International Journal of Future Generation Communication and Networking
Vol.7, No.2 (2014), pp.183-194
http://dx.doi.org/10.14257/ijfgcn.2014.7.2.18

For Deformation Web Attacks based on Feature Recognition IPS Intrusion Prevention Technology Research

TaoYan [1] and Yi-fei Zhang[2]

[1]*Henan University of Urban Construction*
LongXiang Road Pingdingshan City Henan Province, China 467036
[2]*Henan University of Urban Construction*
LongXiang Road Pingdingshan City Henan Province, China 467036
[1]*yyt@hncj.edu.cn,* [2] *zyf@ hncj.edu.cn*

Abstract

The paper analyzes the characteristics of various types of Web application system security vulnerabilities. Based on the deformation of Web attack , according to the principle of Web application vulnerabilities occur, attack methods and targets, the attack characteristics is extended, Proposed a structural model of IPS intrusion prevention based on the feature recognition. The experiments showed that the feature recognition-based Intrusion Prevention System can ensure higher performance in high-speed attack traffic network environment.

Keywords: *Web security vulnerabilities, Deformation of the attack, Feature recognition, Intrusion Prevention*

1. Introduction

Along with the development of the Internet, e-government, enterprise portal, community BBS application, e-commerce, online shopping, online banking, securities trading and other kinds of information sharing platform based on the HTML file format (WEB) application system have deeply improved, and gone deep into the dribs and drabs in people's lives. In the web browsing, most applications are not static web browsing, but a dynamic process involving the server side. At this point, if the programmer lack of the safety awareness of programs such as Java, PHP, or ASP, which can directly bring out web application security problems endlessly [1]. In terms of the loophole attack, a attacked web application server proving Web service can leads to a Web failure service, web export attacked by D.DOS, or a paralysis of the Internet application service and so on, which an directly or indirectly bring out the huge impact.

According to the latest CNCERT/CC statistics report, website implanted backdoor attacks are increasing, stealing the site user information has been become the key point of hackers. In china, the number of being alerted site is 16388, the number of being secretly implanted into the backdoor of the website is 52324 in 2012. In 22308 fishing page in China websites, the number of the card information detected by many hackers to defraud the user is 18000. According to a sampling date, more than 14.197 million hosts overseas have been controlled by about 73000 Trojan or Botnet server. The resulting web page tampering, hang horses, confidential data leakage, such as security incidents occurred frequently, not only seriously affect the external image, sometimes even cause huge economic loss, or a serious social problem, seriously endanger national security and interests of the people.

ISSN: 2233-7857 IJFGCN

Different technologies have been used to ensure safety in all aspects of enterprise web applications. In order to protect the security of the client machine, the user may install antivirus software; In order to ensure that user data transmission to the enterprise web server safely, the user often use SSL technology to encrypt the data communication layer or use firewall and IDS/IPS to ensure that only allow specific can access, unnecessary exposure port and illegal access will be stopped in here; At the same time, enterprises can adopt a certain identity authentication mechanism authorized users to access web applications.

But even there is anti-virus protection, firewall and IDS/IPS, all enterprises still have to allow part of communication through firewall, protection measures can turn off unnecessary exposure of port, but the port of web applications must be open. The part of the passed communication may be well-intentioned, also may be malicious, which is difficult to discern. At the same time, the web application is composed of software, so, it contain bugs inevitabley, which may be exploited by a malicious user. The user may steal, or manipulate, or destroy the important information in the web application by performing a variety of malicious.

2. Web Application Vulnerabilities, Defense Technology and Defects

Web application attacking is the attacker through a browser or attack tools, send a special request to the Web server in the URL or other input area (such as forms, *etc.*), in order to discover the existence of the loopholes of Web applications, so as to further manipulate and control sites, check, modify unauthorized information[2].

2.1. Web Application Loophole Classification

2.1.1. Information Disclosure Loopholes: Information disclosure loopholes is that owing to the Web server or application does not correctly handle special requests, may leak sensitive information of the Web server, such as user name, password, source code, server information, configuration information.

There are three mainly reasons resulting in information disclosure:

There are some problems in Web server, which lead to some system files or configuration files exposed to the Internet;

Web server have loopholes by itself, if we import some special characters in the browser, we can access unauthorized files or dynamic script file source;

There are some programming problems in Web site, it may submit a request to the user without appropriate filter, direct use of the user data that is submitted.

2.1.2. Directory Traversal Loopholes: Directory traversal loopholes is that the attacker sends a request to a Web server, through in the URL or has special significance in the additional ".. / "directory, or additional".. / "some deformation (*e.g.*,".. \ "or".. / / "and its encoding), may cause the results that the attacker can access unauthorized directory, and execute the command outside the Web server's root directory.

2.1.3. Command execution Loopholes: Command execution loopholes is that through the URL request, unauthorized commands in the Web server implementation, may access to the system information, tamper with the system configuration, control the whole system, or make the system paralytic.

There are mainly two kinds of cases in command execution loopholes:

Through a directory traversal loopholes, it may access to the system folder, perform the specified command system;

Attacker submitted special character or commands, Web application doesn't test or bypass filtering Web application, analyses the requests submitted by the user, which can often lead to executive order arbitrarily.

2.1.4. File Contains Loopholes: File contains loopholes is that if an attacker sends a request to a Web server, add illegal parameter in the URL, the Web server program variable filtering is not rigid, it may handle illegal filename as a parameter. The illegal file name can be a file in the local server, or a malicious file outside. Because of these loopholes are caused by the loose filtering of PHP, so only based on the PHP Web application file contains the possible loopholes.

2.1.5. SQL Injection Loopholes: SQL injection vulnerabilities is that due to the Web application does not judge the legitimacy of the user input data, the attacker may input area (such as URL, forms, etc.) through Web page, insert special characters with a carefully constructed SQL statements, and obtain private information or tamper with the database information, through the database interaction. SQL injection attacking is very popular in Web attack. an attacker can use SQL injection loopholes to obtain the administrator privileges, add Trojan and all sorts of malicious programs on the Web, or steal sensitive information enterprises and users.

2.1.6. Cross-site Scripting Loopholes: If web programs in application haven't littered or restricted the statements and variable submitted by the user, the attacker may submit malicious code to the database through Web page input area or in an HTML page. When a user clicks a malicious code link or page, the malicious code can be executed automatically by the browser, so as to achieve the purpose of the attack.

The Harm of Cross-site scripting loopholes is very big, it is especially widely used in network bank, through cross-site scripting loopholes, attacker can access important account user pretended to be the victim, and steal important information of the enterprise.

According to a survey of previous loopholes research institutions, the universal degree of SQL injection vulnerabilities and cross-site scripting loopholes are both in the top, and the damage is so big as well.

2.1.7. The Deformation of the Web Attack: In the actual attack, the attacker usually make a deformation of Web attacks in order to escape from being attacked detection devices, such as the URL encoding technology , modify the parameters, *etc*.

2.2. Web Application Vulnerabilities Defense Technology and Defects

For common Web application loopholes and identification technology, defense measures can be taken from the following several aspects:

2.2.1. For Web Application Developers: The common reasons of most Web application loopholes are developers have not tested the user input parameters or the testing is not strict in Web application development. So, Web application developers should set up very strong safety consciousness to develop writing secure code; carry on the strict testing and limit for URL, query keywords, HTTP headers, post data and etc submitted by the user. Only accept within the scope of a certain length, using the appropriate format and encoding of characters, blocking, filtering, or ignore any other features [3]. By writing secure Web application code, can eliminate most of the Web application security issues.

2.2.2. For the Web site administrator: Website Web administrator is responsible for the daily maintenance work management, who should track and run a variety of software security patches supported by Web site timely, to ensure that an attacker cannot through software vulnerability to attacks on websites [4].

In addition to a loophole in the software itself, incorrect configuration such as the Web server, database, etc may also lead to Web application security issues. Web site administrators should make a software configuration test carefully in order to reduce the security problems.

In addition, the web administrator should also periodically audit the web server logs, detecting whether there is abnormal access to early detection of potential safety problems.

2.2.3. Using the Network Attack Prevention Equipment: The above two kinds of measures which was mentioned are ideal situations. In reality, however, the web application system loopholes are inevitable existing: part of a web site exit a large number of security loopholes, which have not been found by web developers and web administrators. Because the web application is adopted by HTTP agreement, and common firewall equipment can't defense the web attack, so we can use the IPS intrusion prevention device to realize the safety protection [5].

3. Research of IPS Intrusion Defense Technology based on Feature Recognition

3.1. IPS Intrusion Prevention Technologies

IPS key technologies includes merged global and local host access control, IDS, global and local security policy, risk management software and the controlling board supported global access and used for manage IPS. Just as IDS, IPS also need to reduce false positives or false negatives, in which many advanced intrusion detection technology are wildly used, such as heuristic scanning, content checking, state or behavior analysis, and combined with the conventional intrusion detection technologies such as signature-based detection and anomaly detection [6].

3.1.1. Two Types of IPS System:

- Host-based IPS on protected installed directly by the agent in the system. It tied together with the operating system kernel and services closely, to monitor and intercept the kernel system calls, or the API to achieve prevent and record attack role. It also can monitor the data flow and the specific application of environment (such as a web server location of the files and registry entries), in order to protect the application to avoid the signature does not exist, the common attacks.
- Network-based IPS integrates functions of standard IDS, and IDS is a combination of IPS and firewall, which can be referred to as embedded IDS or gateway IDS (GIDS). Network-based IPS equipment can only prevent malicious information flow through the device. In order to improve the efficiency of IPS equipment, the measure that information flow forced to flow through the device must be used. More specifically, the protected information flow must represent the date access to a computer system

which are among them: the specified network domain, needs a high level of safety and protection and/or the network is likely to exist in the field of internal explosion configuration address effectively the network into the smallest protection area, and can provide the largest range of effective coverage.

3.1.2. Characteristics of Web Attack IPS Technology:

- Embedded operation: only in embedded mode operation of the IPS equipment to realize real-time security protection, real-time blocking all suspicious of packets, and the rest of the data flow interception;
- In-depth analysis and control: IPS must have in-depth analysis ability, in order to determine which malicious traffic has been blocked, depending on the type of attack and strategy to determine which traffic should be blocked;
- Intrusion character library: high quality intrusion feature library is the essential condition that the IPS run efficiently, IPS should also periodically update library invasion characteristics, and quickly applied to all sensors;
- Efficient processing capacity: IPS must have the ability to efficiently handle packet, affect the performance of the entire network to a minimum.

Along with the enterprise security protection in the process of the attention to the application layer, based on the deep packet inspection (DPI) technology and the depth of flow detection technology (DFI) become one of hot technology in the field of security.

At present, based on port application protocol recognition is the most usual means. Such as find a data message in the source or destination port for 80, is considered to be the HTTP protocol related message, to the HTTP protocol analysis engine test protocol decoding and attack. But with a variety of network applications gradually rich, this kind of method based on port protocol type to identify the packet belongs to expose its disadvantages [7]:

- Administrator for avoiding risk and so on reasons, as prescribed by the application of some common USES of RFC ports (such as HTTP using 88, 8080), port and mapping applications do not correspond, causing the application unable to recognize or identify mistakes.
- Some new application does not use fixed ports (such as SIP, etc.), but adopt the way of dynamic negotiation in the process of the protocol run, such as creating data channel. This case USES the port mapping to identify application doesn't work.
- In recent years, peer-to-peer (P2P) protocols emerge and evolve. They are in the process of the protocol run, data packet encryption methods such as dynamic negotiation, to the application of recognition is put forward more severe challenges.
- All kinds of application version updates frequently, which requires the application of identification must keep dynamic refresh in time. So keep the application identification has become an important problem need to be solved timely.
- In the face of these new applications, often need to take agreement automatic identification algorithm for identification. How to satisfy the demand for accuracy and high performance?

In response to the fixed end is used to identify the protocol of defects, in actual use process, with DPI and DFI two main techniques:

Firstly, DPI (Deep Packet Inspection), namely the deep packet inspection. Make a comprehensive judgment according to the different application protocol of “fingerprints” based on the analysis of the message header.

Secondly, DFI (Deep Flow Inspection), namely the depth of Flow detection. It is a kind of applications of identification technology based on traffic behavior. Different application types reflect on the condition of session connection and data flow is different.

3.1.3. Advantages and Disadvantage of these Two Technologies: Based on the recognition method for in-depth study of existing agreements, as well as to the network protocol based on the profound understanding of, the combination of the above two Application identification technology, We proposed a general protocol identification model - UAAE (Universal Application Apperceiving Engine). Applications in the identification model, the application for recognition can be divided into the following categories [8]:

- Fixed port agreement: agreement such as BGP, RIP, etc., the port is relatively stable, can according to the port number for the rapid identification.
- Characteristic state machine discover protocols: most of the P2P protocol port is not fixed, some even deliberately USES some standard protocol of well-known ports, such as BT, Emule, thunderbolt, skype, *etc.*, will be using port 80 for protocol interaction, so we need to rely on in-depth data analysis to identify the application protocol. Different protocols have different "fingerprints", these "fingerprints" can be load port, part of the fixed string or binary sequence, such as BT handshake Protocol message characteristic word for ".BitTorrent Protocol". Unlike other applications identification system, UAAE not just rely on a single packet handshake is used to identify the application protocol characteristics, also by the characteristics of the state machine for a more accurate identification [9].
- Negotiation protocol: currently, more and more agreement USES the control channel and data channel model: landing control channel used in interaction, data link and command interaction, *etc.*, through the control channel can negotiate a data interaction of one or more data channels. Traditional FTP protocol belongs to the model, the vast majority of VoIP applications also belong to this model, such as Skype, UUCALL, *etc*.
- Tunnel protocol: firewall and NAT equipment deployment makes in the network, there are many application layer tunnel, the tunnel is the result of the application protocol level between nested. Such as HTTP Tunnel, the surface is a 80 port connection, but, in fact, there may be any kind of application data.
- Flow model discovery protocol: a growing number of P2P traffic USES encryption transmission, such as thunderbolt, Skype, *etc.*, through the application is unable to identify the "fingerprints" way. Only through flow characteristics, such as connection rate, the relevance of each link and data transmission byte distribution and flow characteristics for identification.

In UAAE model, for the above category 1-4 DPI technology is adopted to improve the identification of a specific application; For DPI technology can't identify the applications and data flow, the flow through the DFI) technology of recognition, namely the category 5.

3.2. The IPS Intrusion Defense Model based on Feature Recognition

In the actual attack, the attacker to escape attack detection of devices, usually to deformation of Web attacks, such as the URL encoding technology and modify the parameters, *etc.*, based on the principle of Web application vulnerabilities occur, attack methods and attack targets, t extended he characteristics. Even against attacks from the attacker to modify the parameters, the format, statement, *etc.*, the same hole deformation principle under the various attacks can also be effectively blocked [10]. This makes IPS extended their defense flexibility also significantly enhanced, greatly reduce the omission of situation.

3.2.1. UAAE Model: In the TCP/IP protocol suite, the agreement is a hierarchical relationship, overall as a tree, as shown in Figure 1:

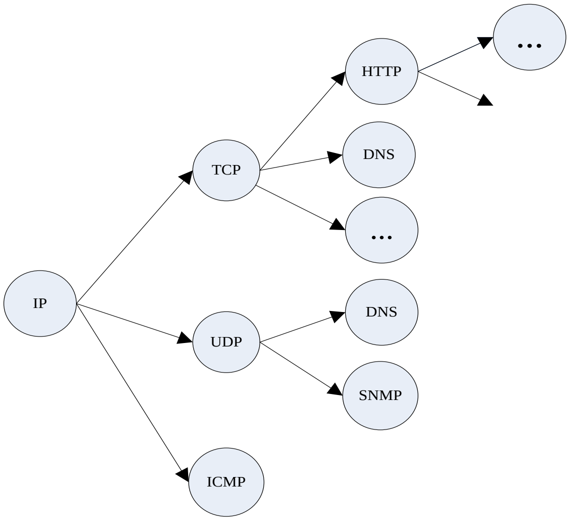

Figure 1. IP Protocol and Protocol Characteristics of Trees

For each protocol that can be used to identify the "fingerprints". Application "fingerprint" namely protocol characteristics, it can be for one or more of the definition, characteristics of tree form the agreement also. For each of the features you can specify the corresponding parent agreement (agreement), UAAE through a complete set of grammar to describe the identification method of agreement.

Traditional way of protocol parsing is combined with the protocol header parsing, layer by layer to obtain the upper protocol, this is UAAE derived the basic idea of agreement [11]. And build the application protocol based on the tree and the tree of protocol characteristics, the application protocol agreement with the actual hierarchical structure tree is completely consistent.

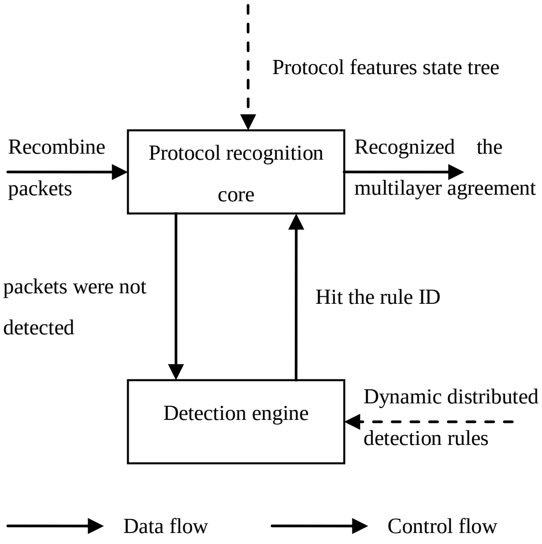
Protocol features state tree
Recombine packets
Protocol recognition core
Recognized the multilayer agreement
packets were not detected
Hit the rule ID
Detection engine
Dynamic distributed detection rules
Data flow
Control flow

Figure 2. UAAE Engine Basic Architecture

As shown in Figure 2, packet core after restructuring to the protocol identification module, the module data packets to the detection engine test; Detection engine according to the received packets characteristic tree for dynamic distributed protocol for testing, then to the protocol identification of core modules back hit test results; Protocol to identify core modules according to the test results and its internal protocol feature tree hierarchy are derived and identify the corresponding agreement. The detection engine can be built by ASIC or FPGA chips etc, so as to realize the detection of high performance. At the same time, if the agreement to identify the core module found a session after several messages still haven't learned the most upper layer applications, will the DFI technology on flow depth analysis and recognition, so as to determine the corresponding unknown flow belongs to the application.

UAAE engine in combined application on the basis of the model classification and recognition, can also effectively flexible intelligent decision making. It can identification method and validation methods for priority, high priority identification results of dynamic intelligent replacement low priority identification results, make the application of the result of the recognition accuracy is greatly increased. Login Skype, for example, using TCP port 80 protocol, the TCP handshake, UAAE according to port to identify the session is an HTTP session. Further testing the content of the session, once you identify the Skype login feature, UAAE can make judgment according to priority immediately intelligently, the session identification for skype login protocols.

3.2.2. Web Attack Defense Framework: IPS intrusion prevention device has a complete Web attack defense framework, can timely find various was exposed and the potential Web attacks. For Web attack the overall defense framework is illustrated below.

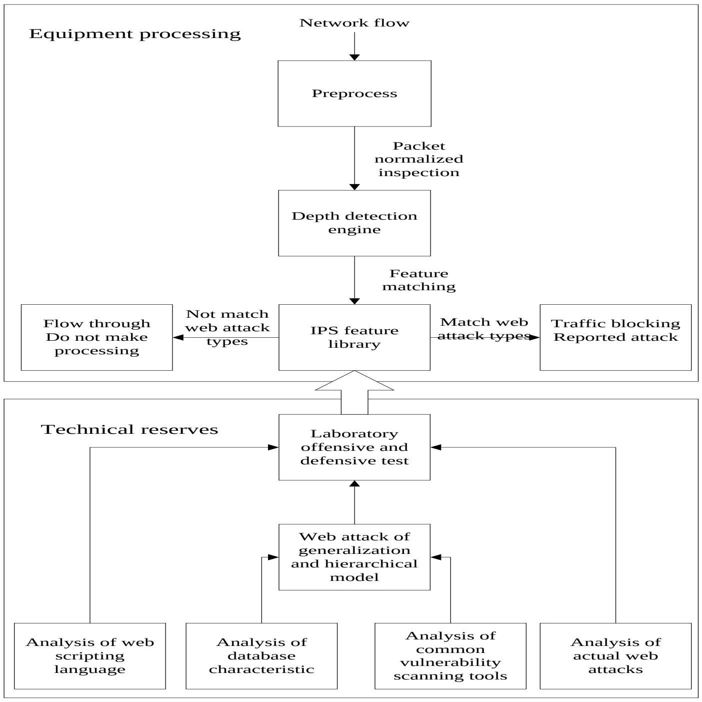
Equipment processing
Network flow
Preprocess
Packet normalized inspection
Depth detection engine
Feature matching
Flow through Do not make processing
Not match web attack types
IPS feature library
Match web attack types
Traffic blocking Reported attack
Technical reserves
Laboratory offensive and defensive test
Web attack of generalization and hierarchical model
Analysis of web scripting language
Analysis of database characteristic
Analysis of common vulnerability scanning tools
Analysis of actual web attacks

Figure 3. Web Attack Defense Framework

This Web attack defense framework has the following features:

Firstly, it can complete web attack detection model structure, accurate identification of various web attacks. According to the characteristics of web attacks, considering the various Web attack principle and form, on different vulnerability model developed a generic, hierarchical web attack detection model, and integration to the features in the library. The model abstracts the general form of web attack can accurately identify the attack on the mainstream, so as to make the model generalization.

Secondly, the detection method can accurately identify the deformation of web attacks.

Thirdly, it can ensure tracking the latest vulnerabilities and technology, effectively prevent the latest attack. As web attack frequency of rising, the harm has a tendency to spread gradually. The IPS equipment in defense in depth and breadth of put forward higher request, not only to defend on an existing web attacks, more effective to prevent the latest attack, unpublished. At present, most IPS equipment manufacturers have established a complete set of offensive and defensive test environment, and can timely found potential web security vulnerabilities. At the same time, continue to track the latest web attack techniques and tools, timely update the web attack the feature library, the first time the latest web vulnerability response, ensure the user's network from attacks.

Lastly, it can ensure efficient operation of the normal business. Detection engine is the key to IPS the equipment operation, the engine uses a more efficient and accurate detection algorithm, and deep analysis for traffic passing through the device, and by matching and attack characteristics, whether there is any abnormal test flow. If flow rate is not matching to the attack features, is allowed to flow through, will not interfere with the normal business network, the accurate defense at the same time to ensure the efficient operation of the normal business.

3.3. System Deployment

By the deployment of intrusion prevention system, not only can effectively resist the external aggression, but also can implement comprehensive safety monitoring. For illegal attacks from external and internal network, you can through the network intrusion prevention system network cascade on the trunk line network, real-time data traffic to stop all kinds of illegal network attacks, the attack behavior to stop outside the network gateway, implementation of the protection of information assets in a network. To come from inside the network attack, also can through by-pass network intrusion prevention system, in the key parts of the network, the network security situation of internal monitoring and analysis, to filter out office in all sorts of virus and Trojan attacks, ensure the safety of network.

3.4. Experiment and Result Analysis

In order to assess the effectiveness of the proposed model, and through the NIPS system based on the OCTEON multi-core platform (hereinafter referred to as the M - NIPS), and on the basis of the improvement implemented a IPS intrusion defense model based on feature recognition systems, hereinafter referred to as (FR - IPS). This attack strength and forward delay through the network to analyze the relationship between the working efficiency of the model, and let 10 units attack the host through the Blade Informer software according to the different contract rate at the same time send Dos attack simulation data to the internal host behind NIPS. Figure 4 shows the M - NIPS and FR - IPS with the increase of intensity of network attack, their forward the change of time delay. Can see the M - NIPS in attack power reaches 400 packet/s network, packet forwarding delay began to increase rapidly, and FR - IPS forwarding delay is not due to the increase of attack traffic increased, when the packet/s 1600 attack power, is to make sure that the packet delay is less than 100 ms. The experimental results show that FR - IPS attack traffic in high-speed network environment still can guarantee the high working performance.

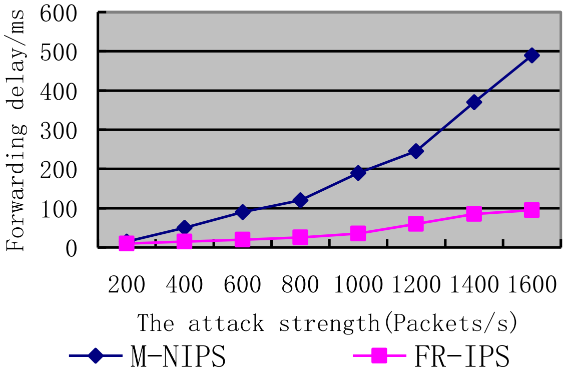

Figure 4. Attack Power and Forward Delay Variation Relationships

4. Conclusion

Along with the application of the web technology, the web application security is facing the challenges increasingly, web systems always suffer from all kinds of attacks, in this case, it is necessary to develop a complete web attack defense solutions, through the security of web applications, web server software, web attack prevention equipment mutual coordination, to ensure the safety of the site. This paper proposes a IPS intrusion defense structure model based on feature recognition, according to the deformation of web attacks, according to the principle of web application vulnerabilities occur, attack methods and attack, the attack characteristics is expanded, even against attacks from the attacker to modify the parameters, the format, statement, etc, the same hole deformation principle under the various attacks can also be effectively blocked. It makes defense scope of IPS, defensive flexibility also significantly enhanced, greatly reduce the omission, experiments show that the attack traffic in high-speed network environment IPS intrusion defense system based on feature recognition still can guarantee the high working performance.

References

[1] D. Stuttard and M. Pinto, “The web application hacker's handbook: discovering and exploiting security flaws”, Wiley. Com., **(2007)**.

[2] F. Larue, M. Di Benedetto, M. Dellepiane and R. Scopigno, “From the digitization of cultural artifacts to the Web publishing of digital 3D collections: an Automatic Pipeline for Knowledge Sharing”, Journal of multimedia, vol. 7, no. 132, **(2012)**.

[3] F. Zhang Yong-Zheng and B. X. C. H. I. Yue, “Survey and Evaluation on Computer Vulnerability Database”, Computer Science, **(2006)**.

[4] H. Shi, B. Chen and L. Yu, “Analysis of Web security comprehensive evaluation tools”, Proceedings of 2010 International Conference on Networks Security, IEEE CPS, Wuhan, China, **(1915)** April 285-289.

[5] K. Yang and F. Jiang, “Study on the uniform description of security vulnerabilities”, Computer Engineering & Science, vol. 28, no. 11, **(2006)**.

[6] J. Du and Y. Lu, “Taxonomy of Web-based application vulnerabilities”, Computer Engineering and Applications, vol. 45, no. 10, **(2009)**.

[7] T. Wang, Y. Li and Y.-B. Sheng, “Honeynet-based Network Security Defense Model”, Application Research of Computers, vol. 26, no. 3012, **(2009)**.

[8] H. G. Kayacik, A. N. Zincir-Heywood and M. I. Heywood, “A hierarchical SOM-based intrusion detection system”, Computer Engineering and Science, vol. 20, no. 439, **(2007)**.

[9] Y. Du, J. Liu, R. Zhang and J. Li, “A Dynamic Security Mechanism for Web Services Based on NDIS Intermediate Drivers”, Journal of Computers, vol. 6, no. 2021, **(2011)**.

[10] A. Mohammed Al-Canaan and A. Khoumsi, “Multimedia Web Services Performance: Analysis and Quantification of Binary Data Compression”, Journal of Multimedia, vol. 6, no. 447, **(2011)**.

[11] M. T. Qassrawi and H. Zhang, “Detecting Malicious Web Servers with Honeyclients”, Journal of Networks, vol. 6, no. 145, **(2011)**.

Authors

Tao Yan, born in Luoyang City, Henan Province,china, in January, 1970,graduated from Shenyang University with the Master degree of information management in June 2007, shenyang. He has been a teacher in Henan University of Urban Construction from 1991. he is a Associate professor.

Yi-fei ZHANG, born in Pingdingshan City,HenanProvince,china, in November, 1979,graduated from Huazhong University of Science and Technology with the Masterdegree of Science in December 2007,Wuhan. He has been a teacher in Henan University of Urban Construction from 2003. he is a lecturer.