ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 3, Issue 2, March 2014

Securing Web Proxy Based Network from Http Attacks with Provision for Detecting Attacker Nodes

Dr. P. Balakumar, Yedu Krishnan.R

Abstract—* A perfect novel server-side defense scheme is proposed to resist the Web proxy-based DDoS attack. The approach utilizes the temporal and spatial locality behavior of the requesting nodes to identify attacks and IP of the node as a unique identity to identify the particular attacker node, which makes the scheme more accurate than existing schemes. A new Forward-backward TSL based HsMM algorithm is proposed to describe the time-varying traffic behavior of Web proxies. Soft control is a novel attack response method proposed in this work. It performs behavior reshaping that tries to converts a suspicious traffic into a relatively normal one before rudely discarding them. If behavior reshaping does not succeed then the particular request is permanently identified as an attack and based on the proposed methodology to find the particular attacker node, the system identifies the particular attacker client and can notify the proxy about the attacker. A Session hijacking handler technique is effectively used for identifying session hijack attacks. The Proposed variation in HTTP protocol supports for identifying which client is intruder rather than detecting the innocent web proxy.

***Index Terms*— DDoS, Hidden semi Markov Model, Soft control, Temporal and Spatial Locality.**

I. INTRODUCTION

In computer networks, a proxy server[2] act as 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. A client usually connects to the proxy server, requesting some service, such as a file, connection, web page, or other resource available from a different server and the proxy server evaluates the request as a way to simplify and control its complexity and effects across the network. A Web proxy may be turned easily into an attacker by two steps: In the first step the attacker sends attack requests to a Web proxy and forces it to forward the attack requests to the origin server. In the second step the attacker disconnects connections between itself and with the proxy.

Mainly two methods can be used to penetrate through the Web proxies: requesting dynamic documents or setting up a Cache-Control : no-cache in the headers of HTTP requests. A single host can simultaneously trigger a lot of Web proxies to attack a Web server without the need of invading them. The attraction of such an attack lies in three aspects : i) It enables the attacking host to break through the client-side restrictions by connecting different Web proxies via HTTP protocols; ii) Resisting such an attack by the mid Web proxies is not a good approach, may be due to lack of cooperation mechanisms between server and proxies. iii) Such an attack may confuse most of the existing detection systems designed for the traditional DDoS attacks due to two reasons: first, the origin server cannot directly observes and diagnose the terminal hosts shielded by the hierarchical proxy system; secondly, the attack traffic is mixed with the regular client-to-proxy traffic by each proxy that forwards the traffic[18].

In the final aggregated proxy-to-server traffic, it is very difficult to identify the difference between the normal traffic and the attack traffic. Thus, the victim server is hard to accurately identify and filter the attack requests. The Web proxy-based HTTP attack is more flexible and covert than most of existing DDoS attacks. The challenges of detection of attacks lies in three aspects which can be identified as follows: i) Real attacking hosts are unobservable to the origin server since they are shielded by the hierarchical Web proxies; ii) A Web proxy may be passively involved in an attack event and may unconsciously act as an attacker [18]; iii) Observed from the victim server, both normal and abnormal traffic comes from the same sources (i.e., Web proxies). Although most of the large-scale official proxies are usually configured to be secure, they cannot avoid being abused for the proxy based attacks. This type of attacks may bring new challenges to existing network security systems [13]. Motivated by these issues, a novel resisting scheme is proposed to protect the origin server from Web proxy-based HTTP attacks in this work. The proposed scheme is based on network behavior analysis. It maps a Web proxy's access behavior to a hidden semi-Markov model which is a typical double stochastic processes model. The output process of an HsMM[14] profiles the observable varying process of a proxy-to-server traffic.

The hidden semi-Markov chain of an HsMM describes the transformation of a proxy's internal behavior states that can be considered as the intrinsic driving mechanism of a proxy to server traffic. In such behavior model, detecting the abnormality of a Web proxy can be achieved by measuring the deviation between an observed behavior and the Web proxy's historical behavior[15] profile. Long-term and short-term behavior assessment methods are proposed. Long-term behavior assessment provides warnings on a large scale whereas the short-term behavior assessment locates abnormal request sequences embedded in the proxy-to-server traffic. Here we propose a TSL based behavior analysis with IP based filtering to accurately detect the particular attacker client. The scheme also proposes a new soft-control for attack response.

The scheme reshapes the suspicious sequences according to certain prefixed criteria's. It converts a suspicious sequence into a relatively normal one by partly discarding its most likely malicious requests instead of denying the entire sequence. Hence a behavior reshaping occurs each time before discarding an entire request sequence. This server side technology can be made implemented in large proxy-based live networks. The proposed system will be able to be used as an efficient defending mechanism in the server –side against http attacks such as application layer DDoS attacks. It is also a solution for Session Hijacking attack.

II. PROPOSED SYSTEM

The main reason of flood attacks is the vulnerability in the protocol. Consider a UDP Flood or SYN Flood attack which uses the nature of protocol's design to saturate the network traffic. In the case of SYN Flood attack the attacker uses the TCP 3 way handshake's first initiation step to spoof IP addresses and to drain server side resources [9]. When the subject comes to the UDP Flood attack, the attacker uses the stateless design of the UDP protocol [9] to spoof IP addresses and to drain server side network resources. Due to this reason, to accomplish an effective security solution, every mitigation method for the flood attacks must be implemented in a consideration and perspective of system/protocol design.

Since HTTP protocol serves at the application layer, it is possible to detect and analyze packet payloads only by application layer security devices like IPS or WAF. For other security devices which do not serve at the application layer, there are no inspection and analyzing chance on the HTTP flood attacks [9]. The only detection way for these devices is TCP connection counts made for the HTTP responses [11]. As a result of detection, HTTP Flood attack attempts can be prevented by and blocked on different layers of OSI model other than application layer. There are many situations in the real world scenarios that the HTTP flood attacks are not mitigated properly. Some of them might be related with security configuration weakness of the security device and others might be depending on an absence of a security device. These situations might be handled with the other security enhancements at the different level of the information technology architecture scenarios. This is how the web application level comes in and can be identified for studying.

To create a resistance at the web application level against the HTTP flood attacks, the basic idea might be summarized into 3 steps: i) Detect IP addresses of the abnormally excessive requests according to a some defined rule; ii) To reduce the attack surface, just return these requests with a low resource used response (like a blank page or else); iii) Block detected IP addresses by using other components at the other mitigation levels (WAF,web server/service, etc.). While reducing attack surface by sending low resource used responses as described above, this implementation will also save resources of the backend infrastructures like the SQL Servers, or many other infrastructures like the distributed servers or the e-mail/ media/application servers, etc. This is extremely important and critical for the network architectures which share the backend infrastructure members with other infrastructures like intranet or distributed web application servers. Saving the resources for the backend infrastructure will prominently reduce the amplitude of the HTTP flood attacks. It is a good security practice to bring the HTTP flood attack awareness for the web application and implement additional precautions to every mitigation level including the web application level.

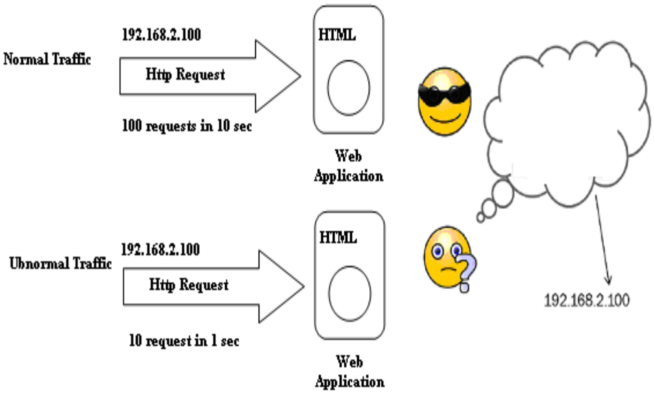

Fig. 1. Http flood attack awareness for web application

A. The Rule Creation Concept

The critical point for the web application level HTTP flood attack mitigation is the false positives. In order to avoid false positives, all the detection rules must be well defined and be tested with the real world traffic usage scenarios. 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.

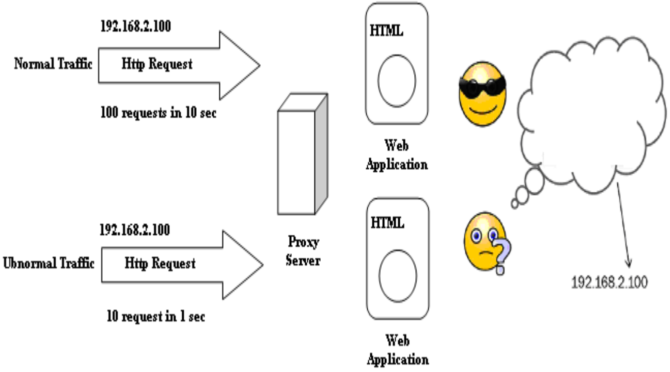

Fig.2. Proxy Server included network

In order to accomplish a healthy rule base against HTTP flood attacks, the initial step should be defining the normal traffic. A basic abnormal traffic rule based on these baseline values could be sampled as 10 requests in 0.1 seconds. According to the normal traffic baseline values, it states that 1 request in 40.000 microseconds from a single IP address cannot be considered as a HTTP flood attack. The abnormal traffic rule above allows 1 request in 10.000 microseconds at 10 times from a single IP address. Based on the rule creation concept, the rule also has a tolerance factor pointed out by 10 times description. Thus the tolerance factor (the request count) can give an opportunity to mitigate false positives.

Besides HTTP flood attacks, this web application level implementation can provide an opportunity to slow down the Brute Force Attacks and Web Vulnerability Scanners. The detected IP addresses shared with other security components would also provide an opportunity to block attacker's access to the web application. In addition to this traffic rule, we also implement different mechanism for detecting attacks such as unsupported HTTP method, oversized Header and Body data size, Large or small time out interval, minimum incoming data and SQL Injection.

III. SYSTEM MODULES

A. Attack Class Creation

Initial work deals with identifying and creating the attack classes. Based on the attack histories and proxy-origin server abnormalities certain attack classes can be predicted. The different attack classes are for the following:

- Unsupported HTTP method
- Oversized Header and Body data size

- Large or small time out interval
- Minimum incoming data
- SQL Injection
- Command Execution
- Path Traversal

The above shown are some unavoidable attack classes. Unsupported http method sometimes leads to http flood attacks. Oversized header and body may lead to request abnormality. SQL injection is another kind of attack which is a code injection technique. In such kind of attack the attacker inputs malicious SQL statements into the application, which breaks the security and the attacker thereby steals the information or gets unauthorized access. Thus all these kind of attacks must be given high consideration. As a result here we give top priority for creating above given attack classes

B. Training Phase

The main step involved in the training phase is the attack learning. The attack learning is a complex process in which the request sequence characteristics are studied. The main characteristics here we take into account is the TSL or the Temporal and Spatial behaviors. The incoming requests are treated as queries which are then compared against each attack classes. The request is then assigned with the attack type that provides the best match using probability distribution. Then identify the temporal and spatial sequence and their distance.

C. Detection Phase

During the comparison high priority is given to TSL behavior rather than normal behaviors. If any sequence found within the distance, that pattern is identified as an attack. If attack is found in the incoming request, then perform Temporal & Spatial Behavior Pattern Identifier analysis and organize incoming request as valid and invalid sequences.

D. Behavior Reshaping

This includes a soft control scheme which reshapes suspicious request sequence according to normal behavior. This process is done by partly discarding most likely malicious request instead of denying entire request sequence. This partly discarding is done based on a threshold value which is referred throughout. The request is cut into parts based on this threshold value in order to attain reshaping.

E. Http Protocol Extension

Design and implement a new HTTP protocol for detecting client based attack instead of Proxy based. Modify existing HTTP protocol by adding custom headers in HTTP protocol. These custom headers contain the IP of the client when forwarding from proxy server to server. So web server can group request from each client separately and easily detect attack based on client IP which is mandatory along with the request sequence.

F. Session Hijack Handler

IP based identification of the client attacker is the key feature of this proposed scheme. Since IP is used as the identity factor, it has less relay on session while dealing with application. Hence the proposed scheme is a solution for session hijacking also. The session hijack handler here uses a long random number or string as the session key. This reduces the risk that an attacker could simply guess a valid session key through trial and error or attacks. Regenerating the session id after a successful login is made mandatory. This prevents session fixation because the attacker does not know the session id of the user after login.

To overcome the shortcomings of the existing algorithms, a new Forward-Backward TSL based HsMM is proposed which utilizes the TSL behavior and detects attack by utilizing the features of a Forward-Backward algorithm and HsMM algorithm, which is an extension of the existing HsMM algorithm.

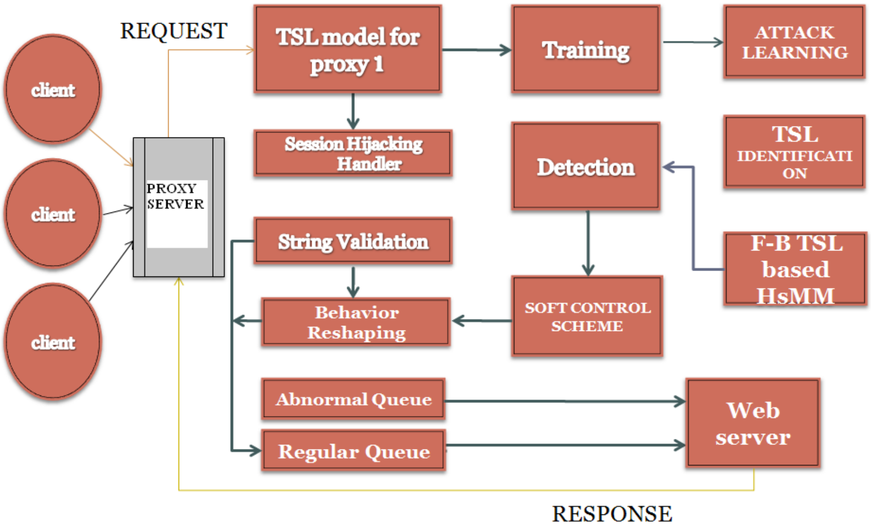

Fig.3. System Architecture

IV. COMPARATIVE STUDY

In existing systems, there is no provision for identifying which client is intruder. It only detect the innocent web proxy [18]. Here we propose a new HTTP protocol (modified) for identifying the particular attacker node. It identifies the particular client attacker and has provision to notify the proxy that a particular client is an attacker. If the proxy is not blocking the client there is provision for blocking the entire proxy. The proposed system also has a new technique for identifying session hijacking attacks. Then a new algorithm Forward-Backward TSL based HsMM, which is highly accurate is utilized. The new method reduces the number of parameters to be estimated, able to characterize the dynamic evolution of the proxy-to-server traffic rather than the static statistics.

The http protocol extension allows the particular attacker detection. The approach utilizes the temporal and spatial locality to extract the behavior features of the proxy-to-server traffic. Soft control is a novel attack response method proposed in this work that converts a suspicious traffic into a relatively normal one by behavior reshaping rather than rudely discarding. Thus our proposed scheme provides a mechanism which is capable of efficiently detecting DDoS attack along with identifying the particular attack creating nodes. Thus the case of punishing or blocking innocent proxies in other existing systems is avoided in the proposed scheme enabling efficient attack management system.

V.DISCUSSION

Several other algorithms and methods are used for detecting and preventing DDOS attacks. Former methods include some router based methods such as D-WORD[17]. These router based algorithms [7] had to be implemented in routers and are not capable to detect application layer DDoS. The proposed method in the project can be implemented in direct servers and is capable for handling most of application layer DDoS.

Other existing algorithms for application layer DDoS handling include the M-algorithm [17], which uses the browsing behavior to detect DDoS. But these are not successful in case of proxy server networks. The proposed method in the project is very much suite for proxy environment. Some other existing methods are using access matrix and entropy variations for anomality calculations. Examples are Shannon entropy and Kullback-Leiber distance matrix [16]. These are capable for detecting low rate DDoS but fails in huge network environment. But the proposed system very much detects low rate DDoS attacks and also succeeds in huge network environment when compared to the performance of the existing schemes available. Many other methods uses Hidden Semi Markov model which is a good algorithm for anomality detection. The proposed scheme uses a new forward backward hidden semi-Markov model parameterized by Gaussian-mixture and Gamma distributions.

The new method reduces the number of parameters to be estimated, and can characterize the dynamic evolution of the proxy-to-server traffic rather than the static statistics. It also ensures a soft control mechanism with behavior reshaping. Most of the efficient techniques utilize the Temporal and Spatial behavior. The proposed scheme also includes Temporal and Spatial behavior analysis. But the methodology here provides a more accurate analysis so that even low rate DDoS could be detected soon. Some methods need to use separate servers called CAT-servers [19] for implementing the DDoS detection technique. But the proposed scheme could be easily implemented in the origin server itself.

Here a novel server-side defense scheme is proposed to resist the Web proxy-based distributed denial of service attack. Proposed approach utilizes the temporal and spatial locality to extract the behavior features of the proxy-to-server traffic and makes the scheme independent of the traffic intensity and frequently varying Web contents. Then a new Forward-Backward TSL based HsMM algorithm is proposed to describe the time-varying traffic behavior of Web proxies.

Attacker/Non Attacker Request Ratio	FalsePositive (Existing)	FalseNegative (Existing)	FalsePositive (Proposed)	FalseNegative (Proposed)
100	45	43	10	10
75	40	39	10	10
50	37	35	10	10
25	35	35	8	8
10	25	30	6	6
5	20	20	1	1
0	0	0	0	0
0.2	15	17	2	2
0.1	23	25	4	4
0.04	36	40	8	8
0.02	40	45	9	9
0.0133	41	46	10	10
0.01	44	47	10	10

Table.1.Expected Performance Table

Two diagnosis approaches at different scales are introduced to meet the requirement of both fine-grained and coarse-grained detection scenarios. The Soft control is a novel attack response method proposed in this work. The method converts a suspicious traffic into a relatively normal one by using the behavior reshaping mechanism rather than rudely discarding. Thus a perfect system with attack detection along with capabilities such as soft-control mechanism can be obtained. We can expect efficient and more accurate result from the proposed scheme. Based on the capabilities of the system which is explained above, an expected performance graph of the system can be obtained as follows:

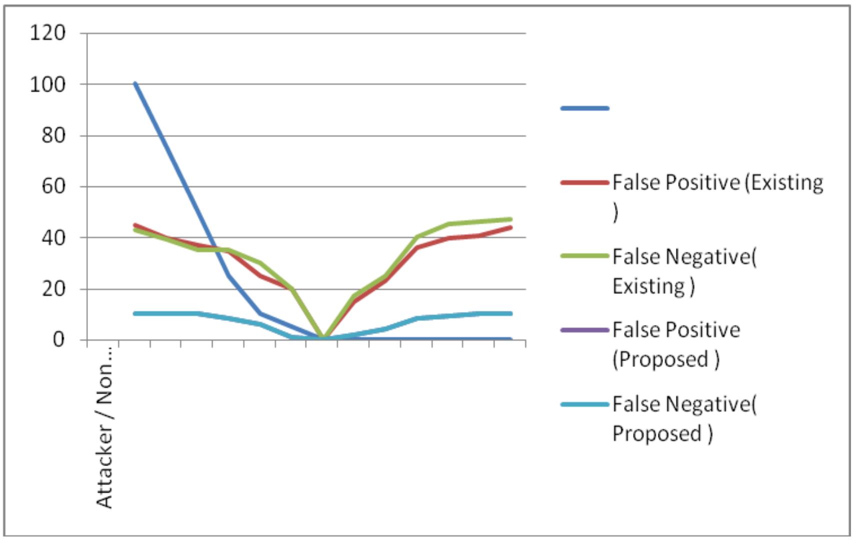

Fig. 4. Expected Performance Graph

VI. CONCLUSION

In this paper, we tried to filter the attack traffic from the aggregated proxy-to-server traffic, which is considered as a new problem for the DDoS detection. An accurate resisting scheme was proposed based on TSL. Forward-Backward TSL based HsMM and soft-control were proposed to improve the detection performance. The main advantages of our approach shown in the experiments include: 1) its detection performance is better than the pure statistical methods; 2) it is independent of the traffic intensity and the frequently varying Web contents; 3) it can realize the early detection.; 4) not only finding the attacker proxy, the scheme can detect the particular attacker client; 5)soft control mechanism utilizes behavior reshaping. Hence the proposed scheme can block the particular attacker node or client and not the innocent proxy. The experiments can be further extended and in future the technology can be made improved so that it can be capable of handling other serious http attacks such as IP spoofing.

ACKNOWLEDGMENT

The authors would like to sincerely thank the anonymous reviewers for their insightful comments and constructive suggestions, which significantly helped the authors to improve the quality of the paper.

REFERENCES

[1] Amey Shevtekar, Karunakar Anantharam, and Nirwan Ansari (2005)," Low Rate TCP Denial-of-Service Attack Detection at Edge Routers", IEEE Communication Letters.

[2] Anirban Mahanti, Carry Williamson (2000) "Temporal Locality and its Impact on Web Proxy Cache Performance".

[3] Chia Yuan Cho, Juan Caballero, Vern Paxson (2004), "Insights from the Inside: A View of Botnet Management from Infiltration", Carnegie Mellon University ICSI.

[4] Jaeyeon Jung, Balachander Krishnamurthy (2002), "Flash Crowds and Denial of Service Attacks: Characterization and Implications for CDNs and Web Sites" www 2002.

[5] Jian Pei,Jiawei Han,Behzard Mortazavi"Mining Access Patterns Efficiently From Web Logs", Simon Frazer University, Canada.

[6] Jie Yu , Chengfang Fangy, Liming Luy, ZhoujunA "Lightweight Mechanism to Mitigate Application Layer DDoS Attacks."

[7] John Ioannidis, Steven M. Bellovin," Implementing Pushback: Router-Based Defense against DDoS Attacks".

[8] Kejie Lu, Dapeng Wu, Sinisa Todorovic (2007), "Robust and efficient detection of DDoS attacks for large-scale internet," Computer Networks.

[9] P Lersak Limwiwatkul' and Arnon Rungsawangr,(2004), "Distributed Denial of Service Detection using TCP/IP Header and Traffic Measurement Analysis" ,international Syinposium on Communications.

[10] Shui Yu, Member, IEEE, Wanlei Zhou, Senior Member, IEEE, Robin Doss Member, IEEE,and Weijia Jia, Senior Member, IEEE (2010), "Traceback of DDoS Attacks using Entropy Variations" ,IEEE Transactions On Parallel And Distributed Systems.

[11] Shui Yu, Member, IEEE, Wanlei Zhou, Senior Member, IEEE,Weijia Jia, Senior Member, IEEE, Song Guo, Senior Member, IEEE,Yong Xiang, and Feilong Tang,(2012), "Discriminating DDoS Attacks from Flash Crowds Using Flow Correlation Coefficient", IEEE Transactions On Parallel And Distributed Systems.

[12] Sujatha Sivabalan, Dr P J Radcliffe (2013). "A Novel Framework to detect and block DDoS attack at the Application layer"IEEE 2013-Tencon."

[13] Tao Peng and Christopher Leckie and Kotagiri Ramamohanarao (2006), "Survey of Network-based Defense Mechanisms Countering the DoS and DDoS Problems," ACM Transactions on Computational Logic.

[14] Xiaobin Tan, Hongsheng Xi (2011) "Hidden semi-Markov model for anomaly detection", Applied Mathematics and Computation.

[15] XIE Yi and YU Shunzheng, "A Detection Approach of User Behaviors Based on HsMM", ITC19/ Performance Challenges for Efficient Next Generation Networks.

[16] Yang Xiang, Member, IEEE, Ke Li, and Wanlei Zhou (2011), "Low-Rate DDoS Attacks Detection and Trace back by Using New Information Metrics", IEEE Transactions on Information Technology.

[17] Yi Xie and Shun-Zheng Yu, (2009), "A Large-Scale Hidden Semi-Markov Model for Anomaly Detection on User Browsing Behaviors", IEEE/ACM Transactions on Networking.

[18] Yi Xie ,S.Tang and J.Hu,(2013), "Resisting web proxy-based Http attacks by temporal and spatial locality behavior," IEEE Transactions on Parallel and Distributed System.

[19] Yu Chen, Member IEEE, Kai Hwang, Fellow IEEE, and Wei-Shinn Ku, Member, IEEE "Collaborative Detection of DDoS Attacks over Multiple Network Domains" IEEE Transactions on Parallel And Distributed Systems.

[20] "FortiWeb Protection against DoS/DDoS Attacks" A Fortinet White Paper (2013).

AUTHOR BIOGRAPHY

Dr.P. Balakumar received Phd degree in October 2011. He is currently working as an Associate Professor and Head of the Department of Computer Science in Mahendra Institute of Technology, Mallasmudram, Namakkal district, Tamil Nadu, India. His research interests include various fields of computer science.

Yedu Krishnan.R received the BTech in Information Technology degree from Anna University of Technology, Tamil Nadu, India in 2011 and currently doing ME in Computer Science and Engineering in Mahendra Institute of Technology under Anna University, Chennai. Research interests include network security, web services and database technology.