ISSN (Online) : 2319 - 8753
ISSN (Print) : 2347 - 6710

International Journal of Innovative Research in Science, Engineering and Technology

An ISO 3297: 2007 Certified Organization   Volume 3, Special Issue 5, July 2014

International Conference On Innovations & Advances In Science, Engineering And Technology [IC - IASET 2014]

Organized by

Toc H Institute of Science & Technology, Arakunnam, Kerala, India during 16th - 18th July -2014

Resisting Proxy-Based Attacks Based On Temporal and Spatial Locality Behaviour

Sai Subha.V[1], Arathi Gandhi[2]

Second Year,Mtech, Dept of Computer Science, Mohandas College of Engineering & Technology, Anad ,Tvm, India[1]

Assistant Professor, Dept of Computer Science, Mohandas College of Engineering & Technology,Anad,Tvm, India[2]

ABSTRACT: In computer networks, a proxy server is a server that may be a computer system or an application that acts as an intermediary for requests from clients seeking resources from other servers**.** In this paper, it mainly focuses to resist proxy-based DDOS attacks. It extracts the behavior features of traffic from proxy to server based on temporal and spatial locality. A hidden semi-markov model is used to describe traffic behavior of proxy to server. Here, it assessed the short-term behavior for improving quality of services of normal users. It converting suspicious sequence into normal one by discarding most illegimate request, instead of discarding entire sequence.

KEYWORDS: Traffic analysis, traffic modelling, DDOS attack, detection of attackers and response.

I. INTRODUCTION

Denial-of-Service (Does) attacks continue to be key threat to Internet applications. In such attacks, especially distributed Dodos attacks, a set of attackers generates a huge amount of traffic, saturating the victim's network, and causing significant damage. Overlay networks have been proposed to protect applications against such DDOS attacks .These overlay networks are also known as proxy networks. The key idea is to hide the application behind a proxy network, using the proxy network to mediate all communication between users and the application, thereby preventing direct attacks on the application. The Web proxy-based HTTP attack is more flexible than most of existing DDOS attacks. It is difficult to detect the attack. The detection can occur in several ways[1] i.e., real attacking host are unobservable by origin server as they are shielded by proxies, proxy can passively involved in an attack, may act as an attacker and also web server is difficult to find, because both legal and illegal traffic came from web proxies. Also web proxy's can turned into an attacker [1].Attacker first sends request to proxy and it forward request to web server. In this step, it can turn into an attacker by requesting dynamic documents or setting cache-control: no-cache in headers of http request.

In web proxy-based DDOS attack due to two reasons, first, the attacking hosts are unobservable to origin server as they are shielded by the web. Proxies. Second, traffic comes from the same sources (web preserver is There are several image, proxies).In the final Proxy to server traffic, it contains both legal and illegal traffic, so the server is hard to find and filter the malicious request accurately. This paper analyses the traffic behaviour of proxy-to-server is based on hidden semi-markov model [6].Hidden semi-markov model describes the varying observable states of traffic from proxy-to-server. It describes the transformation of the proxy's behaviour states. Based on this behaviour, it can assess the abnormality of web proxy. It can do by measuring the difference between the observed behaviour and proxy's historical behavior algorithms. Abnormality of proxy can checked based on long-term and short-term behaviour assessment. Long-terms assessment based on large scale, while short-term assessment detects abnormal request from proxy-to-server traffic. Here,it proposed the soft-control scheme for attack response[1]. It reshapes suspicious sequence into normal one by discarding most malicious request, instead of discarding entire sequence**.** Thus, it can protect quality of services of legitimate users without discarding the entire sequence.

II. RELATED WORKS

Distributed Denial of Service (DDOS) attacks are becoming an increasingly disturbance of the global Internet. It focus on the anomaly detection on user browsing behavior[8].Web user behavior is influenced by structure of Website (e.g., the Web documents and hyperlink) and the way users access web pages. Web user's access pattern can formulate in terms of hidden semi-markov model and p-algorithm. P-algorithm is used for dynamic generation of CAPTCHA to web pages. If an abnormality is found in the incoming request, user is served with captcha to specific web pages. It detects anomaly of user browsing behavior based on http request and page viewing time. A low-rate distributed denial of service (DDOS) attack has significant ability of concealing its traffic. [4], it is based on generalized entropy and information metrics to detect DDOS attack by measuring difference between legal and illegal traffic. The information metrics can effectively detect attacks early and detect the accuracy (false positive rate). Also, the IP traceback algorithm is used to find all attacks as well as attackers from their own local area networks (LANs) and discards attack traffic. It can detect low rate DDOS attacks.

In [6], HsMM has variable states and also associated with duration of each state. HsMM is a stochastic process with a discrete time finite-state homogeneous Markov chain. The state sequences are not observable and is called hidden. It influences another process that produces a sequence of observations. In [3], the sequence order of web page requests can be used for detecting DDOS (App-DDOS) attacks. The attributes of web page request are extracted without considering the web page request sequence. It considers a pca model for formulating web browsing behavior. The reconstruction error is used for discriminating DDOS attacks. In [2], a method is used for discriminating flash crowds from DDOS attacks by using flow correlation coefficient. DDOS attacks possess higher similarity with that of flash crowds. It used detection method ie, flow correlation coefficient used to measure the similarity among suspicious flows to differentiate DDOS attacks from flash crowds.

It used the Pareto law and Gaussian distribution for evaluating flow correlation coefficient. Flow correlation coefficient decreases if the attack flows come from different botnets. Existing algorithm requires large computation for estimating the parameters of hidden semi-markov model. In [9], a forward-backward algorithm is used for profiling web access behavior with explicit duration and estimate parameters .It requires only computations and memory capacity to evaluate all the forward and backward variables, where is the number of Markov states is the maximum duration between successive state transitions. In [4], an effective IP traceback scheme against DDOS attacks based on entropy variations. It works on IP traceback depend on packet marking, i.e. is either probabilistic packet marking or deterministic marking. It identifies the DDOS attacks via, detection algorithm.

.

III.PROPOSED SYSTEM

The proposed scheme is to protect the origin server from the web proxy-based http attacks. The incoming traffic comes from the same sources (web proxies) consist of both legal and illegal traffic. The goal is to discriminate the most malicious request, instead of denying the entire sequence. With the profiling of web access behavior, it can be regarded as a combination of external and intrinsic driving mechanism[1]. The external manifestation includes the temporal and spatial locality, while the intrinsic mechanism includes normality or abnormality .Structure of HsMM [6] is shown in fig.1. The external manifestations can assessed based on behaviour of proxy-to-server traffic. These are observable, but it is controlled by intrinsic mechanism and cannot be accurately obtained by the origin server. The web proxy's access behaviour can be mapped based on HsMM model. HsMM [6] has a pair of stochastic process, with variable states and also associated with duration of each state. These processes are not observable. HsMM is used to model a web proxy's behaviour, each hidden markov states represents driving mechanism of proxy-to-server traffic.

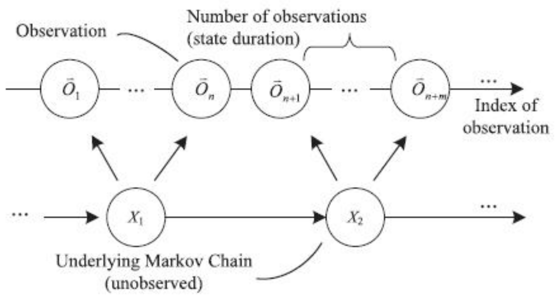

Fig.1. Structure of HsMM

Transition of one markov states to other states represents the changes of driving mechanism. Duration is associated with transition of one state to another state. When the incoming traffic contains suspicious request, then web-proxy's behaviour is equivalent to abnormality state, then it tried to filter those suspicious request from that of normal users. HsMM model is used to describe the stochastic process of proxy's behaviour in order to reduce the parameters. The states output distribution and state duration is parameterized by using GGHsMM.GGHsMM [1] is mixture of Gaussian and gamma distribution. Gamma distribution is flexible and it reduces computational complexity.

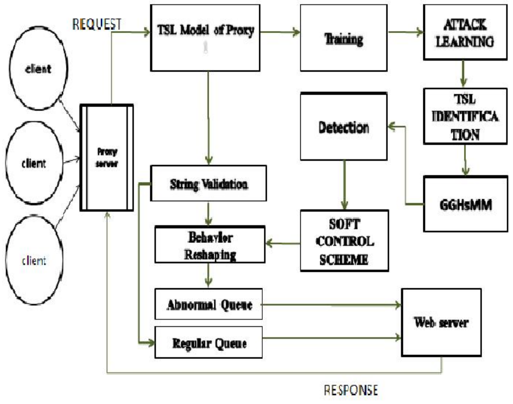

Fig.2. System Architecture

IV. IMPLEMENTATION OF THE SCHEME

A. Temporal and spatial Locality

Web access behaviour can be regarded as combination of external manifestations(e.g. temporal and spatial locality).Temporal locality [10] refers to property that a reference behaviour to be seen in future, whereas the resource metric represents only the frequency of request without indicating the correlation between a reference to a document and time since it was last accessed, Spatial locality [7] refers to property that an object frequently accessed in past are likely to be accessed in future. For example, if a home page is requested, all its possible sequence is likely to be accessed at same time. It denotes the correlation among a sequence of http request.

B. Extraction and Training of Data

The scheme includes three phases: data extraction, training of data, and detection and control of data. System architecture of the scheme is shown in fig.2. Data extraction consist of the extraction of the incoming sequence i.e., it extracts the proxy’s TSL’s and generates a TSL string. It is then handled by nonlinear mapping function and thus formed the final observed process. In training of data, various sequence are extracted, then constitute the different probabilities of sequences are generated by mining the previous history in an array. The input given is the browsing option for available sequence. The output is available path from a particular sequence. Let T be the total number of time windows. Each time window consists of observation sequence. The two parameters are used to measure the normality of proxy’s behavior [1] i.e., behavior index and structure factor. Behavior index denotes total number of request and observation sequence of the T time window, whereas the structure factor denotes the number of request generated by state n of the T time window.

C. Detection of Attackers and Control

The last phase is the detection of data, i.e., detection of attackers .It can be diagnosis by long term and short term observation sequence of the incoming request [1]. In long term observation, normality is evaluated by comparing its consecutive observation sequence in the incoming request. But it works on large time scale. Then, it goes for evaluating the normality based on short term behavior. The incoming request consist of both the legal and illegal traffic coming from the web proxies. If any abnormality is found in the sequence, then it reject the entire sequence. In order to avoid the delay of legitimate users, a soft control [1] scheme is used to protect the quality of services. It reshapes or filter the suspicious sequence from the normal one by partly rejecting the malicious request instead of rejecting the entire sequence. It is done based on algorithm reshaping algorithm. The attackers were controlled by using the firewalls. It provides the accuracy of the system by detecting the attack based on a given threshold under different false positive rate and false negative rate.

D. Analysis of the Scheme

The proposed scheme is implemented in a simulation environment.The proxy-based attack is simulated by NS2.The performance of detection of various attack scenarios. The proxy will compromised, when various attackers enter with the incoming sequence. The attackers were denied and it will allow only the legitimate users by using secure mechanisms. Thus, it control the attackers by setting firewalls. On the other hand, if it use the insecure mechanism, suspicious will enter and there is no detection of attackers takes place.

V. RESULTS

From the experiment, it shows that various attack scenarios based on temporal and spatial locality behavior is analysed.It shows that the incoming sequence from proxy to server is mixed with normal and polluted traffic .The detection of attackers from proxy to server based on tsl behavior was analyzed by using secure and insecure mechanism and plotted on a graph.

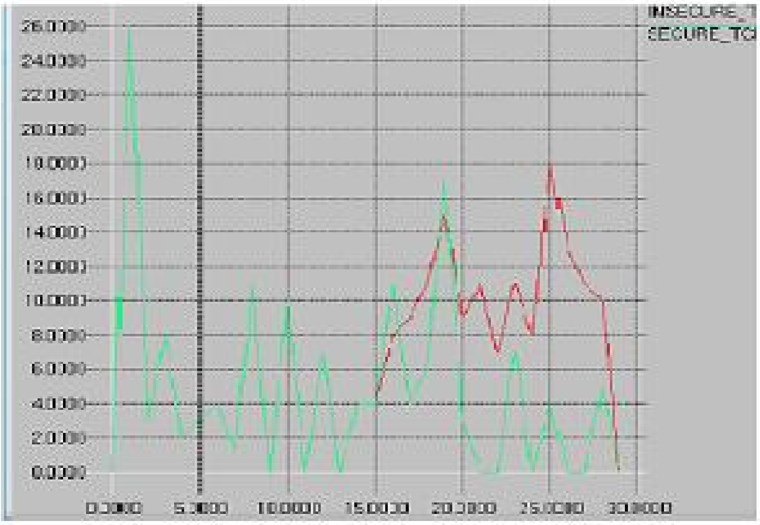

Fig.3.Detection of attackers by using secure and insecure mechanism

From the figure.3, it shows that detection of attack scenarios based on incoming traffic from the proxy to server is plotted in the graph. The red line indicates the detection of attackers in insecure mode, whereas the green line indicates the detection of attackers in secure mode.

It proposed to enhance that the problem with this approach is that data packets from a genuine proxy will also get blocked if there is a sudden and unusual increase in the data volume sent by the proxy. This is an unwanted situation. This can be overcome by analysing the content of the data being sent by the proxy. Many often certain patterns can be identified in case of false data being sent for denial of service. Thus only such packets are dropped even if there is a surge in the packet transfer rate.

VI .CONCLUSION

In this paper, it tried to detect the early detection of attackers by filtering the suspicious traffic from the final aggregated proxy-to-server traffic. It extracts the TSL's behaviour of web proxies and filtering by means of behaviour reshaping. Detection performance is improved by decreasing false positive rate and false negative rate.

REFERENCES

[1] Yield Xian and S.Tang,” Resisting we b proxy-based http attack by temporal and spatial locality behavior, “Transaction on parallel and distributed system, July 2013, vol.24, no.27.

[2] S. Yu, W. Zhou, W. Jiao, S. Goo, Y. Xiang, and F. Tang, “Discriminating Dodos Attacks from Flash Crowds Using Flow Correlation Coefficient,” IEEE Trans. Parallel and Distributed Systems, vol. 23, no. 6, pp. 1073-1080, June 2012...

[3] S. Lee, G. Kim, and S. Kim, “Sequence-Order-Independent Network Profiling for Detecting Application Layer DDos Attacks,” EURASIP J. Wireless Comm. and Networking, vol. 2011, no. 1,p. 50, 2011

[4] S. Yu, W. Zhou, R. Doss, and W. Jia, “Traceback of DDos Attacks Using Entropy Variations,” IEEE Trans. Parallel and Distributed Systems, vol. 22, no. 3, pp. 412-425, Mar. 2011.

[5]. “Low-Rate DDos Attacks Detection and Traceback by Using New Information Metrics,” IEEE Trans. Information Forensics and Security, vol. 6, no. 2,pp. 426-437, June 2011.

[6] S. Yu, “Hidden Semi-Markov Models,” Artificial Intelligence, vol. 174, no. 2, pp. 215-243, 2010.

[7]] Y. Zhong, X. Shen, and C. Ding, “Program Locality Analysis Using Reuse Distance,” ACM Trans. Programming Languages and Systems,vol. 31, no. 6, p. 20, 2009..

[8]. Y. Xie and S. Yu, “A Large-Scale Hidden Semi-Markov Model for Anomaly Detection on User Browsing Behaviors,” IEEE/ACM Trans. Networking, vol. 17, no. 1, pp. 54-65, Feb. 2009.

[9]. S.-Z. Yu and H. Kobayashi, “An Efficient Forward-Backward Algorithm for an Explicit-Duration Hidden Markov Model,” IEEE Signal Processing Letters, vol. 10, no. 1, pp. 11-14, Jan. 2003..

[10]. A. Mahanti, D. Eager, and C. Williamson, “Temporal Locality and Its Impact on Web Proxy Cache Performance,” Performance Evaluation, vol. 42, nos. 2/3, pp. 187-203, 2000.