russian).

[3] Hristov A. et al. Developing and experimenting simulation model of DDoS attacks in IIoT networks using Python, Proceedings of the 31st scientific conference "TELECOM 2023", IEEE Conference, Rec # 59629, 16 – 17 November 2023, Sofia, Bulgaria, DOI: 10.1109/TELECOM59629.2023.10409747.

[4] Andreev D., R. Trifonov, M. Lazarova, Challenges Regarding AI Integration in V2X Communication, Proc. Of 12th International Scientific Conference "Computer Science", 13 – 15 September, Sozopol, Bulgaria, accepted.

[5] https://myslide.ru/documents_3/ec615bd9669892ebf5501065200c6989/img20.jpg.

[6] https://statanaliz.info/statistica/proverka-gipotez/kriterij-soglasiya-pirsona-khi-kvadrat/.

[7] Taha, H. Operations Research: an introduction- 8th ed. 2007, Pearson Prentice Hall, ISBN 0-13-188923-0.