ANNALES UMCS INFORMATICA DOI: 10.17951/AI.2016.16.2.39

Methods for increasing security of web servers

Mariusz Nycz
Faculty of Electrical and Computer Engineering
Rzeszow University of Technology
Rzeszow, Poland
mnycz@prz.edu.pl

Mirosław Hajder
University of Information Technology and Management in Rzeszow
Rzeszow, Poland
miroslaw.hajder@gmail.com

Sara Nienajadlo
Rzeszow University of Technology
Rzeszow, Poland
sara.n@op.pl

Abstract— **This article is addressed in most part to people dealing with security of web servers. This paper begins with presenting the statistical dimension of the issue of data security in the modern Internet. This paper begins with presenting statistics dealing with issues of data security on the modern World Wide Web. The authors main focus in this work is presenting the challenges of dealing with security and protection of web communication. The work analyses the security of implementing SSL/TLS (Secure Socket Layer/Transport Layer Security) protocol and proposes a new method of increasing security of web servers. This article is addressed to people dealing with analysis and security of web servers.**

Keywords—security; web serve; intrusion detection; intrusion prevension

I. THE RANGE OF RESEARCH

The amount of information available on the Internet is increasing every year. One of the main sources of growth is internet sites portals made available through web servers. Lately, we have been observing a substantial increase of interest with internet sites by cybercriminals, which can be attributed to a large amount of people visiting those sites. For this reason, neglecting security of similar projects may result in a significant decrease of security for the whole Web information system. Many modern web pages have vulnerabilities that are, in multiple cases, the consequence of bugs present in free-of-charge CMS (Content Management System) modules. For that reason, among others, loopholes in security of said internet sites are one of the most common methods used to penetrate web systems. The large scale of this problem is also the consequence of many web designers focusing most of their attention during the design phase on functionality and aesthetics of the page, rather than quality of coding.

The range of known attacks conducted against www sites and servers that are hosting them is very broad. The most common types of attacks are: switching the main page, deleting the site's file system, replacing information available on the site for a different one, placing trojan horses for the purpose of spreading malicious software, sending spam using the sites resources, overloading the server with artificial traffic to render it inaccessible.

Because of high priority of security, there are many methods and resources capable of providing it in relation to online portals and sites. One of the most widely used tools is automatic security vulnerability scanners. Automatic scanning for gaps in security is a complicated task, especially if the goal is to create a universal method of scanning. Though documents standardizing the construction of safe and functional websites exist, those guidelines are routinely ignored. A situation when a website is not using any tools of providing or verifying security is, unfortunately very common. The best results in provision of security can be achieved by using holistic solutions, especially when they are based on various methods and resources. That is why authors of this work concentrated on developing architectural solutions in the area of hardware-software, allowing for more effective filtering of threats, while maintaining the same level of accessibility and liveability of the whole system [5], [6], [7], [8], [9], [10], [11]. This chapter presents the results of research conducted by the authors, and includes the steps taken:

- Passive collecting of information about web servers functioning in the Subcarpatian voivodeship area using legal methods of data collection. The main focus of data acquisition were: domain names, subnet addresses, services used, types of operating systems, roles of specific nodes, used security methods, etc.
- Analysis of chosen websites in regards to the presence of information assisting in preforming an attack, with focus on the social engineering aspect.
- Researching the usage of identification and authentication mechanisms on websites, used for the purpose of gaining access to subdomains of those site.

- Analysing SSL or TLS protocols used to provide security of the HTTP (Hypertext Transfer Protocol) protocol using Foundstone SiteDigger. This program is used to search for confidential data using API of the search engine provider.
- Developing and implementing a prototype of hardware-software architecture attack analyser.
- Preparing and analysing effectiveness of websites inner architecture resistant to damage and effects of an attack.

II. Results of resource analysis

The presented results were formulated on the basis of analysing 459 websites belonging to multiple owners from the Subcarpatian voivodeship. Table 1 presents the market share of each web server software program. Achieved results were compared with data provided by Netcraft (world) and amudom (Poland). Results show overrepresentation of the Apache server on the Polish market.

Table 2 shows the market share of operating systems used for server maintenance, on which websites were installed. Data for the world market was published by W3Techs.

From the perspective of popularity in Subcarpatian voivodeship, similarly to the rest of the world, the dominating operating systems are unix class with a market share comparatively bigger than the rest of the world. Table 3 presents the percentage of websites using the secure htpps protocol. The achieved results were compared to data presented by websites W3techs (world) and amudom (poland).

TABLE I. Web-server software market in 2016

Product	market share by %		
	Subcarpatian voivodeship	*Netcraft*	*amudom*
Apache	49	35	54
nginx	23	16	18
IIS	5	32	1,3
Other	23	17	26,7

TABLE II. Operating systems market for web-servers in 2016

Operating system	market share by %	
	Subcarpatian voivodeship	*W3Techs*
Unix-like	74,0	66,6
Microsoft	24,0	33,4
Other	2,0	0,0

TABLE III. Websites secured by HTTPS protocol in 2016

Website status	Secured websites by %		
	Subcarpatian voivodeship	*W3Techs*	*amudom*
Secured	4,7	73,7	6,20

Results presented on table 3 are alarming- the percentage of secured websites in their total number is more than ten times below the world average. The author's opinion is that this state of affairs will change in the next couple of months, because of Google's decision not to position websites that had not been secured. However the situation as it is now does not encourage optimism.

Analysing the data provided two conclusions may be reached:

Disregarding the poor usage of safe communication based on HTTPS protocol, both Poland as a whole and Subcarpatian voivodeship, from the perspective of tools used, are in the mainstream of the world Internet. For providing a high availability of a system, the usage of firewalls and IDS/IPS systems is not sufficient. A special architecture of server access is needed [1], [2], [3].

III. Architecture of server protection

The first method implemented to secure access to the servers is an elaborate system of switches and firewalls, providing a high level of availability of servers used. The systems architecture is presented on fig. 1.

The architecture presented above is intended for services that require constant functionality and liveability of a system processing large amounts of information, both by outside and inside users. In an effort to provide uninterrupted communication with outside networks, access to the Internet is realized by two independent links that in order to provide stability of delivery communication are using BGP (Border Gateway Protocol) protocol.

To protect webservers from attacks firewalls block access by using ports different then 80 and 443. Direct access to application servers from the outside of the network is prohibited. To achieve this, intermediary web servers with additional interface must be used. A channel of this type is used, among other things, for user access to databases located on the application servers.

In order to provide a necessary level of defence against damage in graph-model of the system, its vertical connectivity needs to be increased. In practise this can be achieved by increasing the number of active filtering elements (routers, switches and firewalls) and by combining them properly. In the system presented in fig. 1, thanks to doubling the components, damage to any of the elements in the subsystem does not result in loss of availability of the service [4].

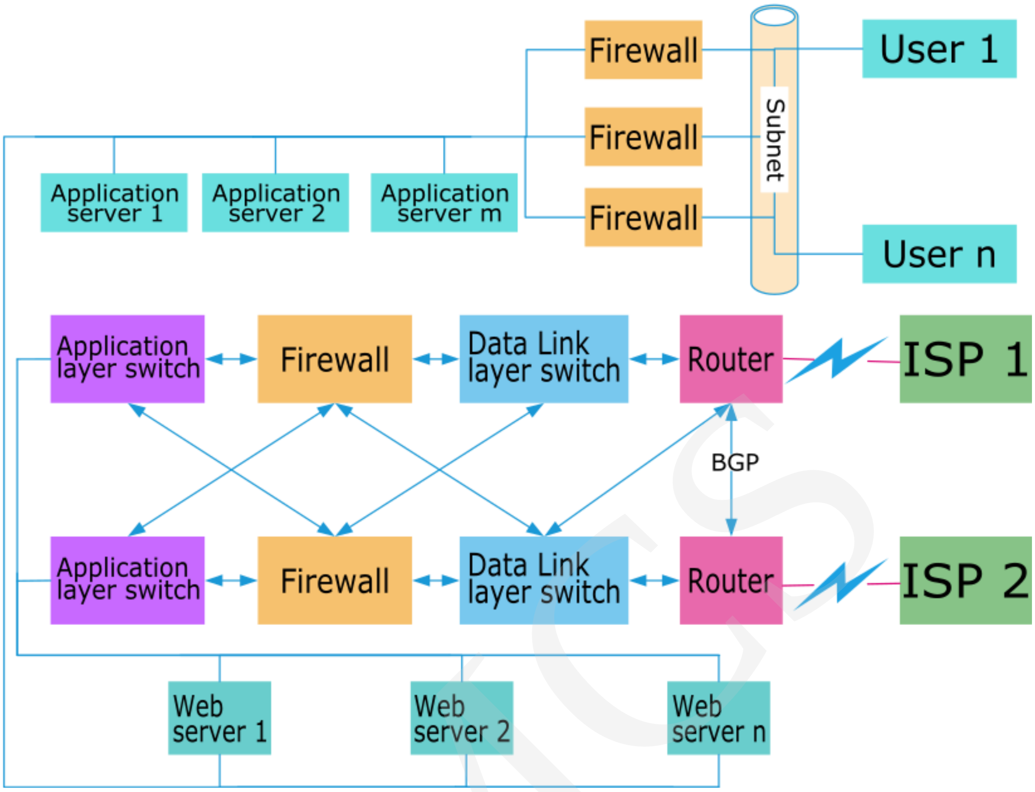

Fig. 2. Architecture of an elaborate firewall

High level of availability is also assured by redundancy of servers, the number of which is chosen based on the amount of requests sent to the server within a specific time. Redundancy of firewalls in addition is used to stop overloading of the server and assuring a constant connectivity of the model , maintaining access to data for inside and outside users. For similar reasons two independent Internet links are used with a set method of using BGP protocol [4], [12], [13], [14].

IV. Attack detection subsystem

To detect an attack directed at a protected set of servers as quickly as possible, system presented in fig. 2 was equipped with the described system of anomaly detection. If the detection of anomalies is performed using artificial neuron networks, general algorithm of functioning (assuming that the network was taught before) has a structure presented in fig. 3. Other described detection methods, might be presented in a similar way.

Specific flowcharts' boxes are represented in fig. 2. Their brief description and the way the neural network learns is presented below.

Web application usage data are being collected by the set of software/hardware sensors, as a part of the intrusion detection system. The data includes: time, geolocation and character of the requests send. They are then further processed. Next step includes creating attack factors streams using path creating algorithm. Collection of the factors used is then made dependant on the hour or the day of the week.

At first the learning process is based on simulated attacks, later it is included in the system as a part of it. The learning process is presented on fig. 3.

As a input data, the algorithm uses B_{pr} - border value of Summary square error and N – number of learning cycles.

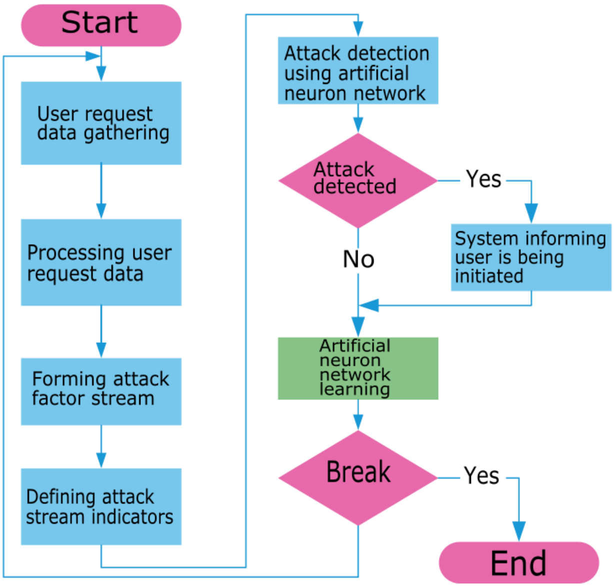
Start
User request data gathering
Processing user request data
Forming attack factor stream
Defining attack stream indicators
Attack detection using artificial neuron network
Attack detected
Yes
No
System informing user is being initiated
Artificial neuron network learning
Break
Yes
End

Fig. 1. A generalized algorithm for the operation of a web-based attack detection system

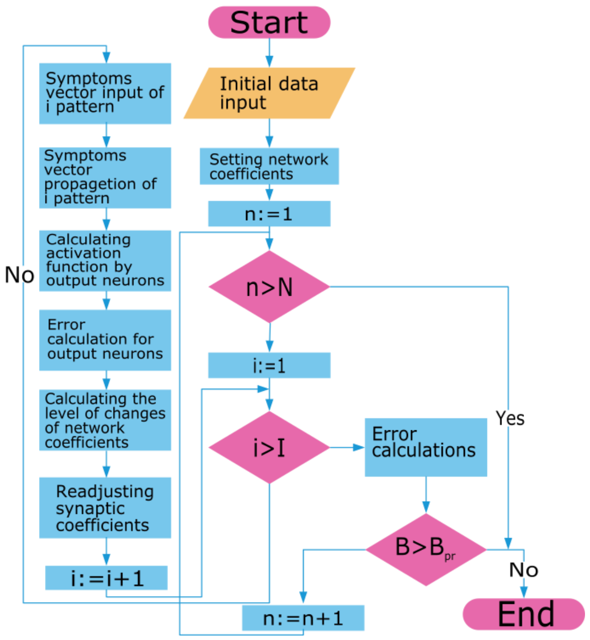
Start
Initial data input
Setting network coefficients
n:=1
n>N
i:=1
i>I
Error calculations
B>B_pr
Yes
No
End
n:=n+1
Symptoms vector input of i pattern
Symptoms vector propagetion of i pattern
Calculating activation function by output neurons
Error calculation for output neurons
Calculating the level of changes of network coefficients
Readjusting synaptic coefficients
i:=i+1
No

Fig. 3. Algorithm for learning neural network

Neural network input contains symptoms' value of ith pattern, later delivered to the networks' next layer. For every output data, the error value is counted and then compared with predicted value. If the comparison equals zero the Compliance response is achieved, if not the correction factor is implemented. The process mention above is repeated for each pattern. Analysis of all the patterns is considered as a full learning cycle. After the end of the learning cycle value of the square error B is compared with border value B_{pr}. The algorithm finishes when the set number of iterations is done or when the error value B is smaller the B_{pr} [4], [15], [16], [17], [18].

V. Summary and Further Reading

Designing web sites using common and automated tools and programs has become accessible to user with not enough technical knowledge to properly and securely use them. Affordable hosting services and easy to use websites management methods are primary reasons for steadily growing number of web sites and services. The designing focuses mainly on visual side of the project, usually overlooking its security aspect, which with ever-growing number of security incidents is a very big oversight. That being said the demand for providing secure services constantly increases. Overall worldwide popularity of SSL/TLS protocol is getting greater by the month, unfortunately polish trends are not as optimistic but it is our hope it will get better soon. Nowadays web applications are primary and sometimes the only source of knowledge about a service. That is why it is of great importance to provide it with appropriate level of accessibility and liveability. The architecture presented in this paper offers undisturbed access to the service. Intrusion and anomaly detection system can allow the system to detect new, not yet known threats, based on normal behaviour patterns. The system tracks users' patterns and using neural network identifies anomalies and attacks not detected by firewalls. Our further work will focus on expanding number of analysed elements of web servers and creating profiled users groups.

References

[1] (2017, Apr.) June 2016 Web Server Survey. [Online]. https://news.netcraft.com/ archives/2016/06/22/june-2016-web-server-survey.html

[2] (2017, Apr.) Statystyki polskiego internetu. [Online]. https://www.amudom.pl/ statystyki-polskiego-internetu

[3] (2017, Apr.) W3Techs - World Wide Web Technology Surveys. [Online]. https://w3techs.com/

[4] M. Hajder, P. Hajder, and M. Nycz, "Inteligentna analiza danych jako metoda detekcji ataków na sieci," in Innowacyjna gmina. Bezpieczeństwo i ekologia. Rzeszów: Wyższa Szkoła Informatyki i Zarządzania z siedzibą w Rzeszowie, 2013, pp. 7-25.

[5] G. Weidman, "Bezpieczny system w praktyce", Gliwice: Helion, 2014.

[6] M. Gregg, "The Network Security Test Lab" John Wiley & Sons, Inc: New York, 2015.

[7] P. Hope, B. Walther, "Testowanie bezpieczeństwa aplikacji internetowych. Receptury", Helion: Gliwice, 2012.

[8] L.Kępa, P. Tomasik, S. Dobrzyński, "Bezpieczeństwo systemu e-commerce, czyli jak bez ryzyka prowadzić biznes w internecie", Helion: Gliwice, 2012.

[9] C. Sanders, "Praktyczna analiza pakietów. Wykorzystanie narzędzia Wireshark do rozwiązywania problemów z siecią", Helion: Gliwice, 2013.

[10] G. Weidman, "Bezpieczny system w praktyce. Wyższa szkoła hackingu i testy penetracyjne", Helion: Gliwice, 2015.

[11] W. Stallings, "Kryptografia i bezpieczeństwo sieci komputerowych. Koncepcje i metody bezpiecznej komunikacji", Helion: Gliwice, 2012.

[12] B. Sullivan, V. Liu, "Web Application Security, A Beginner's Guide", McGraw-Hill Education: New York, 2011.

[13] J. LeBlanc, T. Messerschmidt, "Identity and Data Security for Web Development: Best Practices", O'Reilly Media: New York, 2016.

[14] P. Hope, B. Walther, "Web Security Testing Cookbook: Systematic Techniques to Find Problems Fast", O'Reilly Media: Sebastopol, 2008

[15] T. Rashid, "Make Your Own Neural Network", CreateSpace Independent Publishing Platform, 2016.

[16] M. T. Hagan, H. B. Demuth, M. H. Beale, O. De Jesús, "Neural Network Design", Martin Hagan, 2014.

[17] M. Smart, "Neural Networks for Complete Beginners: Introduction for Neural Network Programming", Mark Smart, 2017.

[18] N. Adams, N. Adams, N. Heard, "Data Analysis for Network Cyber-Security", Imperial College Press: London, 2014.