Web Application Security: Protection from Advanced Persistent Threat

Ms. Shailja Yadav[1], Mr. Sharvan Kumar[2]
[1]*M.Tech in Computer Science Engineering, KIIT College of Engineering, Gurgaon*
[2]*Assistant Professor (CSE), KIIT College of Engineering, Gurgaon*

Abstract- **A group of quite skilled, sophisticated, well organized and motivated person's experts in the field of cyber security exploit that vulnerability to inject a code into your system with a sole intent of extracting some valuable information. These are called web based attacks and unfortunately such a web based attacks are quite common. These Web attacks very dangerous, unfortunately these attacks can't be treated like any other types of attacks. We require web protection applications to protect our system from such attacks; this will prevent the attacker the access to our system and will remain undetectable for a longer period. The most common tricks the attacker's uses are the Adaptive phishing and spear phishing techniques, Cross site scripting and SQL injection etc. In our study, firstly we will describe a model that will help us to understand various attack vectors commonly used for launching and triggering of APT attacks. Secondly we will describe a framework which can be used a framework, which can be used as guidance to analyze, prevent and detect the APT activities inside the application, system, and the network.**

Index Terms- **Advanced Persistent Threat (APT), Web Application Security, Cross site scripting (XSS), Cross site request forgery (CSRF), SQL injection Attack (SQLIA).**

I. INTRODUCTION

Advanced Persistent Threat is a more sophisticated entity in the cyber world; it was originated from the military sector. APT are prominent attackers who are well organized, quite skilled and are well funded. They have emerged as the most dangerous and fast growing security threat [2]. These attackers bypass the security mechanisms with some proficient tricks and thus are quite hard to be detected and in many cases these could only be identified after a long period of time [1].The APT attacks are quite difficult to prevent but these can be refracted towards other targets so that the attack is unprofitable and difficult. The intruders use these techniques to extract sensitive information. The aim of the attackers in most such case may vary from monetary gain, to get the victim`s details or to gather sensitive information [2]. Mostly these attacks are launched through the email based tricks such as spear phishing, in most cases these are backed by other external technique of exploitation. These APT attackers most commonly use the web based exploitation which starts from social engineering and phishing. Taking advantage of economy of scales they break into many sites.

The attackers get a back door entry to the system mainly because of the insufficient training, lack of knowledge and unawareness of the user through the web application using social networking Many times the employees carry the pending office work to their homes and access their organization's site from homes this attracts the attacker to their web. Another trick, most commonly used, by the attackers, along with the social engineering to get the user's credentials is phishing. The most widely used trick by the APT is Spear Phishing. It is done through an e-mail having unauthorized link, clicking to that link, the user is connected to malicious locations or downloads the malicious program code to the user's computer. The APT attack keeps changing itself and thus it is very difficult to discover it and it looks to be entirely new. These APT attacks are extremely sophisticated, continuous and are well organized. Amending past vulnerabilities won't be of any good, our entire focus must be on correcting the existing vulnerabilities and fixing it. Goals of the attack are stealth and to look as close as possible to the legitimate traffic, many times the difference is so minute that the device can't differentiate between them. The attacker aim is not just attack and leave but to have a long term access to the system. APT's are so swift that it can cause very serious damage to the system before the user knows that his system is hit by the APT. The APT has a quite wide range of tools

from very common malware to quite complex ones which they can use; they can even adopt zero-day threat tactics to get their aim.

The two most common and very dangerous categories of attacks are SOL injection (SQLIA) and Cross site scripting (XSS) attack. They are capable of attacking any web application. The attacker, in SQLIA, take advantage of flaws in the programming, executes the malicious queries which lead to injection of malicious code into the system. In case of XSS category, an attacker executes the malicious code to the user's machine by exploiting the insufficient data validation check which ultimately leads to injection ssof malicious code to the user machine. Another type of attack, which is used by the APT to launch the code, is Cross Site Request forgery (CSRF). This attack occurs when a malicious site cause unwanted activities over a legitimate website. The internet user fails to identify this attack; it is termed as sleeping giant.

The entire paper is organized in seven sections:

Section II : Discusses the literature review work.

Section III : This section describes attack and latter the Corresponding defenses, using attack tree & defense tree concept.

Section IV : It describes the proposed protection mechanism for web applications from APT.

Section V : This section describes protection with methodology.

Section VI : It describes the challenges.

Section VII : is the conclusion of the whole paper.

II. RELATED WORK

During past few years many organizations, institutions and individual researcher have conducted extensive studies on Advanced Persistent Threat. The studies in [1], gives us complete information about APT but does not mention about its connection with web security threat. Study in [12] deals with Phishing and Spear Phishing attacks. It also proposed a detention model, this system works well in real time processing of traffic and monitoring but mitigation of web based attacks are not covered in this section. The study in [11] gives in depth study of existing techniques and their limitation for preventing these attacks. The study focuses to network and system security, it also considers applications like web browser are vital to attackers due to its extensive nature and therefore, require an extra bit of security for it. Study in,[13] adopts data leak prevention algorithm to improve prevention techniques against APT. The new generation attack and threats by APT are more complex in nature.

This paper discusses about the detection of malicious codes but the attacks by which these codes are injected are not covered. The model given in[3] explains in detail on a spear phishing attack by the attacker to launch their code into the users network but does not cover other attacks like SQLIA etc. The intended study in[14] Deals about the exploitation of zero day vulnerabilities and presents an idea which addresses the cyber counter intelligence process used as a tool to handle such attacks.

Intrusion detection systems with counterintelligence sensor and fingerprint database can give desired protection but consideration of web flaws too are quite vital to make it more robust.

III. ATTACK MODEL

This section, we will describe the attack and then will learn the corresponding defences, we will explain this with the help of attack tree concept and defence tree concept. Here, we will introduce a new model to explain how the APT attacks the users, and the defence's available with the user organization to counter any such attacks. This is a hybrid model given by Edward Ambroso and Bruce Scheneir [8].

Attack Tree: Here, we will learn about the APT attack using attack tree concept [8, 9]. Attack tree model helps us to know about the attacker, attacks, and the attacker's aims, their security assumptions and the best option to spend the security budget. We will explain, with the help of the figure-I, the attack tree for targeting the web application vulnerabilities to execute their attack.

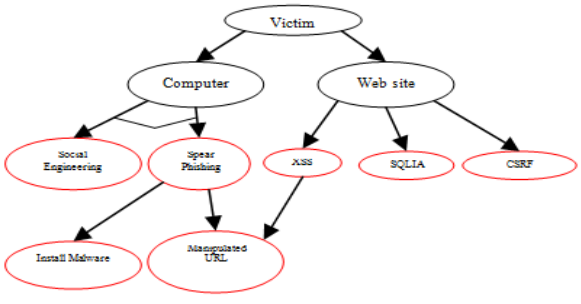

Fig. 1: APT aimed to launch malicious code in victim's machine

The figure above explains in detail how the APT's attack the web users. This model is based on the technique given by Bruce Sheneier[8]. In the above diagram the attackers launch their malicious codes with the help of leaf node and get the desired result. The parent node is linked to the child node with an arrow, when the two arrows are linked with an arc, it is AND relationship, whereas arrows without the arc shows OR relationship.

To extract the required information from the victim, the APT attacks him either through the computer or the website. The APT uses some social engineering along with spear phishing to launch their malicious code, this sends some malicious link on the user's email out of ignorance the victim might click on that link once this happens the victim is further redirected to other locations or some malware is installed into his machine. The Victim can easily be targeted through his website or any other application used by him. The attacker successfully launches the malicious code into the web server, or their targeted machine, by exploiting the existing vulnerability of the website through SQL injection attack, XSS attack and CSRF attack. In the SQL injection attack, an unintended database query fed to the database server can disclose sensitive information that can be used for more complex and dangerous attacks. These attacks may be used separately or can be backed by one or more attacks.

Defence Tree: In this model each node is known as a defence node, it is a countermeasure for the corresponding attack tree. Here, we will explain the defence against APT attack with a tree based diagram. This defence tree is quite similar to the attack tree. Figure 2 explains various defence techniques, which we can use as countermeasure for the attacks discussed in the previous section. In this diagram, the attacks are shown in red oval shapes whereas the defence in green rectangles.

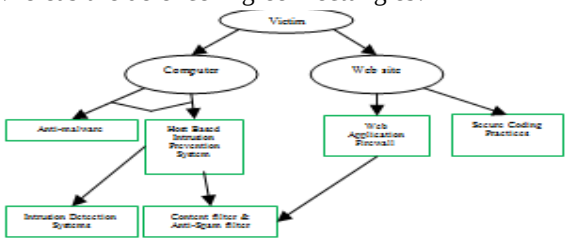

Fig. 2: Victim aimed to defend against APT

To minimize the possibility of attack, Anti-Malware and Host based Intrusion Prevention System (IPS) work together. The suspicious items are identified by them through the security mechanism, which have some abnormal activities; for example, there is an uploading of a large sized file in the internal network, it is treated as abnormal activity because no such activity has been reported in the past history. Another example can be, an active network traffic having known binary source code distributed in different packets. To cope up with manipulated or suspicious URLs, content filtering methods are also quite effective. This is based on the content screening and it denies the access of the web pages or the email that is suspicious or deemed to be objectionable.

Internet firewall categories are divided into 2 parts; one is web filtering. It is used for screening of web pages or the sites and the second is e-mail filtering; it is used for screening e-mails for spam and other unwanted items.

Intrusion Detection System (IDS) is a security system for computers and its networks. In this system information is collected from various parts of the computer and the network. This information, then it is analyzed for any possible malicious security threats. Completely prevention of APT is very difficult but it can surely be detected and corrected adopting related security measures such as: Assessing, monitoring, analyzing, tracking and alarming against abnormal or suspicious events. The firewall provides security to web applications against threats commonly listed on OWASP top ten. The firewall also provides automated and adaptable security against SQLIA, XSS and CSRF.

IV. PROTECTION FRAMEWORK

In this section we will discuss a protection framework for web applications against APT. Our proposed model is depicted in the diagram below; mapping to attack and defence model.

Prevention and detection are two separate activities, but they work together to provide the system a complete protection. The Prevention system gives feedback to the Detection system, and the Detection system collects the input data from the Prevention system.

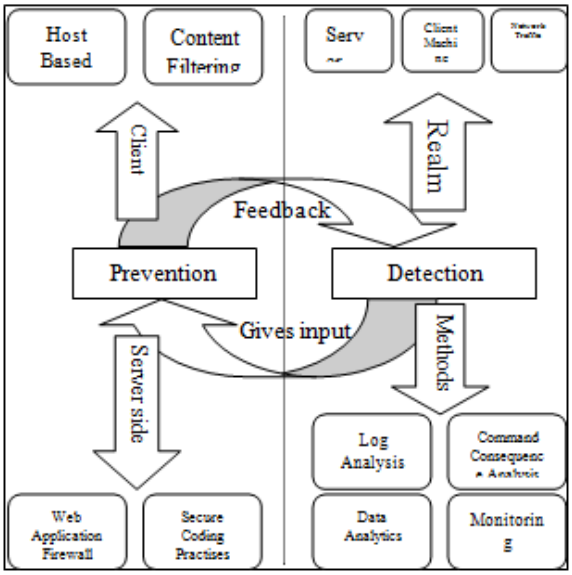

Fig. 3: Protection Framework

Client side prevention

Host Based IPS: It is software, installed into the user's system; it continuously monitors any abnormal or suspicious activity of the inbound traffic by analysing the signature, anomalies and protocols.

Content Filtering: It is software, installed into the user's system; it continuously monitors any abnormal or suspicious activity of the inbound traffic by analysing the signature, anomalies, and protocols.

Server side prevention

Web Application Firewall: This is a software plug-in or content filter. It works on a set of principles or rules governing the whole HTTP communication. It performs strict checks for common attacks like SQLIA and XSS attacks many other attacks can also be recognized and stopped by customizing web application firewall.

Secure Coding practices: These are the advices and guidelines for the web application developers so that to reduce the vulnerabilities that comes due to bad programming. For example, to prevent the SQLIA or OWASP, always use prepared statement, Callable Statement and stored procedure to execute data based queries. When we comply with these guidelines, it is not a hundred percent protection but chances of attack are minimized.

Realm

Detection requires locations where the validation checks can be performed. The most suitable three locations, where detection mechanism can perform:-

First is at the server, it could be application server, web server, database server or a combination of these. This location, since it holds very high level sensitive data, is very crucial. Most commonly, an attacker aims to target this location because here, the chances of success are more. Therefore the detection should focus this area.

Second is the client's machine because here the attacker's signature or footprint may found. To perform their malicious activity, the APT targets an individual, who is more vulnerable and easy to exploit in comparison to others.

Third is the Network traffic; this is a very important realm to find out the malicious activities inside the network by analysing the data packets, controlling, prioritizing and measuring the traffic to determine the cause of the attack.

Methods

In this segment, we propose some of the methods which should be used for detection of APT attacks. One is, deep log analysis, this gives us a sense of legitimacy of the activities being carried out helps us to identify the malicious pattern. These patterns serve as indicators and detect the activities done by the attacker. Another method commonly used to detect the attacks, is the command consequence analysis method. Commands are a set of instructions that work together to complete a given task.

Commands should be delivered in an ethical way so that there is no harm to the computer environment. Unethical delivery of command and its respective consequence create alert signals for the analyst to react accordingly.

Data analytics is digging out meaningful information from a various pattern generated during communication. It uses statistical based approach to visualize the operations performed solely based on signatures, anomalies & policies [1]. In order to maintain effectiveness of detection requires to regular updating the signature databases. A defined static policy governs the organization's traffic flow, identifies any change in the behaviour of intended traffic and gives appropriate response against them. Monitoring of active data packets flowing through

the network should be carried out continuously. Program code within the application should have been monitored, by program like reference monitor, ensure the integrity of existing code and while on execution code does not do any harm [11]. If malicious code causes damage to computer resource such as important files, software code, device drivers' etc. recovery is the only possible solution. In this regard several tools are available in the market that could properly address and access the problem.

V. METHODOLOGY

An intense research was conducted in web application testing and auditing, in which approximately 30 websites were tested for vulnerability which can be a threat to both the web server and the end user. The results of these tests are solely based on the responses given by the owner's of the websites where the attacks were performed.
The study and the result as shown in the diagram 15 give right direction to align our research. This gives a tree diagram based on some decisions to address the injection of malicious code through web pages, as shown in figure 4. In the following paragraphs we will describe the types of attack performed on these sites and results between the impact and ease of attack. A chunk of websites which had a responsible disclosure program were for the top 10 vulnerabilities and beyond. Vulnerabilities like XSS, CSRF, Using servers with known Vulnerabilities; SQL, Click-jacking, Authentication flaws were performed. A complete picture of the above vulnerabilities is given below.

XSS (Cross site Scripting): A Cross Site Scripting is an exploit where the attacker injects a malicious code to a link that appears to be legitimate. The attacker than sends this link to the victim, which on being clicked executes the malicious code on to the victims' system typically allowing the attacker to gather sensitive information about the victim. The vulnerability appears due to scripting flaws, code embedded inside HTML head or body to enhance a page's dynamic loaded capabilities and extensibility of their features.

SQL Injection: In SQL attack, the attacker sends his input in such a way that it gets executed with the original SQL query as a part. As a result the database behaves abnormally. In such a case, the attacker can even bypass the authentication if the wasp is vulnerable to login SQL injection. A compromise in the database will result to a lot of information discloser.

CSRF (Cross Site Request Forgery) : CSRF (Cross-site request forgery) Also called as one Click or session riding or XSRF is an attack where the victims' browser is fooled by the attacker, and it makes a request that the user didn't intend to. For any website your browser is your identity. The website recognizes the user on the base of IP address of its traffic, header, and cookies and links it requests. The CSRF attacks hold the mix identity by manipulating the victim's browser into making request against a website on the attacker's behalf. The attacker relation to the site is immaterial; in fact the website never sees traffic from the attacker. Some users consider phishing attack as CSRF, although it can be a part of it; the difference is phishing attack we manipulate the user into initiating a request from the browser whereas CSRF forces the browser into initiating request. It's not that the attacker has gained access to the victim's browser but yes he has made the browser to do something that the user is unaware of.

Click-jacking: In this attack, the attacker hijacks the clicks of the end users. The actual page is downloaded into an invisible frame which can easily trick a user into altering the setting of their account, linking other's accounts to his own account or even make him delete his own account. For example the victim tries to click on the "click me" button but instead clicked on the invisible "delete my account" button. The figure below shows a decision tree to inject malicious code through click jacking.

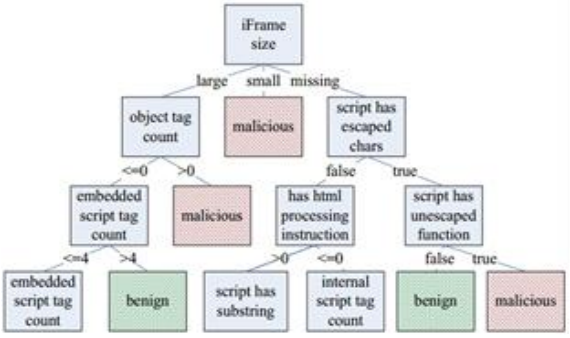

Fig. 4: Decision tree [Source: [15]]

Using servers with known Vulnerabilities: A vulnerability where the server on which the web application is hosted itself is vulnerable to attacks. This vulnerability is possible in cases where the web servers used are mostly outdated and several public exploits are available for the same. Revealing the specific software version of the server may allow the server machine to become more vulnerable to attacks against software that is known to contain security holes.

VI. CHALLENGES

To verify our data, we have performed a various web based attacks on 82 websites, exploit their web pages and we were successful in order to find a way where we can inject a malicious code or script. We responsibly disclosed, to the owners, the vulnerabilities that exist in their web pages distinguished by URL and discussed the framework that we proposed here in this paper to make a better protected environment of their organization. We reached out to 82 organisations but only 30 have acknowledged us, rest of them did not respond, some of them responded with a justification that identified flaws is not vulnerability and they do not feel that it could do any harm towards its security requirements. This Framework is appreciated by almost everyone. A few organisations state two big constraints in implementation it, one is the time required and other is the budget.

VII. CONCLUSION

This paper gives a detailed protection framework in order to protect web application from APT.

We have explained frameworks that give protection from

1. SQLIA,
2. XSS,
3. CSRF,
4. Spear-phishing and common attacks which used to launch malicious code into the targeted machine or networks.

This paper presented APT in a manner aspect that has not been dealt by others and gives customized approach to deal with such attacks or events.

REFERENCES

[1] Paul Giura and Wei Wang, “a context based detection framework for advanced persistent threat” in “IEEE ASE international conference of Cyber Security” Washington DC Dec-2012 pp 69-74 .

[2] Sam curry, Bret Hartman, David P. Hunter, David Martin, “mobilizing intelligent security operations for advanced persistent threat” RSA security brief, February 2011.

[3] Nalin Asanka Gamagedara Arachchilage, Melissa Cole “Designing a mobile game for home computer user to protect against phishing attacks” Brunel university Uxbridge, Middlesex, UK in “International Journal for E-Learning Security (IJeLS)”, Volume 1, Issues ½, March/June 2011 pp 19-27

[4] Mahmood Khonji, Youssef Iraqi, Andrew Jones “Mitigation of spear phishing attacks: a context based authorship identification framework” in “6th International conference on internet technology and secured transactions” from 11-14 December 2011, Abu Dhabi, United Arab Emirates pp 446-452

[5] Christian seifert, Ian welch, Peter Komisarczuk “Identification of malicious web pages with static heuristics” victoria university of wellington, P.O. Box 600, Wellington 6140, New Zealand.

[6] OWASP foundation (2003 - 2013) “OWASP top ten 2013” [online] Available: https://www.owasp.org/index.php/Top_10_2013-Top_10 Last modified on June 2013, Accessed on March 2014

[7] William Zeller and Edward w. felten “cross-site request forgeries: exploitation and prevention” Woodrow Wilson school of public and international affairs, Princeton University revision 10/15/2008

[8] Bruce Schneier “Attack Trees” SANS network security Conference on UNIX and NT Network Security, New Orleans, Louisiana. Wednesday, October 6th, 1999, Session

[9] Allessandra Bagnato1, Barbara kordy2, per Hakon Meland3, Patrick Schweitzer2 “Attribute decoration of attack-defense tree” 1TXT e-solutions, corporate research division, I-16100 Genoa, Italy. 2University of Luxembourg, Luxembourg. 3SINTEF ICT, Norway. International Journal of Secure Software

Engineering, volume 3(2), pages 1-35. IGI Global, 2012

[10] Shon Harris "CISSP exam guide" sixth edition [ISBN 978-0-07-178172-2], January 2013, published by Tata mc-graw hill.

[11] Gary McGraw1 and Greg Morrisett2 "Attacking Malicious Code: a report to the InfoSec research council" 1Reliable software technologies. 2Cornell University. Submitted to IEEE software and presented to IRC, May 1, 2011

[12] Paul Giura and Wei Wang, "using large scale distributed computing to unveil advanced persistent threat" in "IEEE ASE international conference of Cyber Security" Washington DC. [ISBN 978-162561-001-0] Dec-2012 pp 1-13.

[13] Tarique Mustafa "Malicious data leak prevention and purposeful evasion attacks: An approach to advanced persistent threat (APT) management" founder & chief executive officer, nexTier Networks, Inc, USA in IEEE 2013 [isbn 978-1-4673-6195-8]

[14] Johan sigholm and martin bang "Towards offensive cyber counterintelligence adopting a target-centric view on advanced paersistent therats" Department of military studies, Stockholm, Sweden in European Intelligence and security informatics conference 2013 [ISBN 978-0-7695-5062]

[15] Chistian seifret, Ian welch, peter kmoisarczuk "Identification of malicious web pages with static heuristics" victoria university of Wellington, New Zealand, IEEE 2008, [ISBN 978-1-4244-2603-4/08]