

## Finding and improvement of user Based HTTP Attacks on Web Proxy by Using SSL performance

**Dr.k.kiran Reddy**

Associate Professor, dept.of .CSE  
MLRIT, dundigal, Hyderabad

**Dr.P.Bhaskara Reddy**

Director  
MLRIT, dundigal, Hyderabad

### ABSTRACT

Distributed Denial-of-Service (DDoS) attacks continue to be a key threat to Internet applications. In such attacks, a huge amount of traffic generated by a set of attackers, which causes significant damage to the victim's network. In the network communication scenario, request from client seeking resources from other servers passed through intermediate servers called proxy servers. This Overlay proxy network being used to protect applications against such DDoS attacks. Client can access web server through different web proxies which facilitates access to content on the World Wide Web. Web servers have no technique for identifying malicious client. This paper proposed to resist the web-proxy based DDoS attacks. Here Hidden Semi-Markov Model (HsMM) with Gaussian Gamma parameters used to configure the web access behavior sequence and find which proxy cause attack using temporal and spatial access behavior. On existing scenario detection of attacks is based only on the proxy server behavior rather than the actual client. In such cases, innocent web proxies may be blocked. To avoid this problem, a client based approach is employed for detecting spatial and temporal attacks. This paper proposed a threshold based algorithm called Threshold based attack detection (TBAD) for detecting actual attacking client rather than an innocent proxy by adding custom headers in HTTP protocol. Thus by detecting the short term and long term spatial and temporal attack behaviors and reshapes suspicious request to normal one. Thus by effectively using this client based detection approach the QoS of legitimate users can be protected.

**Key words:** DDoS Attacks, HTTP Attacks, Temporal & Spatial Locality, Web Proxy, GGHsMM, TBAD.

### I. INTRODUCTION

Distributed Denial of Service (DDoS) attacks, now are the recent critical threats on the internet communication. DDoS attack relay on interrupting or suspending the services of a client connected to the communication network. By exploiting the HTTP request the attacker find out a legitimate path and thus by easily bypass firewalls and other security measures. In computer networks, a proxy server act as an intermediate server for bypassing the HTTP requests from

clients to web servers. Proxy servers serve as an overlay networks to protect applications against such DDoS attacks. In the current communication era most of the proxies relay upon the web based contents and are called web proxies. Client can access web server through different proxies. Web servers have no technique for identifying malicious client.

This paper proposed to resist the DDoS attacks on web proxy using a client based method. Here Hidden Semi-Markov Model (HsMM) is too used to configure the web access behavior sequence and find which proxy cause attack using temporal and spatial access behavior. To reduce the parametric complexity in the HsMM method Gaussian Gamma parameters is added to it. In existing system a proxy based approach is used to analyze the temporal and spatial behavior. It includes detection of long term and short term access behavior using the Markov model. In this approach only the proxy behavior is analyzed and based on attack detection, the proxy server is blocked. As a result entire clients connected through that proxy is being dropped from the communication. The innocent clients are blocked from getting the services.

## II. DDoS- HTTP ATTACKS

In the current scenario website and web applications are rapidly growing, as internet is an integral part of the modern society. According to the view of security experts all risks related to confidentiality, availability, and integrity because of web based attacks. The purpose of a web based attack is significantly different than other attacks; in most traditional penetration testing exercises a network or host is the target of attack. Web based attacks focus on an application itself and functions on Application Layer of the OSI. Most of all attacks occur at the application layer. One approach for dealing with HTTP-based attacks is to identify malicious code in incoming HTTP requests and eliminates bad requests before they are processed. Different types of DDoS attacks:

**Application-level floods:** Application-level floods now an important DoS attack. Such DoS cause buffer overflow, consume all memory or CPU time. Various other kinds of DoS mostly relay on flooding packets, bandwidth oversaturation, brute force or system resource depletion. Bandwidth oversaturation floods relay on attacker with higher bandwidth than the victim.

Another type of DoS attack is “banana attack” which redirects outgoing messages from the client back onto the client, preventing outside access. A kind of application-level Denial of Service attack is XML DoS (XDoS ) controlled by Web Application Firewalls (WAFs).

**Smurf Attack:** A smurf attack is another type of DoS attack on the public Internet. It relay on network devices that sent packets to host computers via the broadcast address of that network. Then the network acts as a smurf amplifier. Attackers send huge numbers of faked IP packets to the victim. Thus legitimate packets are blocked.

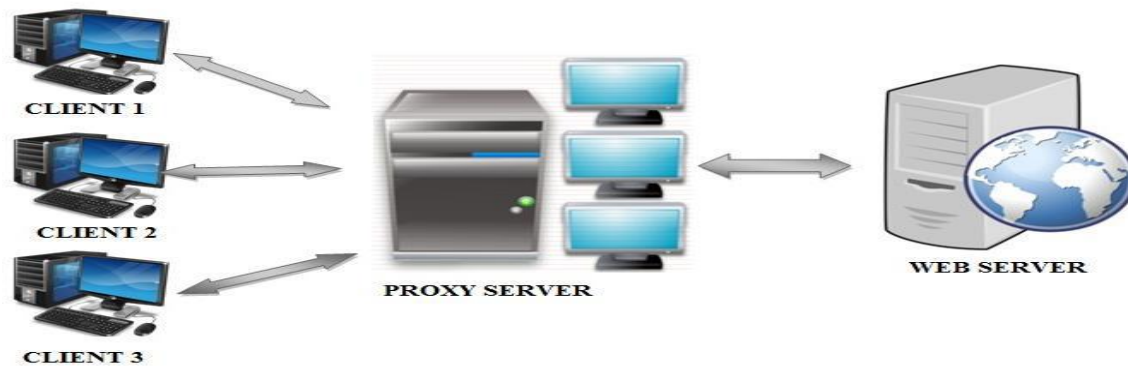


Fig 1.1 Proxy –Web Server Communication

**Teardrop attacks:** Teardrop attacks engaged with sending mangled IP packets to target machine. It attacks on operating systems and cause serious crashes because of error in their TCP/IP packet header.

**Nuke Attack:** Another type of DoS attack is Nuke which targeted on internet .It consists of invalid ICMP packets or fragmented packets which sent to the targeted system. This is achieved by sending corrupted data by using a modified ping utility. As a result targeted system slows down its performance. This attack also delays the legitimate user request and affects the network communication.

### III. WEB PROXY

In computer networks, a proxy server is also a server that acts as an intermediary for requests from clients looking for resources from other servers. A client connects to the proxy server, requesting services, such as a file, connection, web page, or other resources available from a different server. Then the proxy server evaluates the request as a way to simplify and control its complexity. Today's, most proxies available are web proxies, facilitating access to content on the internet.

Attack on a web proxy occurs in different ways: one way is the attacker sends requests to a Web proxy and it makes the web server to forward the attack requests to the origin server. Another way of attacking web proxy is that the attacker disconnects connections between itself and the proxy. The attackers chose this type of attack because of three aspects: It enables the attacker to break through the client-side restrictions by connecting different Web proxies via HTTP protocols; Resisting this type of attack in the middle of Web proxies is not a practical approach, due to lack of cooperation mechanisms between server and proxies, especially uncontrollable private proxies. This attack may confuse most of the existing detection systems designed for the traditional DDoS attacks due to two reasons: first is that the origin server cannot directly observe and diagnose the terminal hosts shielded by the hierarchical proxy system .second reason is the attack traffic is mixed with the regularclient-to-proxy traffic by each proxy that forwards the traffic.

#### **IV. TEMPORAL AND SPATIAL LOCALITY**

Temporal and spatial locality analysis used to extract the proxy to server behavior. Temporal locality of reference has been widely applied in many fields, for example, program behavior reference pattern of Web access and Web proxy cache replacement strategy. Temporal locality refers to the property that a referencing behavior or pattern of web request in the recent past will be referenced again in the near future. The resource request popularity metric represents the frequency of the requests without indicating the correlation between document reference and the time since it was last accessed. Here, we resort to the concept of structure. Here the files are to be placed on a stack such that, whenever a file is requested, it is pulled from its current position and placed on the top of the stack. If the file is not present in the stack then it added to the stack. If the file is found on the stack, then stack distance for the request is calculated by taking the distance of file from the stack top otherwise it said to be undefined.

Spatial locality refers to the behavioral property that frequently accessed objects and its neighboring objects in the past are likely to be accessed in the future. Spatial locality indicates relation among a group of HTTP request patterns. The access behavior of web proxies can be mined using the property of spatial locality. While analyzing the traffic frequency from proxy to server, there is no visible difference between the normal and the attack traffic except their internal purposes. So it is difficult to the victim server to identify the malicious attack.

Compared to existing Distributed DoS attacks the proxy-based HTTP attack are more pliable and surreptitious. The problem of detecting these attacks relay on different aspects: one is that, the attacking hosts are no visible to the origin server because they are shielded or covered by intermediate web proxies. Thus these intermediate web proxies may submissively involved in the attack and act as an attacker. Another aspect is that both the valid and invalid request comes from the same sources. Thus by this hierarchical shielding of web proxies, the detection of actual attacking client is difficult.

From these issues a novel scheme of resistance is proposed based on a client based detection to protect the servers from HTTP attacks .It maps the both the client and proxy behaviors to a hiddensemi-Markov model (HsMM). HsMM is a double stochastic processes model which outlines the observable request process of client- proxy-server traffic. HsMM chain of Markov Model delineate the internal state behaviors of Client and Proxy which can be considered as congenital driving mechanism for client server communication through proxies. By analyzing this HsMM model abnormal behavior of the request from proxies can be measured by taking difference between the proxy's historical behavior and observed behavior.

#### **V. RELATED WORK**

Temporal locality and its impact on web proxy cache [1] performance deals with temporal locality characteristics present in document referencing behavior at web proxies and the impact of this temporal locality on caching document. First, drift metrics are developed to

characterize the popularity profile of documents. Second, a measurement of short-term temporal locality is developed that characterizes the relationship between recent-past and near-future document references. This paper relays on the impact of temporal locality of request pattern in cache performance. But proposed paper deals with both temporal and locality. Here behavior analysis used for improving cache performance, and used at proxy side. But in proposed paper it is used for anomaly detection and used at server side. In [2] A Lightweight Mechanism to Mitigate Application Layer DDoS Attacks attack refers to the attempt to prevent a server from offering services to its legitimate users, typically by sending requests to exhaust the server's resources such as bandwidth or processing power. This system calculates trust behavior of users instead of request. Bases on these trust value, server manage next request from same user .Same detection will be on the proposed system but it based on request not the user. In [3] Anomaly- based network intrusion detection: Techniques, Systems and Challenges deals with the intrusion detection behavior. In this scenario, anomaly-based network intrusion detection techniques are a valuable technology to protect target systems and networks against malicious activities. This deals with well-known anomaly-based intrusion detection techniques. Intrusion Detection Systems (IDS) are security tools like firewalls, antivirus software, and access control schemes, which are intended to strengthen the security and integrity of information and communication systems. Our proposed system also deals with detection of anomalies by using Hidden Markov Model. This system considers only proxy behavior. But our proposed system considers the web browsing behavior of client also. Sequence-order-independent network profiling for detecting application layer DDoS attacks [4] deals on Distributed denial of service (DDoS) attacks, which are a critical threat on the Internet which exploits application-layer. Comparing with proposed system this system detects DDOS attack. With the profiling of web browsing behavior, the sequence order of requests from web page can be used for detecting the application-layer DDoS(App-DDoS) attacks. This system considers web browser behavior but the proposed system deals with web proxy. This system consider sequential behavior, proposed paper consider both temporal and spatial behavior.

In [5] Traceback of DDoS attacks using entropy variations deals with the source of Distributed Denial-of-Service (DDoS) attacks in the Internet are a big challenge. Comparing with proposed system this system present a detection mechanism for DDOS attacks using trace back methods. This system deals with a generalized detection not consider special behavior of proxy. But the proposed system considers the web proxy behavior. The use of entropy based detection and chain model present in this paper can be used for Hidden Markov Model in my proposed work. Discriminating DDoS attacks from flash crowds using flow correlation coefficient [6] presents a flow similarity based approach is taken to discriminate DDoS attacks from flash crowds. The difference from proposed method is that it use flow from proxy. This system uses the HsMM technique. In the proposed system same HsMM technique with Gaussian Gamma parameters is employed. In [7] Low-rate DDoS attacks detection and traceback by using new information metrics deals with DDoS attack in the distributed, cooperative environment. The anomaly-based detection metric typically models the normal network traffic behavior and deploys it to compare differences with incoming network behavior. But proposed system considers the web proxy behavior. The use of entropy based detection and chain model can be used for Hidden Markov Model in the proposed system. In A Large-Scale Hidden Semi-

Markov Model for Anomaly Detection on User Browsing Behaviors [8] deals with Distributed denial of service. Comparing with proposed system this system provides HiddenSemi-Markov Model for Anomaly Detection on User Browsing Behaviors. The proposed method implements same mechanism, but anomalies in Web proxy not in individual browser.

Measuring the Normality of Web Proxies Behavior Based on Locality Principles deals on web proxy locality behavior [9], Web Proxy and cache play important roles in the modern Internet. Comparing with proposed system this system proposed a method for finding malicious web request by using locality behavior of Proxy deals on locality. But not considering spatial behavior. In Monitoring the Application-Layer DDoS Attacks for Popular Websites [10] deals

with continuous critical threat to the Internet. A new application-layer-based DDoS attacks derived from the low layers, utilizing legitimate HTTP requests to overwhelm victim resources are more undetectable. This system introduces a scheme to capture the spatial-temporal patterns of a normal flash crowd event and to implement the App-DDoS attacks detection. This system also a provide Hidden semi Marko Model for detecting Application Layer DDOS attack at Popular website. Same HsMM model is used in the proposed system. But in proposed system it is used to implement in web Proxy to analyze the temporal and spatial behaviors.

## **VI. THE PROPOSED SCHEME**

### **A. Existing System –Disadvantages**

Temporal and spatial locality analysis used to extract the proxy to server behavior. Temporal locality of reference has been widely applied in many fields including program behavior reference pattern of Web access and Web proxy cache replacement strategy. Temporal locality refers to the property that a referencing behavior or pattern o web request in the recent past will be referenced again in the near future. The resource request popularity metric represents the frequency of the requests without indicating the correlation between document reference and the time since it was last accessed. Here, we resort to the concept of stack structure. Here the files are to be placed on a stack such that, whenever a file is requested, it is pulled from its current position and placed on the top of the stack. If the file is not present in the stack then it added to the stack. If the file is found on the stack, then stack distance for the request is calculated by taking the distance of file from the stack top otherwise it said to be undefined.

Spatial locality refers to the behavioral property that frequently accessed objects and its neighboring objects in the past are likely to be accessed in the future. Spatial locality indicates relation among a group of HTTP request patterns. The access behavior of web proxies can be mined using the property of spatial locality.

Here this system uses hidden semi-Markov model, model without state information. That is requests are grouped based on proxy server ID. And this model has no idea about original request source. Because proxy server hide this information from web server due to the HTTP

protocol limitation. By using this model, we detect both temporal and spatial behavior in this system. And this model also reshape incoming request to remove attacking request instead of blocking proxy server. Requests from attackers and non attackers through proxy server are reached at web server side in a mixed manner. Web server has no facility to split this requests. From this request, web server detects spatial and temporal behavior using this markov model. But this will increase false positive and false negative ratio when number of non attacking client increases.

## B. Proposed system

The architecture includes three phases: Training and Traffic Capturing Phase, Parameter Extraction and Analyzing Phase, and third is the Attack Detection and Mitigation Phase.

Different client's requests their files to web server through proxy. In the training phase all the web request are to be analyzed and trained on its behavior. All the traffics are to be captured and sorted using Traffic capture module.

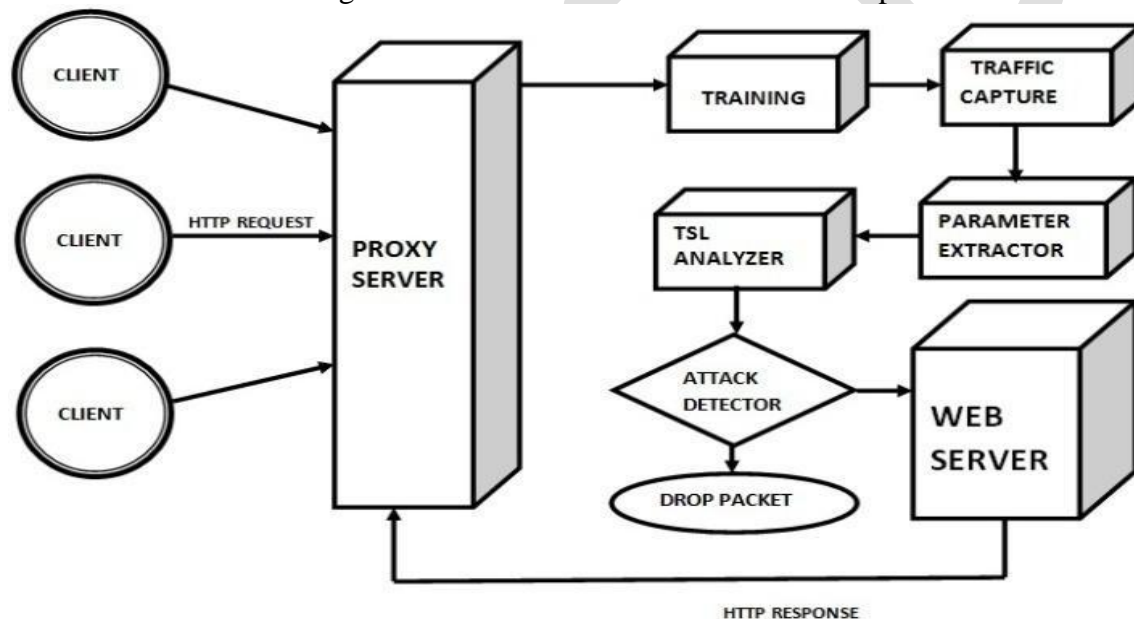


Fig 1.2 System Architecture

Parameter Extractor extracts the features such as remote IP address and the arrival time of packets. Then the packets are subjected to the Analyzer which decides whether to drop or allow the packets into the network based on the threshold set. The analyzer module contains an IP frequency list to store the number of occurrences of individual IP address over a period of time. It checks the frequency of each IP address in IP frequency list and further decides whether to block or allow the packets. Threshold value is set to restrict the number of HTTP requests to a particular IP address for a given period of time. The attack detection and control done using the proposed Threshold Based Attack Detection (TBAD) Algorithm.

Most of the attacks related to DDoS are flood attacks. The main reasons of flood attacks are due to the vulnerability in the protocol. By using the nature of protocols design a UDP Flood or SYN Flood attack saturates the network traffic. In a SYN Flood attack, attacker employs the TCP three way handshake's protocol mechanism .It first spoof the IP addresses in initiation process and then to drain server side system resources. In a UDP Flood attack, attacker uses the stateless UDP protocol to spoof IP addresses and to drain server side network resources. To accomplish an effective security mechanism, every mitigation method for the flood attacks must be implemented in a perspective of system/protocol design.

HTTP protocol serves at the application layer of the OSI Model. So it is easy to analyze and detect packet payloads by application layer security devices like IPS or WAF (Web Application Firewall). There are no inspection and analyzing on the HTTP flood attacks for other devices which do not serve at the application layer. The only detection way for these devices is TCP connection counts made for the HTTP responses. As a result of detection, HTTP Flood attack attempts can be prevented by and blocked on different layers of OSI model other than OSI application layer. There are many situations in the real time scenarios that the HTTP flood attacks are not detected and mitigated properly. Some of them might be related with system security configuration and others might be depending on an absence of a security mechanism. These situations are handled with the other security enhancements at the different level of the communication technology architecture. Here web application level security comes in. Unlike from a network-layer protection, an application-layer solution needed for the application that it is protecting. Web application is a group of technologies which serves for the web services. It is important to justify whether the attack is a HTTP flood attack or not.

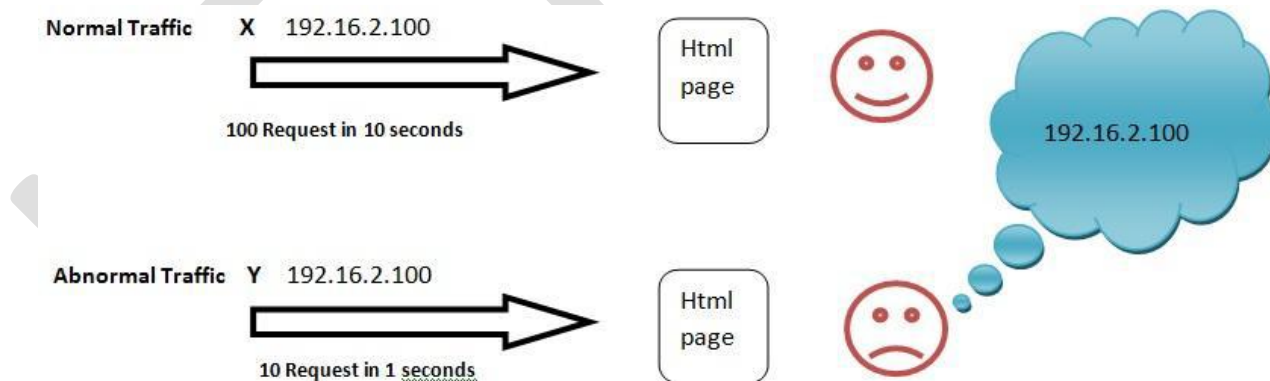


Fig.1.3- HTTP Flood Attack Awareness for the Web Application

In the case of HTTP flood attack, a HTTP request payload carried by a TCP packet should be interpreted by the web application. Attack space for HTTP flood attacks always begins with the web service and its backend infrastructure. To create a resistance at the web application level against the HTTP flood attack formalized into three phases: First, detect IP addresses of the abnormal requests according to a previously defined rule. Second, To reduce attack surface, requests are returned with a low resource used response like a blank page or else. Third, block the detected IP addresses by using some other components at the other mitigation levels such as

WAF, web server etc. Saving the resources for the backend infrastructure will prominently reduce the amplitude of the HTTP flood attacks.

### C. The Rule Creation Concept

The critical point for the web application level HTTP flood attack mitigation is the false positives. The detection rules must be clear and be tested with the real world traffic usage scenarios to mitigate false positives. Also a good understanding for the rule creation concept is highly suggested. But here we implement an enhanced HTTP protocol in this proxy server. So proxy server doesn't hide application id from web server. So web server got client identity of each request. So client can group requests based on this application ID.

### VII. IMPLEMENTATION A. Algorithm

Here implement a new algorithm (Threshold based attack detection-TBAD) for detecting attacks modules of the proposed TBAD, viz., Traffic capture, Parameter extractor, and the Analyzer module. First the packets are captured at kernel level by the traffic capture module. The output of the traffic capture module, the outbound TCP packets alone, are filtered and sent to the parameter extractor module which extracts the features such as remote IP address and the arrival time of packets. Then the packets are subjected to the Analyzer which decides whether to drop or allow the packets into the network based on the threshold set. The analyzer module contains an IP frequency list to store the number of occurrences of individual IP address over a period of time. It checks the frequency of each IP address in IP frequency list and further decides whether to block or allow the packets. The flow chart for the basic functioning of TBAD is given in Fig. 1.4. Threshold value is set to restrict the number of HTTP requests to a particular IP address for a given period of time. Based on the parameters extracted from the packets  $\Delta T$  values are calculated, which gives the time interval between current and previous instance of a packet for a particular IP address. Using HTTP GET different IP addresses are extracted from the filtered packets.

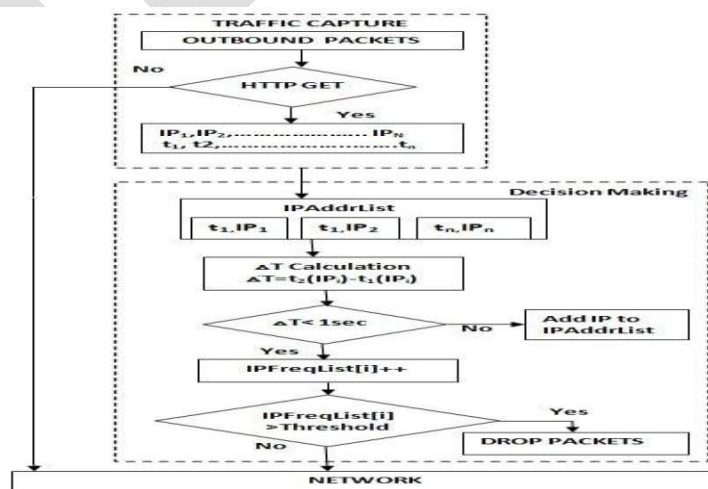


Fig1.4 TBAD Algorithm Flow Chart

If the time interval is less than 1 second, then the IP frequency list value for the corresponding IP address is incremented by 1. Otherwise the IP address is added to the IP frequency list. The time interval to refresh list is set to 1 second and also the IP frequency list values are reset to null after the elapse of every second. The threshold value (N) is set to 20 based on the experiments conducted in our lab, viz., only 20 HTTP GET requests are allowed to a particular IP in one second. The HTTP GET packets are allowed to an IP address till the corresponding IP frequency list value reaches N. If the IP frequency list value for an IP address exceeds N, then the packets are dropped.

## VIII. EXPERIMENT AND ANALYSIS

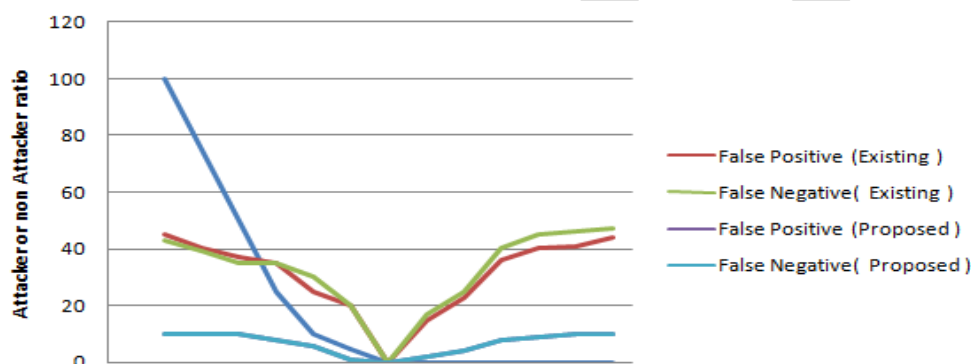


fig 1.5 Performance graph In the existing system, both false positive and false negative ratio increases when difference between number of attacking and non attacking packets increases. But from figure 1.5, it is clear that it have no effect on proposed system.

## IX .CONCLUSION

Application layer attack has become a major threat to the internet in today's world. The focus of this project was to come out with an effective solution for the detection and prevention of clients from inadvertently taking part in such attacks. Accordingly, a Threshold Based Attack Detection (TBAD) was proposed and implemented in Windows OS using J2SDK. Experiments were conducted by generating HTTP GET attacks and using TBAD for its mitigation. It was evident that the TBAD suppressed the flooding packets and thus prevented the client system from taking part in such an attack. The ongoing work is to implement TBAD in order to find different type of DDoS attacks.

## REFERENCES

- [1] A. Mahanti, D. Eager, and C. Williamson, "Temporal Locality and Its Impact on Web Proxy Cache Performance," (2000) Performance Evaluation, vol. 42, nos. 2/3, pp. 187-203, 2000. Computing, vol. 5245, pp. 61-73, 2008.
- [2] J. Yu, C. Fang, L. Lu, et al., "A lightweight mechanism to mitigate application layer DDoS attacks," (2009) Scalable Information Systems, vol. 18, pp. 175-191, 2009.

- [3]P. Garcia-Teodoro, J. Diaz-Verdejo, G. Macia-Fernandez, and E. Vazquez, "Anomaly-Based Network Intrusion Detection: Techniques, Systems and Challenges," (2009) Computers and Security, vol. 28, nos. 1/2, pp. 18-28, 2009.
- [4]S. Lee, G. Kim, and S. Kim, "Sequence-Order-Independent Network Profiling for Detecting Application Layer DDoS Attacks," (2011) EURASIP J. Wireless Comm. and Networking, vol. 2011, no. 1, p. 50 2011.
- [5]S. Yu, W. Zhou, R. Doss, and W. Jia, "Traceback of DDoS Attacks Using Entropy Variations," (2011) IEEE Trans. Parallel and Distributed Systems, vol. 22, no. 3, pp. 412-425, Mar. 2011.
- [6] S. Yu, W. Zhou, W. Jia, S. Guo, Y. Xiang, and F. Tang, "Discriminating DDoS Attacks from Flash Crowds Using Flow Correlation Coefficient," (2012) IEEE Trans. Parallel and Distributed Systems, vol. 23, no. 6, pp. 1073-1080 July 2012
- [7]Y. Xiang, K. Li, and W. Zhou, "Low-Rate DDoS Attacks Detection and Traceback by Using New Information Metrics," (2011) IEEE Trans. Information Forensics and Security, vol. 6, no. 2, pp. 426-437, June 2011.
- [8]Y. Xie and S. Yu, "A Large-Scale Hidden Semi-Markov Model for Anomaly Detection on User Browsing Behaviors," (2009) IEEE/ACM Trans. Networking, vol. 17, no. 1, pp. 54-65, Feb. 2009.
- [9]Y. Xie and S. Yu, "Measuring the Normality of Web Proxies Behavior Based on Locality Principles," (2008) Network and Parallel
- [10]Y. Xie and S.-Z. Yu, "Monitoring the Application-Layer DDoS Attacks for Popular Websites," (2009) IEEE/ACM Trans. Networking, vol. 17, no. 1, pp. 15-25, Feb. 2009.



KIRANREDDY, B.E, M.E.(CAD), MBA, M.Tech.(CSE). Department of CSE, MLR Institute of Technology, Dundigal, Hyderabad. He is doing Ph. D in Computer Science & Engineering, JNTUH. His research interests include Information Security, Mobile computing, Distributed Systems, Cognitive Science and MANETS. published 12 research papers in various international journals.



Dr. P. Bhaskara Reddy, the Director MLRIT is a young and dynamic Professor of ECE, has 26 years of Industry, Teaching, Research and Administrative experience in Reputed Engineering Colleges & Industry. In 24 years of experience served various positions from Asst. Professor to Principal

Research & Guidance: Published 1 Book (International Edition) "Information Technology in Technical Education – Economic Development by "LAMBERT Academic Publishing" Published 9 Laboratory Manuals, 63 Research papers at National and International Level on Education, Electronics Communication, I.T, Computer Networks, E-Commerce etc. Guided 5 Research Scholars for their Doctorates, about 50 M.Tech., M.C.A. and B.Tech projects .

Symposiums Conducted: 6 National Level Technical Symposiums on various topics in Electronics & Communications, Computers etc.