International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-9 Issue-8, June 2020

The Method of Logic Cyber Attack Detection of Abuse Functionality Type on Nginx Http-Server and Apache on the Basis of Fuzzy Logic

Vitalii Fesokha, Igor Subach, Volodymyr Kubrak, Artem Mykytiuk, Stanislav Korotaiev

The article presents a method of abuse detection functionality of the most common open source http-servers Nginx and Apache, which currently implement a full web stack and serve more than 60% of traffic on the Internet. The proposed method is based on the application of the mathematical apparatus of fuzzy set theory and fuzzy inference to the selected analysis parameters corresponding to the properties of a logical cyber attack that realizes the vulnerability of web server data. To obtain the most accurate results of fuzzy inference the direction of adaptation of membership functions to the conditions of server operation (changes on the server and analysis of client behavior) based on the application of the mathematical apparatus of genetic algorithms is determined.

Keywords: http-server, abuse, fuzzy logic, genetic algorithms.

I. INTRODUCTION

According to the statistics of the global http-servers ranking in 2020 [1], more than 60% of http/https of global traffic is accounted for by Nginx and Apache web servers, which are opened on the basis of publicly available licenses. That is why the issue of ensuring the proper level of their cyber security is especially relevant.

A general task that is traditionally solved by a web server is to receive requests from software used by clients, process them with server software, then generate and send a response back to the client. To implement this architectural style on the hosting server part for http-server settings you need:

for Nginx – server blocks corresponding to the registered domain names;

for Apache – virtual servers corresponding to the registered domain names. However, one ip-address can correspond to hundreds of different domain names through special DNS resource records (Domain Name System type "A"). So, for each of them the http-server settings are needed. Thus, after receiving a request from the client the web server checks for compliance with the name of the server block (virtual server) specified in the request header (parameter "host") to determine its handler, and compares request_uri – a unique resource identifier with the corresponding parameters of configured virtual server directives. If the name does not match the "host" parameter, the request is usually processed by the block/virtual server that is configured by default. However, if the default instructions are not configured correctly, such a request will be processed by the first specified server block when using Nginx, and by the last virtual server specified in the configuration file while using Apache [2-4].

In this regard, the functional core of the considered http-servers may be abused, which can be classified as a logical cyber attack (Logical Attacks) – the use of logical vulnerabilities (the expected process when performing specific tasks) or software features. However, the vulnerability can be classified as Insufficient Process Validation – insufficient verification of the execution process. To do this, it will be enough to specify the ip-address correspondence of the victim site in the domain name registration system, e.g., the domain "fake-example.com" specify with the ip-address of the victim site "example.com". In this case, the example.com web stack will be available at http://fake-example.com, and to access the https link the certificate of non-compliance will be received, but the site will be also available. The scheme of the http-request duplication with the corresponding logging is shown in Figure 1.

Such abuse can lead to significant negative consequences for the business. One of that results will be a rapid decline in the rating of the victim's Internet resource in searching systems because of indexing the source "fake-example.com" by a Google, as the duplication on the World Wide Web will be indentified.

The area of concern is the difficulties to reveal the fact of using this abuse of the http-servers functionality and the possibilities of using this type of abuse as well. In the first case, the site owner learns about the problem only after establishing the fact of downgrading the site. In the second, to analyze the customer behavior is appropriate, taking into

Revised Manuscript Received on June 30, 2020.
***** Correspondence Author
Vitalii Fesokha*, postgraduate student, Military Institute of Telecommunication and Information technologies named after the Heroes of Kruty, Kiev, Ukraine. Email: vitaliifesokha@gmail.com
Dr. Igor Subach, doctor of technical science, associate professor, head of department, Institute of special communications and information security National technical university of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine. Email:igor_subach@ukr.net
Volodymyr Kubrak, postgraduate student, Institute of special communications and information security National technical university of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kiev, Ukraine. Email:volodymir.kubrak@ukr.net
Artem Mykytiuk, postgraduate student, Institute of special communications and information security National technical university of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kiev, Ukraine. Email:mukuta8888@gmail.com
Stanislav Korotaiev, engineer, Institute of special communications and information security National technical university of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kiev, Ukraine. Email:meduha1998@ukr.net

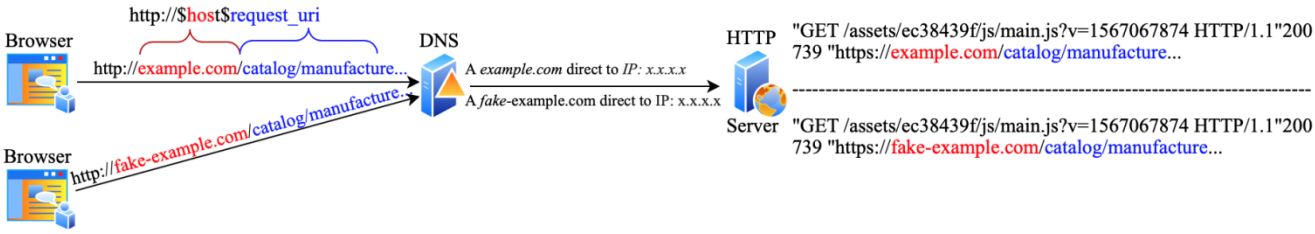

Fig.1 The scheme of the http-request duplication for example.com, fake-example.com

account the web server specifics (number of domains, proxies). If it is theoretically possible to solve the problem of using the web server functionality by setting up a special block/virtual server which will handle requests with incorrect value of the "host" parameter or break "such" connections with the client, then correctness testing of such settings practically are quite complex and often not fully possible.

In consequence, the most efficient approach to solve this problem is using the models that in real time will ensure effective decision-making about the web server state in case of incomplete (unclear) information that is analyzed and simultaneously operated by qualitative and quantitative knowledge.

II. LITERATURE REVIEW

Analysis of recent research and publications [5-7] has shown that the detection of logical cyber attacks, which involve the implementation of expected scenarios of software operation (the presence of acceptable sets of states and many transitions between them) is quite difficult because there is a very narrow approach to it, which is based on the elimination of software vulnerabilities that allow their implementation.

To solve such problems it is significant to apply the proposed solutions in [8-11], based them on the problem formulation, specific activity identifying and nature of the analyzed data limiting, which will cover the studies of logical cyber attacks detecting.

So, there is a task to develop a method for detecting a logical cyber attack such as Abuse of functionality based on fuzzy sets and fuzzy inference, which unlike existing approaches, will solve this problem, as well as work out the testing of software settings in real time.

III. METHOD

A. Preparation of input data

To analyze the state of the web server in order to establish the fact of abuse of its functionality it is remarkable to choose the following parameters for further research, which sufficiently characterize the features in this logical cyber attack and present them with the following linguistic variables:

match_server – the number of inconsistencies in the list of configured server blocks/virtual servers in the configuration file with the parameter "host" in the header of the http-request;

match_uri – the number of consistencies of the parameter request_uri domain name with the real directives in the web stack on the hosting server;

result – the resulting variable of the fact of using the abuse of functionality.

B. The knowledge base with fuzzy rules formation

Under such circumstances, as example, the following fuzzy logical rules describing the states of the http-server with the following linguistic terms and their corresponding values for 15 server blocks/virtual servers will be valid (Fig. 2) [8-11]:

match_server, match_uri – {"L – low [0,15]", "M – medium [15,25,35]", "H – high [35,50]"} on the universal set [0,50];

result – [positive, negative].

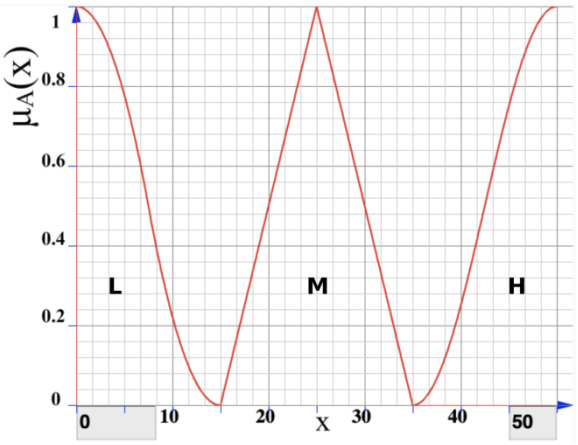

Fig. 2 Graphical representation of the described linguistic terms of membership functions

R1: IF match_server_names is H and match_uri is H Then result is negative with a confidence interval – 1;
R2: IF match_server_names is B and match_uri is B Then result is positive with a confidence interval – 1;
R3: IF match_server_names is C and match_uri is C Then result is positive with a confidence interval – 0.5;
R4: IF match_server_names is B and match_uri is C Then result is positive with a confidence interval – 0.7;
R5: IF match_server_names is H and match_uri is B Then result is negative with a confidence interval – 0.7.

It is enough to parse the log-files of the http-server (access.log) to get the values of the input linguistic variables **match_server, match_uri**.

C. Fuzzyfication

At this stage the degree to which the values of the studied parameters belong to the term sets of mentioned linguistic variables is determined.

It is expedient to study the signs of using the functionality of web servers on the basis of Z (1), S (2) - similar (beginning/ending of the range of values) and triangular (3) (intermediate values of the range of values) type of membership functions with the following values:

match_server, match_uri – {"L – low [0,15]", "M – medium [15,25,35]", "H – high [35,50]"} on the universal set [0,50];

$$\mu^{a_i^p}(x)=\begin{cases}1, x\le a\\ \dfrac{x-a}{b-a}, a<x<b\\ 0, x\ge b\end{cases} \quad (1)$$

$$\mu^{a_i^p}(x)=\begin{cases}0, x\le a\\ \dfrac{x-a}{b-a}, a<x<b\\ 1, x\ge b\end{cases} \quad (2)$$

$$\mu^{a_i^p}(x)=\begin{cases}0, x\le a\\ \dfrac{x-a}{b-a}, a\le x<b\\ \dfrac{c-x}{c-b}, b\le x\le c\\ 0, c\le x\end{cases} \quad (3)$$

D. Parametric adaptation of membership functions

In order to elucidate effective detection of this cyber attack in terms of accuracy it is advisable to provide adaptation (refinement) of the constructed membership functions by experts, on whom the accuracy and reliability of the logical conclusion depend. The changes in operating conditions and/or specifics of http/https-traffic and also changes in the number of server blocks for the Nginx web server or virtual hosts for the Apache web server could be taken into account.

Considering the difficulties of applying classical methods for optimizing nonlinear functions with problems of local extremum, this problem can be solved by applying the mathematical apparatus of genetic algorithms [5]. In addition, the genetic algorithm has no significant mathematical requirements for the types of objective functions and constraints and performs simultaneous search in many areas by using a population of possible solutions, and the transition from one population to another avoids hitting the local optimum.

It is enough to apply the approach providing an adjustment to two parameters for parametric adaptation of membership functions:

the coordinate of the maximum;

compression-tension ratio.

Thus, at each iteration of the algorithm the ranges of term sets of membership functions are evaluated by the correspondence function, which forms a new population [12], that leads to the choice of optimal or suboptimal term values.

E. Fuzzy inference

The stage of execution involves calculations on the obtained membership degrees at the previous stage for all expert decisions in a fuzzy knowledge in order to determine the most appropriate conclusion based on the use of fuzzy max-min operations. The described process can be represented as the following system of fuzzy logical equations (4):

$$\mu^{d_j}(x_1,x_2,...,x_n)=\max_{p=1,k_j}\left\{w_{jp}\min_{i=1,n}\left[\bigwedge_{i=1}^{n}\mu^{jp}(x_i)\right]\right\}, j=\overline{1,m} \quad (4)$$

Thus, when receiving real-time values that correspond to the expert opinions on the rules of R2-R4, the cyber security administrator will make a decision about the abuse of the web server functionality, which is classified as a logical cyber attack.

IV. RESULTS

So, the next stages to conduct the experiment were developed:

OC: Centos 7.7.1908 (Core);

http-сервер: Nginx 1.16.1) –default block;

programming language: Java (openjdk version 1.8.0_252);

fuzzy logic library: JFuzzyLogic.

Conditions: section **II** item **B**.

Iteration analysis: 100 seconds.

For the following vectors of values of input variables: match_server, match_uri, the resulting change took the following values (the fact of using abuse of functionality or without such) in real time.

Table- II: The results of the analysis

№	*count request*	*match_server*	*match_uri*	*result*
1	17	0	0	negative
2	10	10	0 (10 not found)	negative
3	37	34	29 (5 not found)	**positive**
4	50	50	50	**positive**
5	50	12	12	negative
6	30	30	2 (28 not found)	negative
7	40	40	40	**positive**
8	50	25	25	**positive**

V. CONCLUSION

The article represents a method of detecting a logical cyber attack on the most common web servers today, which is based on the operation of the hosting server with fuzzy network activity in real time.

The control of the web server status is carried out on the basis of the analysis parameters which characterize the investigated logical cyber attack. The use of fuzzy sets and genetic algorithms in a combination of mathematical apparatus allows detecting effectively the fact of using the considered abuse.

The practical value of the method is the possibility of detecting logical cyber attacks in conditions of control information uncertainty and ambiguity with an adaptation to changing conditions and features of the object of protection. In addition, this approach solves the problem of testing the web server settings for the possibility of using abuse of its functionality.

REFERENCES

1. Global Web Server Rating 2020. [Online]. Available: https://ru.hostadvice.com/marketshare/server/.
2. Matthew Mombrea. (2015, Oct 1). Why your Nginx server is responding with content from the wrong site. [Online]. Available: https://www.itworld.com/article/2987967/why-your-nginx-server-is-responding-with-content-from-the-wrong-site.html.
3. Why is Nginx responding to any domain name. Active: Aug, 2018. [Online].Available: https://stackoverflow.com/questions/9824328/why-is-nginx-responding-to-any-domain-name.
4. Maksym Baiev. (2019, January 18). Ban on processing non-existent domains. Nginx. [Online].Available: https://mbaev.com/posts/zapret-na-obrabotku-nesuschestwuuschih-domenow-nginx.
5. Ryan O'Leary. (2017, May 17). Abuse of Functionality: The Intersection of Application Security and Ransomware. [Online].Available: https://www.whitehatsec.com/blog/abuse-of-functionality/.
6. (Informit) Ryan C. Barnett. (2006, Feb 17). Mitigating the WASC Web Security Threat Classification with Apache. [Online].Available: https://www.informit.com/articles/article.aspx?p=442984&seqNum=9.
7. (InfoSecPro). Abuse of Functionality. [Online].Available: http://www.infosecpro.com/applicationsecurity/a61.htm.
8. Shanmugavadivu R., N. Nagarajan. “Network intrusion detection system using fuzzy logic”. Indian Journal of Computer Science and Engineering (IJCSE), ISSN : 0976-5166. – 2011. – Vol. 2, No1. – C. 101 – 111.
9. J. Alam, Dr. M. K. Pandey. “Advance Cyber Security System using fuzzy logic”, ACME: Journal of Management &IT, Vol: 10, Issue 1, September 2014 ISSN: 0974-1763.
10. D. K. Levonevskiy ; R. R. Fatkieva ; S. R. Ryzhkov. “Network attacks detection using fuzzy logic”. 2015 XVIII International Conference on Soft Computing and Measurements (SCM). 19-21 May 2015.
11. Ihor Subach, Vitalii Fesokha. “Model of detecting cybernetic attacks on information-telecommunication systems based on description of anomalies in their work by weighed fuzzy rules”. Information technology and security, vol. 5, iss. 1, pp. 29-41, 2017.
12. Y. Mitiushkin, B, Mokin, A. Rotshtein. “Soft Computing: identification of patterns by fuzzy knowledge bases”. Monograph – Vinnytsia. UNIVERSE – 2002. – 145 c.

AUTHORS PROFILE

Vitalii Fesokha, postgraduate student, Military Institute of Telecommunication and Information technologies named after the Heroes of Kruty, Kiev, Ukraine. Bachelor in Computer Science Engineering. Master in Cybersecurity.

Dr. Ihor Subach, doctor of technical science, associate professor, head of department, Institute of special communications and information security National technical university of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Kyiv, Ukraine. Cybersecurity, Decision Support System, Information System.

Volodymyr Kubrak, postgraduate student, Institute of special communications and information security National technical university of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Kiev, Ukraine. Master in Computer Science Engineering.

Artem Mykytiuk, postgraduate student, Institute of special communications and information security National technical university of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Kiev, Ukraine. Bachelor in Computer Science Engineering. Master in Cybersecurity.

Stanislav Korotaiev, engineer, Institute of special communications and information security National technical university of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Kiev, Ukraine. Master in Computer Science Engineering.