Smart Firewall Using Machine Learning

S.SANDEEP, C.SHANDEEP SRINIVAS, R.MEENAKSHI, V.BHARATHWAJ, P.PRASATH

Abstract— In today's modern world no system currently available in the world is 100% secure. In addition, we always can notice that there are huge Attack scenarios. Generally, if a new signature is found on the database, then the behavior will be considered as an attack. Vulnerabilities in most computer systems and, it can be exploited by either non-authorized or authorized users. We propose fuzzy based prevention techniques. In which we have the advantage of automatic intrusion prevention system by having a trained data set of previous attack patterns. Having said that, several tools are being designed and implemented for a variety of exploitation in the diverse range of security attacks. Among these tools firewall is the one which allow us to monitor a range of computer systems: an information system, a network or a cloud computing. The objective of this- project is detecting the different types of attacks with varying parameters using predictive learning. In addition, we will include the attacks like XSS, DNS and Cookie poisoning are used by the attackers to exploit the client-based system, these types of attacks will be prevented using machine learning approaches.

Index Terms— Cookie Poisoning, DNS, Exploit, Firewall, Fuzzy logic, Vulnerabilities, Web application, XSS.

1 Introduction

Web application firewalls (WAFs) in a necessary protection mechanism for online software systems. Because of the relentless flow of new kinds of attacks as well as their increased sophistication, WAFs have to be updated and tested regularly to prevent attackers from easily circumventing them. There is a need of data security in cloud computing since user data are stored on server. In our proposed system we focus on recent cloud attacks such as Cross Site Scripting (XSS) attacks, Domain-Name-Server (DNS) server attacks and Cookie poisoning attacks.

Secure computer systems should assure the following services: integrity, authentication, non-repudiation, confidentiality and availability. Integrity assures that no cyber criminals change data that is stored on computer systems or being transmitted between computers. Confidentiality assures that no information is disclosed to unauthorized people.

2 Overview

Using machine driven algorithms to detect intrusions has several advantages; zero-day malware can be detected through statistical analysis. Machine learning can also help data analysts to analyze large amounts of data by analyzing programs or data in groups that might be malicious or not. Generally, security managers observe the patterns of the IP addresses with histories of intrusive behavior. However, intrusions have become more complex. For example, intrusions can be low and slow which means that an attack consists of intrusive behavior over hours, days or weeks and they can have more than one network source. Complex pattern recognition can be done through machine learning. Through automation large numbers of intrusive signatures or patterns can be monitored.

3 Objective

This project contributes to fight against cybercrime attacks in everyday life. This is done by creating more complicated firewall rules with the use of machine-driven algorithms. This paper contributes by giving insight in how to use machine learning algorithms to create these better firewall rules. This method is singular in the sense of how the dataset is created on which the machine learning algorithms are applied. How the data set is created is also characterized in this paper. The goal of this project is to create a model using machine learning techniques that can differentiate between malicious and normal network traffic, extract a set useful firewall rules from this set.

i. Web application vulnerability is tested
ii. XSS, DNS, Cookie Poisoning patterns are observed
iii. Detection and prevention of XSS, DNS, Cookie Poisoning
iv. To improve the efficiency of firewall
v. Web application security is increased

4 Related work

In [1] Jason Bau, Elie Bursztein, Divij Gupta, John Mitchell "State of the Art: Automated Black-Box Web Application Vulnerability Testing, 2010 "Black-box web application vulnerability scanners are automated tools that probe web applications for security vulnerabilities. In order to assess the current state of the art, we obtained approach to eight leading tools and carried out a study of: (i) the class of vulnerabilities tested by these scanners, (ii) their capability against target vulnerabilities, and (iii) the relevance of the target vulnerabilities to vulnerabilities found in the wild. For testing we used a custom web application vulnerable to known and projected vulnerabilities, and previous versions of frequently used web applications containing known vulnerabilities.

In [2] V. Nithya, S. Lakshmana Pandian and C. Malarvizhi "A Survey on Detection and Prevention of Cross-Site Scripting

- *R.Meenakshi, professor, DepartmentofInformationTechnology,Valliammai EngineeringCollege,Kanchipuram, India.*
- *S.Sandeep, C.Shandeep Srinivas, V.Bharathwaj, P.Prasath is currently pursuing bachelor degree program in Information Technology in Valliammai engineering college, Anna university, India, PH-91-7299865263.*

Attack, 2017”. In present-day time, protecting the web application against hacking is a big challenge. One of the common types of hacking method is to attack the web application via Cross-Site Scripting (XSS). Cross-Site Scripting (XSS) vulnerabilities are used to steal web browser’s resources such as cookies, credentials etc. by injecting the malicious Script code on the user’s web applications. Since Web browsers support the execution of scripts, which is used to enable dynamic Web pages attackers can make use of this feature to enforce the execution of malicious code in a user’s Web browser.

In [3] Shinde, Prashant S., and Shrikant B. Ardhapurkar “Cyber security analysis using vulnerability assessment and penetration testing, 2016” This paper focuses on detecting and preventing the cross-site script attacks in web application. 80 percent of the web applications are vulnerable to security threats, as based on the survey conducted by Open Web Applications Security Project (OWASP). Cross-Site Scripting (XSS) vulnerabilities are due to the lack of input validation that allow attackers to insert malicious scripts in user input and the script is executed at another end. This is frequently found within web pages with dynamic content and it carry out different malicious operations like hijacking user sessions, defaces web sites, redirect the user to malicious sites, password theft etc. In this paper, we detect and prevent the cross-site scripting attack in two phases. In first phase, user given URL is extracted and tested for vulnerability using concolic testing approach. It compares concrete and symbolic values and as a result the vulnerable URLs are sent for prevention. In second phase, the URLs whose vulnerability is unknown are injected into Information Leakage Calculator and the decision is taken based on threshold value. The detected XSS attack URLs are prevented using pattern filtering approach. The way of preventing the XSS attack shows the proposed solution effectiveness and convenience.

In [4] Adam Ali.Zare Hudaib” DNS Advanced Attacks and Analysis, 2014” Nowadays DNS is used to load balance, failover, and geographically redirect connections. DNS has become so common it is hard to identify a modern TCP/IP connection that does not use DNS in some way. Unfortunately, due to the accuracy built into the fundamental RFC-based design of DNS, most IT professionals don't spend much time worrying about it. If DNS is attacked — altering the addresses it gives out or taken offline the damage will be enormous. Whether conducted for political motives, financial gain, or just the notoriety of the attacker, the damage from a DNS attack can be calamitous for the target. In this research they have reviewed different DNS advanced attacks and analyzed them. Also, they surveyed some of the most DNS vulnerabilities and ways of DNS attacks protection.

5 Existing System

Web application firewalls (WAFs) protect web systems from malicious attacks. The WAFs inspect incoming HTTP messages and decide whether blocking or forwarding them to the target web application. The decision is often performed based on a set of rules, which are designed to detect attack patterns which is done manually by the security manager.

Disadvantages :

i. Manual updating of attack pattern
ii. New pattern can not be detected
iii. Firewalls can be useful in repelling intrusions, but they offer no protection against sabotage
iv. Frequent updating of new attack patterns is required.

6 Proposed System

In our proposed system we focus on recent cloud attacks such as Cross Site Scripting (XSS) attacks, Domain-Name-Server (DNS) server attacks and Cookie poisoning attacks. The intrusion prevention technique for XSS and Cookie poisoning is been performed using domain name validation and data encryption using advanced encryption standard. The intrusion prevention technique for DNS is been performed using validating and checking the unwanted external / internal links using link guard, eliminating malicious IP address & automatically block the IP and check for any malicious requests.

6.1 Methodology

Web application firewalls (WAFs) are a crucial protection mechanism for online software systems. Because of the relentless flow of new kinds of attacks as well as their increased sophistication, WAFs have to be updated and tested regularly to prevent attackers from easily circumventing them. In our proposed system we focus attacks such as Cross Site Scripting (XSS) attacks, Domain-Name-Server (DNS) server attacks and Cookie poisoning attacks. In this project we propose ML-Driven, an approach based on machine learning and an evolutionary algorithm to automatically detect holes in WAFs against Cross Site Scripting (XSS), Domain-Name-Server (DNS) server and Cookie poisoning attacks. ML-Driven uses machine learning to incrementally learn attack patterns and build a classifier, i.e., that predicts combinations of attack substrings.

Advantages :

i. Safe and Secure connection establishment.
ii. Dynamic pattern recognition and updating.
iii. New attack pattern will be detected.
iv. Intrusion is detected and prevented.
v. Security of the system is increased.

7 System Architecture

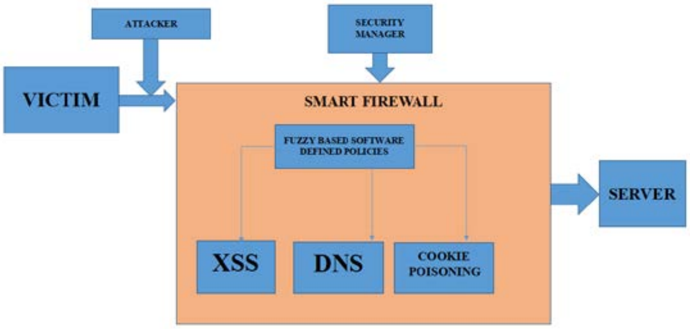

fig 6.1 : This is where the Fuzzy based software defined policy detects and prevents the specified three attacks The modules specified in the above diagram are explained as bellow.

8 Modules

8.1 XSS

Cross Site Scripting (XSS) attacks, this attack injects malicious scripts or code into Web contents has become much popular since the beginning of Web 2.0. The website can be dynamic or static based on the types of services provided. Static websites generally don't experience the security threats while the dynamic website does because their dynamism property in providing user multi-fold services.

Cross-site scripting (XSS) attacks are considered one of the dangerous attack types. It contributes 27% to the total web attacks in 2012 for cloud infrastructure web applications and databases. In XSS, hackers inject malicious scripts, such as JavaScript, VBScript and Flash into a vulnerable dynamic web page to execute the scripts on victim's web browser which later can be compromised and could conduct illegal activities by tricking the victim into clicking a malicious link.

8.2 DNS

The Domain-Name-Server (DNS) server basically performs the task of translation of any domain name to corresponding IP address. But there are many cases when having called server by name, the client has been routed to other evil cloud in its place of the server he asked for. Even though using a DNS security measures such as Domain-Name-System-Security-Extensions (DNSSEC) always reduces the overall effects of DNS security threats and issues but still there are many cases when these security solutions and measures are proved to be not enough when the connection between the sender and the receiver is getting rerouted by a bad connection.

8.3 Cookie poisoning

It this type of attack the change and modification in the contents of cookies is made in order to gain illegal access to any particular application or to a webpage by an attacker. The identity related credentials of the user basically contained by these cookies and once these cookies have accessible by attacker; the integrity related content of these cookies can be used to impersonate any authorized user.

8.4 Fuzzy based software defined policy

A fuzzy intrusion detection system which is host-based and uses data mining methodology and services of the underlying operating system calls. The result of the proposed system shows that the performance is improved and decreases the size of the database as well as time complexity and the rate of false alarms. Fuzzy based software defined policy analyses the attack patterns and software defined policies are been integrated which automatically prevents the attack patterns and block the intruder. In this we proposed fuzzy network intrusion detection method based on class-association-rule mining. The proposed method is dynamic and efficient for both misuse and anomaly detection in networks and it can handle mixed databases which contain both continuous and discrete attributes to mine important class-association rules needed for improvising intrusion detection. The result of the proposed method provides as high detection rate in analogy with other machine learning techniques. Provides better flexibility to some uncertain problems. Detection accuracy is higher

8.5 Intrusion prevention techniques

We have used the Dynamic encryption Generation Technique on the server side, which is used to generate the ciphertext of name attribute in the cookie. The user on the web browser side submits the password and user id to the web server of the web application.

• The web server submits the corresponding data from the browser and generates a cookie.

• Now the web server will dynamically generate encryption value of the name attribute in the cookie and store both these values (original as well as encrypted value) in the form of a table on the server side. Subsequently, the web server will send the encryption value of the name attribute in the cookie to the web browser.

• The web browser will store this encrypted value into its repository. Since the cookies (encrypted version) at the browser 's database now is not valid for the web applications.

Therefore, XSS attack will not be able to imitate the user using stolen cookies which are converted into its hash form.

We proposed DNS Amplification Attacks Detector (DAAD) method that is implemented in the destination side (DNS server). The intrusion prevention technique for DNS is been performed using validating and checking the unwanted external / internal links using link guard, eliminating malicious IP address & automatically block the IP and check for any malicious requests.

Cookie poisoning can be avoided by either performing regular cookie-cleanup or by implementing the encryption scheme for the cookies data.

9 Conclusion

WAFs play an important role to protect online systems. The rising occurrence of new kinds of attacks and their increasing sophistication require that firewalls be updated and tested regularly, as otherwise attacks might remain undetected and reach the systems under protection. We propose ML-Driven, a search-based approach that combines machine learning based automatic intrusion prevention system against vulnerable attacks like DNS, Cookie Poisoning and XSS attacks.

10 Future enhancement

In our proposed system we have detected the patterns of Cross Site Scripting (XSS) attacks, Domain-Name-Server (DNS) server attacks and Cookie poisoning attacks in future work, we will investigate automated approaches to generate effective patches for the WAF under test starting from the learned attack patterns. The above method can also be used for generation of attack patterns other than the aforementioned three attacks. Since the ML algorithms that we use are evolutionary approaches they can be easily adopted for various types of attacks, for pattern generation. In future these patterns can be adopted for enterprise firewall design.

References

[1] Bau, Jason, Elie Bursztein, Divij Gupta, and John Mitchell. "State of the art: Automated black-box web application vulnerability testing." In 2010 IEEE Symposium on Security and Privacy, pp. 332-345. IEEE, 2010.

[2] Nithya, V., S. Lakshmana Pandian, and C. Malarvizhi. "A survey on detection and prevention of cross-site scripting attack." International Journal of Security and Its Applications, no. 3 (2015): 139-152.

[3] Shinde, Prashant S., and Shrikant B. Ardhapurkar. "Cyber security analysis using vulnerability assessment and penetration testing." In 2016 World Conference on Futuristic Trends in Research and Innovation for Social Welfare (Startup Conclave), pp. 1-5. IEEE, 2016.

[4] Hudaib, Adam Ali Zare. "DNS advanced attacks and analysis." International Journal of Computer Science and Security (IJCSS) 8, no. 2 (2014): 63.

[5] Bo Hang,Ruimin Hu,"A novel SYN Cookie method for TCP layer DDoS attack" Published in: 2009 International Conference on Future BioMedical Information Engineering (FBIE)

[6] Suphannee Sivakorn,Iasonas Polakis,Angelos D. Keromytis ,"The Cracked Cookie Jar: HTTP Cookie Hijacking and the Exposure of Private Information" Published on: 2016 IEEE Symposium on Security and Privacy (SP)

[7] Rui Wang; Xiaoqi Jia; Qinlei Li; Shengzhi Zhang, “Machine Learning Based Cross-Site Scripting Detection in Online Social Network” Published in: 2014 IEEE Intl Conf on High Performance Computing and Communications, 2014 IEEE 11th Intl Conf on Embedded Software and Syst (HPCC, CSS, ICESS), 2014 IEEE 6th Intl Symp on Cyberspace Safety and Security

[8] Mukesh Kumar Gupta; Mahesh Chandra Govil ; Girdhari Singh,” Predicting Cross-Site Scripting (XSS) security vulnerabilities in web applications“ Published in: 2015 12th International Joint Conference on Computer Science and Software Engineering (JCSSE).

[9] Zecheng He; Tianwei Zhang; Ruby B. Lee, “Machine Learning Based DDoS Attack Detection from Source Side in Cloud” Published in: 2017 IEEE 4th International Conference on Cyber Security and Cloud Computing (CSCloud).

[10] Ahmad Riza'ain Yusof; Nur Izura Udzir; Ali Selamat ; Hazlina Hamdan , “ Adaptive feature selection for denial of services (DoS) attack” Published in: 2017 IEEE Conference on Application, Information and Network Security (AINS).

[11] Rohilla, Monika, Rakesh Kumar, and Girdhar Gopal. "XSS attacks: analysis, prevention & detection." International Journal of Advanced Research in Computer Science and Software Engineering 6, no. 6 (2016): 264-71,IEEE.

[12] Singh, Tejinder. "Detecting and Prevention Cross–Site Scripting Techniques." IOSR Journal of Engineering 2 (2012): 854-857,IEEE.

[13] Sumitra, B., C. R. Pethuru, and M. Misbahuddin. "A survey of cloud authentication attacks and solution approaches." International journal of innovative research in computer and communication engineering 2, no. 10 (2014): 6245-6253,IEEE.