International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Index Copernicus Value (2013): 6.14 | Impact Factor (2013): 4.438

Protection of Server from Proxy-Based DDoS Attack

Poonam U. Patil[1], Dr. Y. V. Chavan[2]

[1, 2]Department of Computer Engineering, PVPIT, Bavdhan, Savitribai Phule Pune University, Pune, India

Abstract: ***Number of intermediate node i.e. proxy servers are lies between client and server. Due to this hierarchical organization of web proxy, existing system those were considered for DDoS attack detection was not able to preserve the Quality of Service of typical user as those are rudely rejects the all demands on DDoS attack detection. This affects on the quality of service of normal user because traffic from last proxy-to server may consists of attack traffic as well as regular traffic. Therefore; it is indispensable to ensure the QoS to normal user. Server side defense scheme that detects DDoS attack is implemented. This system protects the origin server from proxy based HTTP as well as HTTPS attacks. Soft control response mechanism solves the issues of QoS. It uses Hidden Semi Markov Model (HsMM) to analyze the traffic behavior. Pre-filtering at entry point is also proposed in this work to minimize the time require to check normality or abnormality of requests.***

Keywords: Distributed denial of service, Hypertext transfer protocol, temporal locality and Spatial locality, Hidden semi-Markov model.

1. Introduction

A Distributed Denial of Amenity Attack (DDoS) is dicer to the internet. It is broadly spread on over network through wireless or wired network. It always debilitates preparing limit, memory, bandwidth on the other network. The attacker send the packet to the Maltreated person a low level DDOS attack is smart attack which is difficult to detect. The low level DDoS is distributed all over network and join with the more DDoS attackers so it is more difficult to identify and detecting the attackers. The measurement identification of low level DDoS attacks are mostly divided in to two classes: a] the anomaly based metrics b] signature based metrics: The signature based metric attacks sends the predefined set of signatures. For example the string is send as the signatures and match to the packets. The anomaly-based reorganization metrics has more constraint and initially the attackers prepare recognized framework for progressive acknowledge anomaly network conducted by typically. The rate of false positive utilizing anomaly based reorganization metrics is greater than the rate of utilizing signature based identification metrics. The false positive and the false negative rate of identification are very difficult to identify the correct limit. It is extremely difficult to point out the feature of the typical and anonymous networks parties. Example of an anomaly based detection metric which uses individual limit from predefined set of limits is anomalous deviation of some measurable attributes from network traffics to find the unusual traffic from all ordinary traffic. For this we use decision of statistical technique and devices are difficult [4]. The Gaussian noise technique is mostly used to reproduce the genuine network traffic it is the most part of the acknowledgement. In DDoS attack traffic collection the accumulation and Poisson distribution technique is used [5].

The homogeneous Markov chain is the discrete time limited state in the stochastic procedure. The Hidden Markov Model (HMM) is categorized in to doubly stochastic methodology. The state grouping is hidden and it impact on alternate stochastic process that deliver succession of perception. In understanding Bayesian network the exchange of learning and induction in HMMs is mostly used. The HMM is a imperative class model which is effective in application domain.

HMM has the limit in few application domains due to its verifiable geometric circulation. A Hidden-semi Markov Model (HSMM) is permitted the some fundamental methodology on the Hidden-semi Markov Chain. The HSMM is also called as the "express spam HMM", "variable term HMM", "segmental HMM" .Some kinds of attacks are adventures and vulnerable at application layer instead of network layer but not all DDoS attacks. There is no spoofed IP address with amenities like HTTP and HTTPS. The application DDoS attacks send little honest packet .App-DDoS attacks are takes place on the crowd when large quantity of request to the web server [6].It is hard to detect the App-DDoS from ordinary traffic. There are few DDoS system that uses application layer data [7].

Another "soft-control" plan is presented for assault reaction. The plan reshapes the suspicious groupings as determined by the profile of a proxy's historical behavior. It changes over a suspicious succession into a normally typical one by halfway disposing of its doubtlessly vindictive requirements as opposed to denying the whole grouping. In this way, it can confirm the HTTP & HTTPs appeals of real clients to the best degree convincing from being disposed of[1]. In summing up contradict and the huge part of the current and past works [15], [16], the oddity of this work lies in:

1) It is centered on conflicting Web proxy-based HTTP & HTTPs assaults and understands the early discovery without any participation of mid Web alternatives;
2) The technique is free from frequently changing Web substance. It need not often change model parameters.
3) Long and short behavior evaluation plans empower the multi-granularity determination, while the "soft-control" plan can enhance the quality of service of normal users.
4) Pre-filtering reduces time for processing large number of requests by means of entry point blocking.

2. Related Effort

The DDoS examination is concentrated on IP layer this technique is used to detect the attacks by breaking some feature e.g. arrival rate and header data. The cross connection examination is over in [8] to catch traffic example Time-to-live values and statistical methodology [9] is used for detecting the DDoS attacks[10].The Net-DDoS attacks and App-DDoS attacks does not depends on the frameworks firewall dispatch the attacks when they using the opening amenities like HTTP and HTTPS. Several protocols and application are used when passage through firewall by interfacing standard TCP protocol using port several 80.The attackers demand amenities these request of amenities passes through firewall without distinguished at the same time other customer can't complete their transaction. There are some following factor 1] the Net-DDoS attacks identification strategy are not able to detect the App-DDoS attacks identification at the ground level they fit with separate layer.2] TCP anomaly reorganization method can separate the App-DDoS attacks by passing the HTTP request through TCP protocol.3] to create TCP connection attackers required the IP address and IP packets which can be used to create anomaly identification for IP packet get to be invalid.4] the main reason behind discovering the plan is quality of DDoS attacks vary from ordinary traffic falls on grounds. The foundation of traffics of this situation is expected as steady the App-DDoS attacks can enhancing surviving Net-DDoS attacks e.g. They can calculate the HTTP request rate, HTTP session rate and terms of clients used for detection.

HTTP based DDoS attacks is tolerating more thought the client is evaluated by the trusted organization framework and after that IP address is soothed at the application layer by offering to customer. The outlier are recognized by using advance technique after that HTTP suspected are blocked[11].A four characteristics of web site page request model is evaluated at the profile of normal Access performance [12].The propagation mix up action is used to finding DDoS attacks. Coefficient of relationship flow is used to identify the likeness of affected client after that the result of estimation is separated on the HTTP based DDoS attacks [13]. Multidimensional access structure is used to catch the performance of swarms and detect the HTTP attacks that duplicate unmistakable web page[16].All these arrangement is to connecting the mislaid server the mislaid server precise progress move by the host and host progress does not fit to the predefined instruction of given model. on the other hand portion of the terminal host are at different level of the web proxy server are consequently HTTP request are saw by the misused individual server is typically last web proxy server that interface with individual server and can be confirmed by individual server[1].

System the proxy server operates to send the request to the web server. IP packet size entropy (IPSE)-based DoS/DDoS detection scheme was developed in which the entropy is markedly changed when traffic is affected by an attack. From investigation, they concluded that the IPSE-based scheme is able of detecting not only long-term attacks but also short-term attacks that are beyond the volume-based schemes' capability to identify. Introduced proposed was tested by using two typical Internet traffic data sets from DARPA and SINET, and the test results shown that the IPSE-based detection scheme can provide detection of DoS/DDoS attacks not only in a local area network (DARPA) and but also in academic backbone network (SINET). Countering Distributed Denial of Service (DDoS) attacks is becoming ever more dangerous with the large number resources and techniques increasingly accessible to attackers. Introduced exploration was focused around complicated attacks that are protocol-compliant, non-intrusive, and utilize legitimate application-layer appeals to overwhelm system resources. It distinguish application-layer resource attacks as either request flooding, asymmetric, or repeated one-shot, on the basis of the application workload parameters that they develop[17].Soft control mechanism was introduced in [1].This preserves the Quality of Service of legitimate user because instead of rejecting over all requests on detection of DDoS; it filters the aggregated traffic. It takes the choice of rejection or filtration according to result of analysis. But for executing the process of attack detection takes time. To reduce this processing time new scheme is proposed in this paper. Proposed scheme filters the requests at entry point. So it is very effective in terms of execution time.

3. Proposed SYSTEM

A] System Overview

Fig 1. Show the system architecture. How aggregated traffic of normal and abnormal request is handled by the system? How it response to user? is given below.

a) Filtering:
On detection of abnormal behavior it reshapes the incoming request i.e. according to result it takes the decision of how many requests are to be rejected? And how many are to be survived? Among all request. It maintains the list of discarded request which will help in future to block the new request at entry time (in case if match found).Entry point blocking will reduce the time to process the large number of request.

b) Blocking / Pre-filtering:
This process is done if system contains the list of blocked request. The block list is maintained at the time of reshaping. When next time requests come at server first it checks whether newly arrived requests contain the blocked requests? If yes then it removes matching request from newly arrived request.

c) Temporal locality:
Temporal locality implies that recently accessed documents are more likely to be reference in the near future.

d) Spatial Locality:
References to the property that objects neighboring an object frequently accessed in the past are likely to be accessed in the future.

e) Long-term and Short-term detection strategies:
In this work diagnosis of attack can be done on the basis of two types of detection methods namely short term detection and long term detection. Short term detection is used for small period of time and long term is used for large period of

time. In long term detection decision is made by Chi-Square test method. And in short term detection is decision of normality and abnormality is made on the basis of threshold value.

f) Soft control mechanism

On detection of the abnormality in to aggregated traffic system performs the process of reshaping the suspicious requests. For that it uses reshaping algorithm. After performing the filtering it maintains the list of blocked URLs with IP So that maintained record can be used in future at the time of pre-filtering at entry point when new requests are arrives at server.

Proposed system performs filtering as well as blocking of requests. Temporal locality (TL) and spatial locality (SL) of newly arrived request are calculated first. Determined TL and SL are used to train the data and also to detect the abnormal behavior of requests. Further this data is forwarded to GGHsMM. On the basis of comparison between normal and recognized behavior of sequences it takes the decision of either reshaping (in case of suspiciousness) or surviving the all requests as it is (in case of normality). According to result of analysis reply is send back to client through intermediary.

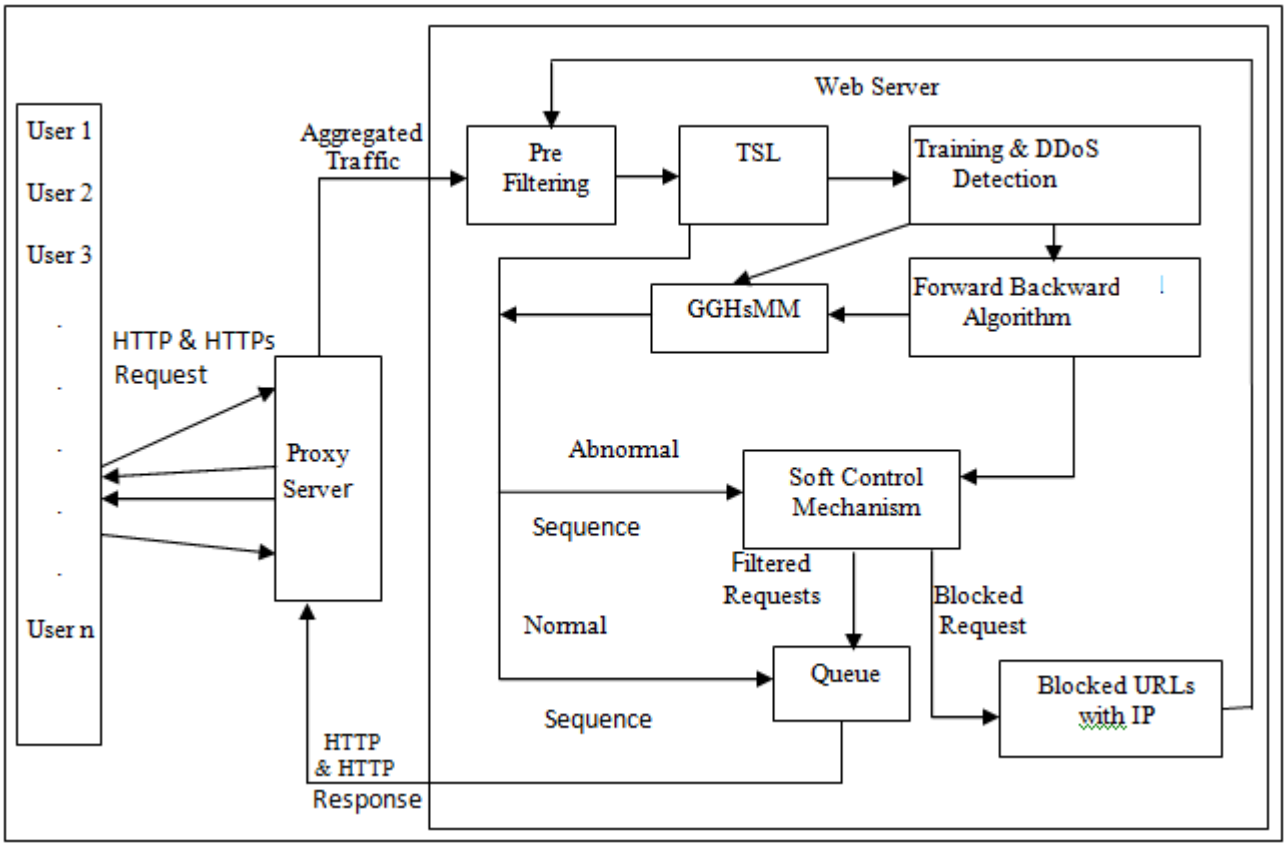

Figure 1 Demonstrations of the proposed system architecture.

B] Reshaping Algorithm

SF- Structure Factor
SR_E- Global Survival Rate
DN -Discard Several
BI- Performance Index
Tw- Total no. of requests
μ- Mean
σ^2 - Variance
σ - Standard deviation

Reshape suspicious reference strings require:
The abnormal reference string: f^w;
The hidden state process of f^w : x^w;
The $BI_w$of f^w and its PDF $\varphi(x, \mu BI, \sigma BI)$;
The Structure Factor: $\overline{SF^i}$; $i \in IM$;
Ensure: Reshape f^w and output:
$\hat{f}$w = {f^w_1,……, $\hat{f}^w_{\hat{T}_w}$}
1. Calculate global survival rate SR_E;
2. Calculate final length of $\hat{f}^w$by $\hat{T}_w = [\hat{T}_w.SR_E]$;
3. for i = 1 to M do
4. DN_i=0;
5. If [$\hat{T}_w.\overline{SF^i}] \leq Num(i, w)$ then Then
6. DN_i= *Num* (i, w)- [$\hat{T}_w.\overline{SF^i}$];
7. End if
8. Randomly mark D$N_i$requests of state i of f^w;
9. End for
10. Discard all marked requests of f^w;
11. Let $\hat{f}^w = f^w$ and output$\hat{f}^w$;

Reshaping algorithm used by the system in case of when system detects the abnormality in to traffic. Normality and abnormality is decided on the basis of threshold value which is calculated by formula

Threshold = μ-2σ

$$\mu = \frac{\sum x}{n}$$

Where,
μ=Mean
x=samples
n=Average of samples

$$\sigma^2 = \frac{\sum (x-\mu)^2}{n}$$

σ i.e. Standard deviation is determined by square root of variance.

Global survival rate is determined by

$$SR_E = \frac{\varphi(BI_w, \mu BI, \sigma BI)}{\varphi(\mu BI + 2\sigma BI, \mu BI, \sigma BI)}$$

$\varphi(x, \mu BI, \sigma BI)$ is nothing but of Probability density function of N (μBI, σ^2BI).

Reshaping algorithm finds the discard number and randomly marks the requests to reject them. It discards the marked requests and survives the remaining requests to user.

C] Mathematical Model

1] Behavior index

$$BI_w \underline{def} P[\hat{X}^w_{1|T_w}, \vec{O}^w_{1|T_w} \mid \lambda]$$

Where,

T_w = Total number of requests

$\vec{O}^w_{1|T_w}$=Observation sequence of the w_{th} time window.

$\hat{X}^w_{1|T_w}$= the optimal hidden state sequence corresponding to $\vec{O}^w_{1|T_w}$

$$L_w = P[\vec{O}^w_{1|T_w} \mid \lambda]$$

denote the likelihood of $\vec{O}^w_{1|T_w}$ Fitting to the parameter set λ.

2] Structure Factor

$$SF^i_w = \frac{Num(i;w)}{\sum_{i=1}^{M} Num(i;w)}$$

Where,

Num (i, w) denotes the number of appeals generated by state i of the w^{th} time window $\sum_{i=1}^{M} Num(i;w) = T_w$ and i € IM.

$$\overline{SF^i} = \frac{1}{W}\sum_{w=1}^{w} SF^i_w, i \in IM$$

D] Expected Experimental Setup

The system is built using Java framework (version JDK 8) on Windows platform. The Netbeans (version 8.0.2) is used as a progress tool. The system doesn't require any specific hardware to run; any standard machine with operating system windows XP or above is capable of running the application.

4. Result and Discussion

Result Analysis

The proposed scheme is implemented on single machine and experiments are performed on dataset of Oracle log. Dataset of Oracle log is taken to generate the number of requests. The experiment includes following aspects:

1) The performance comparison between existing method and previous method.
2) The performance comparison between short term detection with Pre-filtering and short term detection without Pre-filtering.
3) The performance comparison between Long term detection with Pre-filtering and long term detection without Pre-filtering.

Analysis

Comparative analysis between existing system and previous system:-

As we discussed earlier; on detection of DDoS attack; existing explorations rejects the entire incoming requests (even if incoming traffic is an aggregation of normal as well as abnormal requests).Here it affects on the Quality of Service of legal user. So to solve this issue soft control mechanism was given in [1]. Proposed work implemented this soft control scheme in order to preserve the Quality of Service of legal user.

In this work there is need to be provide number of HTTP and HTTPs request as an input data to the system. Table 1 shows the record of denied requests by hard control system and by soft control system

Table 1: Number rejected request by Hard control scheme and Soft Control Scheme

Sr. No.	*Number of Requests*	*Number of Rejected Requests*	
		Existing System (Hard Control)	*Previous system and Proposed system (Soft Control)*
1	37	37	15
2	36	36	21
3	37	37	19
4	37	37	17

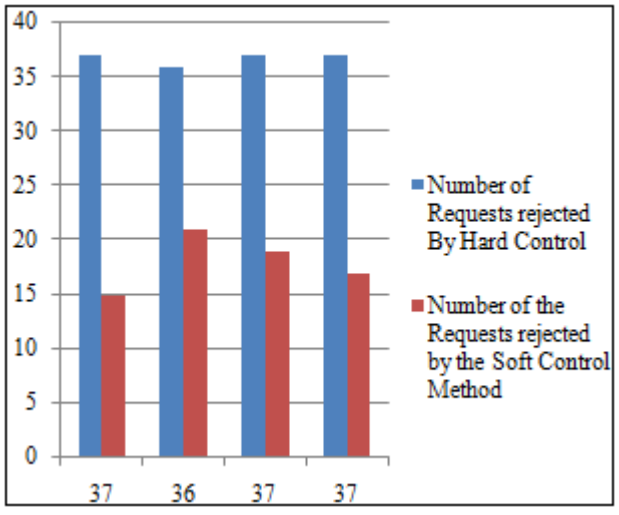

Figure 2: Comparison between Hard Control Mechanism and Soft control mechanism

Y-axis: Number of the request rejected by System
X-axis: Number of the requests

Figure2 depicts the comparison between Hard Control Mechanism and Soft Control mechanism. Blue bar shows number of requests rejected by hard control method and red bar shows number of requests rejected by soft control method. By graph we analyzed that soft control method has less rejected request than the hard control method which helps to improve the Quality of Service.

Evolution

We analyzed our proposed system performance for following cases

1) Processing time of short-term detection without pre-filtering
2) Processing time of short-term detection with pre-filtering.
3) Processing time of Long-term detection without pre-filtering
4) Processing time of Long-term with pre-filtering.

We have performed testing experiment for short term detection on 36 requests of 1sec firstly. At the time of initial execution there was no any previous record of blocking request. So for all new requests; system will do its task of finding normality or abnormality. During this diagnosis; system found abnormal requests on the basis of threshold value. After that it filters the requests by means of reshaping algorithm. System keeps the record of rejected requests and puts the number of blocked request into BlockIp list for future use.

Next time when requests arrive at system it makes the use of previous record of blocked list to check that whether newly arrived requests contains block list entries. In this checking system founds that among 36new requests 8 are matched with Blocked URL with IP. System has deleted these matched requests from new requests. So, now 28 requests are remaining into system for DDoS Detection. Obviously time require to process 37 requests will more than the time require to process 28 requests. From fig.3 we concluded that processing time of 36 is more than 28.

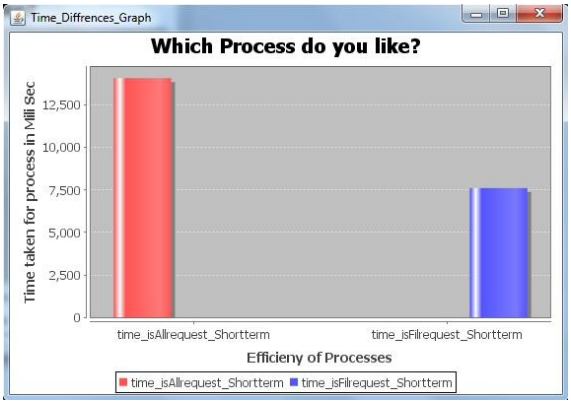

Figure 3: Efficiency of Process in Short term

Same experiment is performed for Long-term diagnosis "without pre-filtering" And "with pre-filtering" As it filters the request initially (if matches found) the time require to process whole incoming request is more than pre-filtered request. These results are depicted in fig 4.

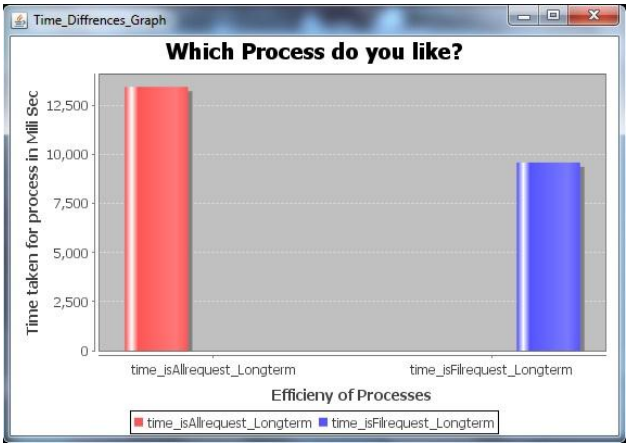

Figure 4: Efficiency of Process in Long-term

Considering the issue of Quality of Service, while implementing this system in real time there is need to delete the record of Block list after Sometime to achieve the basic requirement i.e. quality of service of legal user. Otherwise it may block the legal user requests for longer while. In our work we have assumed count when entries of block list exceeds than that count all records are deleted automatically. So system will take decision on the basis of new records.

Effectiveness of System

Fig.5 depicts the count of allowed requests and blocked request. X-axis denotes the total number of requests (Allowed requests as well as blocked requests) and Y-axis indicates the count of number of requests. From that we can concluded that proposed system is more effective than the system which was explored in [1].

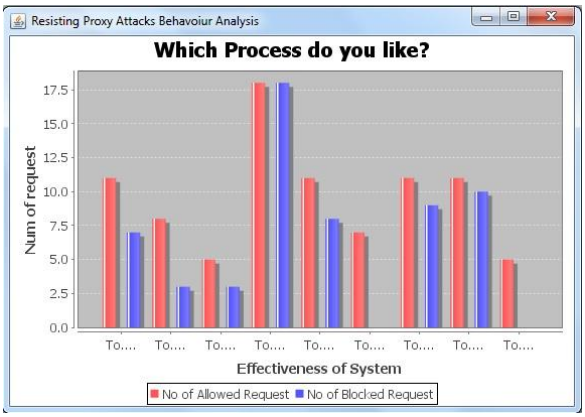

Figure 5: Effectiveness of System

5. Conclusion and Future Scope

Presented work is focused around recognition of DDoS attack and it ensures the origin server. Spatial Locality used to mine the access behavior of users which results into dynamic evolution of proxy-to-server traffic rather than static statistic. Temporal locality is unrelated to the frequency changing of documents. So, system is independent on the frequently varying web contents. Proposed system solves the aggregated traffic (i/e. mixed traffic of attack requests and normal requests) problem of DDoS attack by reshaping the

attack requests from the entire incoming traffic from proxy-to-server. To provide the Quality of Service to the typical client it utilizes soft control mechanism to handle the user demand. On detection of abnormal behavior; it reshapes the requests. On the basis of earlier record if discarded requests are found in newly arrived request at server then those requests are blocked by system at entry point. As it blocks the abnormal request initially; it requires less time to check normality or abnormality of remaining requests.

Finally we have concluded that scheme achieves its goals of providing the quality of Service to normal user and minimizing the time to process the requests (In case of early blocking). Proposed work is well suited for protecting the web server from DDoS attacks.

Proposed work focuses on attack detection, filtration and blocking (Pre-filtering). The reshaping algorithm which is used in our work for filtering discards the requests randomly. In worst case it may reject the requests of real user. Consequently this random rejection of request must be done on the basis of proper criteria. So it motivates us to do such exploration that will reject the request on the basis of particular analysis.

For to achieve better performance even more than proposed system, we have to prevent the happening of such attack instead of providing protection to the system after its happening. In future there is need to develop such mechanism that will prevent the happening of such attack.

References

[1] Yi Xie, Resisting Web Proxy-Based HTTP Attacks by Temporal and Spatial Locality Performance, IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS, VOL. 24, NO. 7, JULY 2013.
[2] X. Jin et al., ZSBT: A novel algorithm for tracing DoS attackers in MANETs, EURASIP J. Wireless Commun. Netw., vol. 2006, no. 2, pp. 19, 2006.
[3] A. Shevtekar, K. Anantharam, and N. Ansari, Low rate TCP Denial-of- Amenity attack detection at edge routers, IEEE Commun. Lett., vol. 9, no. 4, pp. 363365, Apr. 2005.
[4] G. Carl et al., Denial-of-amenity attack-detection techniques, IEEE Internet Comput., vol. 10, no. 1, pp. 8289, Jan./Feb. 2006.
[5] P. Du and S. Abe, IP packet size entropy-based scheme for detection of DoS/DDoS attacks, IEICE Trans. Inf. Syst., vol. E91-D, no. 5, pp. 12741281, 2008.
[6] X Chen, J Heidemann, Flash crowd mitigation via adaptive admission rheostat based on application level observations. ACM Trans Internet Technol. 5(3), 532569 (2005). doi:10.1145/1084772.1084776.
[7] S Ranjan, R Swaminathan, M Uysal, A Nucci, E Knightly, DDoSshield: DDoSresilient scheduling to counter application layer attacks. IEEE/ACM Trans Netw. 17(1), 2639 (2009).
[8] J. Yuan and K. Mills, Monitoring the macroscopic effect of DDoS flooding attacks, IEEE Trans. Dependable and Secure Computing, vol. 2, no. 4, pp. 324335, Oct.-Dec. 2005.
[9] T. Peng and K. R. M. C. Leckie, Protection from distributed denial of amenity attacks using history based IP filtering, in Proc. IEEE Int. Conf. Commun., May 2003, vol. 1, pp. 482486.
[10] B. Xiao, W. Chen, Y. He, and E. H.-M. Sha, An active detecting method against SYN flooding attack, in Proc. 11th Int. Conf. Parallel Distrib. Syst., Jul. 2022, 2005, vol. 1, pp. 709715.
[11] J. Yu, C. Fang, L. Lu, and Z. Li, Mitigating Application Layer Distributed Denial of Amenity Attacks via Effective Trust Management, IET Comm., 2010.
[12] S. Lee, G. Kim, and S. Kim, Taxonomy-Order-Independent Net-effort Profiling for Detecting Application Layer DDoS Attacks, EURASIP J. Wireless Comm. and Networking, 2011.
[13] S. Yu, W. Zhou, W. Jia, S. Guo, Y. Xiang, and F. Tang, Discriminating DDoS Attacks from Flash Crowds Using Flow Correlation Coefficient, IEEE Trans. Parallel and Distributed Systems, 2012.
[14] S. Yu, W. Zhou, R. Doss, and W. Jia, Traceback of DDoS Attacks Using Entropy Variations, IEEE Trans. Parallel and Distributed Systems, 2011.
[15] Y. Xie and S. Yu, A Large-Scale Hidden Semi-Markov Model for Anomaly Detection on User Browsing Performances, IEEE/ACM Trans. Networking, 2009.
[16] Y. Xie and S. Yu, Monitoring the Application-Layer DDoS Attacks for Popular Websites, IEEE/ACM Trans. Networking, 2009.
[17] A. Chonka et al., Cloud security defence to protect cloud computing against HTTP-DoS and XML-DoS attacks, J. Netw. Comput. Applicat.Jun. 23, 2010 [Online]. Available: http://dx.doi.org/10.1016/j.jnca.2010.06.004