AENSI Journals

Australian Journal of Basic and Applied Sciences

ISSN:1991-8178

Journal home page: www.ajbasweb.com

Sustaining Web Services By Differentiating Ddos Attacks From Flash Crowds

[1]Prof.P. Krishna kumar, [2]Prof.Dr.K. Vijayalakshmi, [3]Prof.Dr.R. Bharathi

[1]*ASP/CSE ,PET Engineering College,Vallioor, Tamilnadu, India*
[2]*Dr. K.Vijayalakshmi, HOD/CSE, Ramco Institute of technology, Rajapalayam, Tamilnadu, India.*
[3]*AP/ECE,University College of Engineering Nagercoil, India.*

ARTICLE INFO

Article history:
Received 20 August 2014
Received in revised form
19 September 2014
Accepted 23 October 2014
Available online 16 November 2014

Key words:
Distributed Denial of Service (DDoS) proxy server, HTTP request, flash crowd.

ABSTRACT

The recent Denial of service (DoS) and Distributed Denial of Service (DDoS) attacks on popular websites and web servers shows how defenseless the Internet is under such attacks. This paper presents a novel technique to detect the application layer-based DDoS attacks which become more serious during flash crowd event. This approach deals with the inter-arrival time between two successive HTTP requests from a client, popularity of web page and the maximum request that could be made by a human in a particular duration of time for a particular web page to differentiate flash crowds from DDoS attacks. This technique is applied online for an efficient detection of DDoS attacks and it is seen that this technique prevents 99.9% of attacks. The simulation work is carried out with DARPA 1999 and DARPA 1998 data sets.

To Cite This Article: Prof.P. Krishna kumar, Prof.Dr.K. Vijayalakshmi, Prof.Dr.R. Bharathi., Sustaining Web Services By Differentiating Ddos Attacks From Flash Crowds. ***Aust. J. Basic & Appl. Sci.,*** *8(16): 405-412, 2014*

INTRODUCTION

Distributed Denial of Service has caused severe damage to server and to the development of new Internet services. DDoS attacks which were carried out at the network layer are called Network Layer DDoS attacks. Since many techniques have been implemented to detect DDoS attacks, attackers have shifted their offensive strategies to application layer attacks. An application layer DDoS attack is launched by overwhelming the vicim server with numerous HTTP get requests (HTTP flooding) or by issuing numerous queries through the victim's search engine. Flash crowd refers to the situation when a very large number of users simultaneously access a popular web site. For example when the election results of a nation is declared or during the World cup foot ball or cricket match, the number of users simultaneously visiting a particular popular web site to search the results increases tremendously leading to flash crowds.

Some of the major Denial of Service attack incidents which happened in the year 2009 are listed below. During the 2009 Iranian election protests, foreign activists launched DDoS attacks against Iranian government's official website (ahmedinejad.ir) thus making the website inaccessible. On June 25, 2009, the day Michael Jackson died, the spike in searches related to Michael Jackson was so big that Google News initially mistook it for an automated attack. As a result, for about 25 minutes, when some people searched Google News they saw a "We're sorry" page. In June 2009 the famous P2P site known as The Pirate Bay was rendered inaccessible due to a DDoS attack. Multiple waves of July 2009 cyber attacks targeted a number of major websites in South Korea attack. and the United States. The attacker used botnet and any file update through internet is known to assist its spread. As it turns out, a computer trojan was coded to scan for existing MyDoom bots. MyDoom was a worm in 2004, and in July around 20,000-50,000 were present. MyDoom has a backdoor, which the DDoS Bot could exploit. Since then, the DDoS bot removeditself, and completely formatted the hard drives. Most of the bots originated from China, and North Korea. On August 6, 2009 several social networking sites, including Twitter, Facebook, Livejournal, and Google blogging pages were hit by DDoS attacks.Although Google came through with only minor set-backs, these attacks left Twitter crippled for hours and Facebook experienced trouble for several weeks. A survey was conducted by CISCO technology developer partner- Arbor Networks with the help of human users to know about the attack mitigation technique they use. Figure 1 shows the result of the same.

Corresponding Author: P. Krishna kumar, ASP/CSE, PET Engineering College,Vallioor, Tamilnadu, India.
E-mail: {ponkrishkumar@yahoo.com, vijaya@mepcoeng.ac.in}

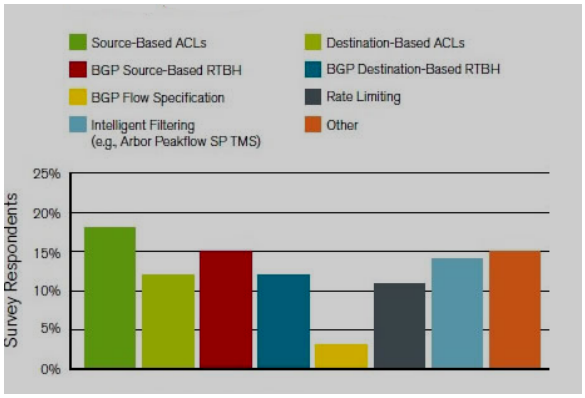

Fig. 1: Primary Attack Mitigating Techniques

According to Arbor survey the three most often referenced obstacles to reduce the attack mitigation time by respondents include (i) Accurately identifying and separating attack flows from legitimate traffic (ii) Communication with upstreams, customers and internal staff (iii) Internal resources and manpower to mitigate attacks.Figure 2 shows the graph of increasing attack size with every year.

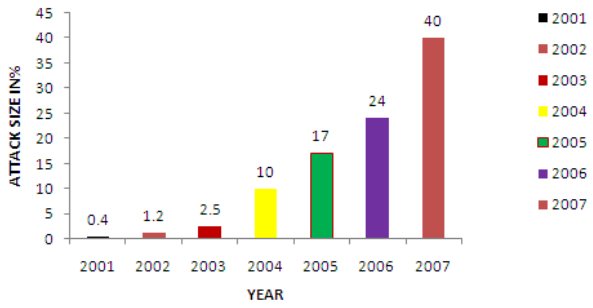

Fig. 2: Year vs attack size

Our contributions in this paper are (i) Proxy server concept employed helps to reduce the load and risk for a Web Server. After verifying the HTTP request based on the set threshold time for the inter-arrival time of HTTP requests, the requests from the client will be accepted or denied.(ii) The threshold values are adaptive to the changes in the Internet traffic i.e the threshold set for the maximum number of Hits is adapted to change with the existing web traffic conditions.(iii) Faster DDoS detection process during flash crowds by maintaining a database with whitelist and blacklist for decision process.

The rest of this paper is organized as follows section II briefly reviews related works. Section III presents the detailed description of the proposed system. In section V the implementation issues and results are discussed. In Section V the merits of the proposed technique is explained. Finally we concluded our paper in section VI and discussed further research needed towards this issue.

II. Related Work:

Previous researches show that the DDoS attacks are launched in Network layer or IP layer. The IP layer DDoS detection schemes attempt to detect attacks by analyzing specific features eg. arrival rate or header information. Recently the attack on popular websites is increasing tremendously. However little work is done related to the detection of application layer DDoS attacks.

Ranjan et al. (2006) used statistical methods to detect characteristics of HTTP sessions. The main principle used is rate limiting and it is used as the primary defense mechanism. Yen et al. (2005) defended the application layer DDos with constraint random requests by the statistical methods. Jung et al. (2002) used two properties to distinguish the DDoS from flash crowds. They are i) flash crowds are due to increase in number of clients distributed throughout the internet sending requests to access the popular website, while DDoS is due to a limited number of clients from a particular group of IP addresses ii) DoS clients originate from new IP addresses or new client clusters compared with flash crowd events which originate from client clusters or IP addresses which have send HTTP requests before the flash crowds.

Yuan et al (2005) used the cross correlation analysis to capture the traffic patterns and to decide where and when a DDos attack possibly arises. Yi Xie et al (2005,2006) used Hidden semi-markov model to detect. Application layer DDos attacks for popular websites Georgios et al proposed a model using three aspects of human behaviour a) request dynamics b) request semantics and c) ability to process visual clues to differentiate DDos bots from human users. Other existing defense methods may be those based on man–machine interaction, e.g. puzzles, passwords, and the CAPTCHAs. However, as Kandula*et al.* in (2005) and Ranjan *et al.* in (2006) have pointed out, those schemes are not effective for the DDoS attack detection because they may annoy users and introduce additional service delays. Our solution is mainly based on inter arrival time of HTTP requests and it is very simple and adaptive to changes in the web documents and web server's popularity.

III. Detailed Description Of The Proposed System:

Wepropose a practical DDoS defense system that aims to sustain web services by differentiating DDoS attacks from flash crowds during a flash crowd event. Our proposed system provides service to legitimate clients with a good quality of service. The following figure 3 shows the presence of Application layer DDoS during flash crowds due to Mydoom worm.

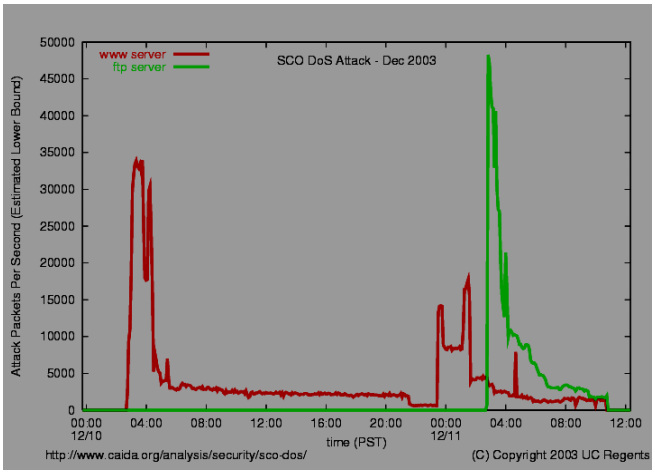

Fig. 3: DDoS Vs flash crowds

The model of the proposed system is shown in the figure 4. It consists of a Web Server, a Proxy Server, Decision Module and Clients on the Internet.

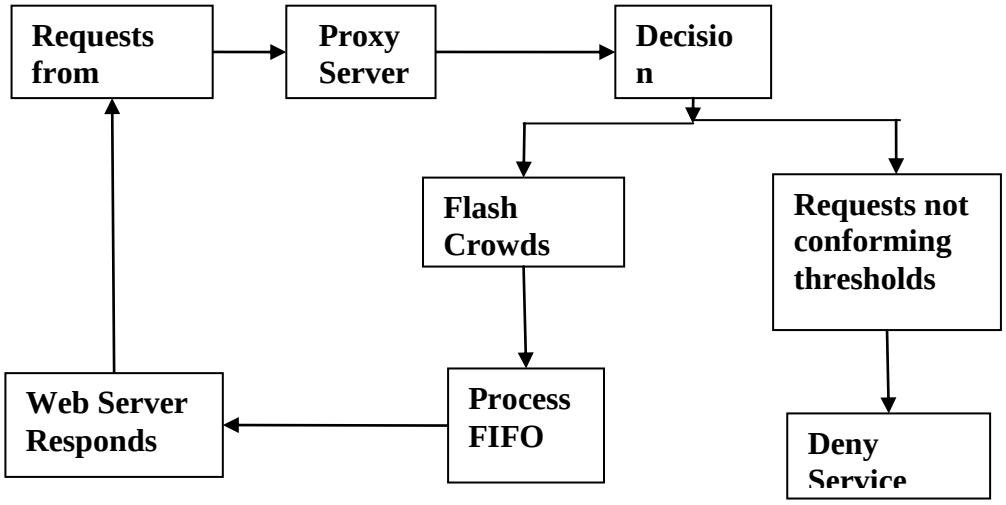

Fig. 4: Model of the Proposed System

Proxy server:

A proxy server is server that acts as an intermediary for requests from clients seeking resources from other servers. A client connects to the proxy server requesting some service such as a file, connection, web page or

other resource availability from a different server. The proxy server evaluates the request according to its filtering rules for example it may filter traffic by IP address or Protocol. If the request is validated by the filter, the proxy provides the resource by connecting to the relevant server and requests the service on behalf of the client. The proxy server provides the following functions

i) Speeds up access to resources for legitimate client by using cached web pages from a web server.
ii) Applies access policy to network services or content for example to block undesired sites
iii) Logs the internet and web page usage

The Proxy server used is also called as 'Web Proxy' as it focuses on World Wide Web traffic. The most common use of a web proxy is to serve as a web cache proxy server provides a means to deny access to URL's specified in black list. A black list is a list or register of persons who are being denied a particular privilege, servi*ce*, access or recognition conversely a white list is a list or compilation identifying persons or organizations that are accepted ,recognized or privileged.

In the proposed model shown in figure 4, client1 sends a HTTP request to access a document to the web server. The HTTP request passes through the router and reaches the proxy server. The proxy server receives the HTTP request and sends it to the detection module. At the detection module the database is searched to check whether the incoming client's IP address is in the black list of database. If the client's IP address is not in black list , the inter arrival time of two or more successive requests from that particular client is measured. If the inter arrival time threshold is satisfied then the decision is taken to confirm the HTTP request. At the same time the HTTP request for the particular web page or website, the address of the client issuing the requests and time of arrival of requests are recorded in a log file. A counting scheme is enabled to count all successive page hits or requests from a client.

A graph of webpage Vs hit rate value is plotted using the counting mechanism. The data obtained in a log file for web pages vs hit rate is used to adjust the hit for a particular web page. If the demand is very high and no DDoS attackers are visible then a high value of request rate is set. If the demand for a popular web site is very high but if there are presence of DDos attackers, then the threshold is set to a low value. Thus the threshold values are adaptable to changes in internet conditions

In the following we state the assumptions used in designing the proposed systems

1) We assume that the proxy server plays a major role in the process of detection of DDos in the presence of flash crowds. The proxy server keeps the web server keeps the web server robust against flash crowds and bandwidth DDoS attacks

2) We assume that there can be large number of attackers. The attacker is identified by the inter arrival time successive HTTP requests.

3) We assume that the threshold values used are adaptable to changes in web traffic

Mathematical Analysis of the Proposed Model:

A request is a human-initiated event such as a click on a web page. Requests for embedded objects such as pictures on a page are considered as part of the original request. The Request Hit Rate is defined as

$$RHR = Pit \Big/ \sum_{i=1}^{N} Pit \quad (1)$$

where $\mathbf{P_{it}}$ is the request number or hit count of the i th web page at the i th time unit. N is the number of web server's documents. An access matrix can be defined as follows

$$\mathbf{A_{NxT} = [a_1\ a_2\ a_3\ \ldots\ldots\ldots a_N]^T} \quad (2)$$

where $\mathbf{a_i = (a_{i1}\ a_{i2}\ \ldots\ldots\ a_{iT})^T}$

The request hit rate is also defined as the ratio between the average request number per user on the ith document at t th time unit and the average request number per user at the t th time unit.

We assume $a_{it} = R_{it}$because it is more suitable to detect the attacks that repeatedly request the same pages such as home page or "hot pages".Log file maintained by proxy server contains the details of IP address of the destination server and time at which HTTP request is made. Page filter implemented in our approach is used to view the access of web pages by clients.

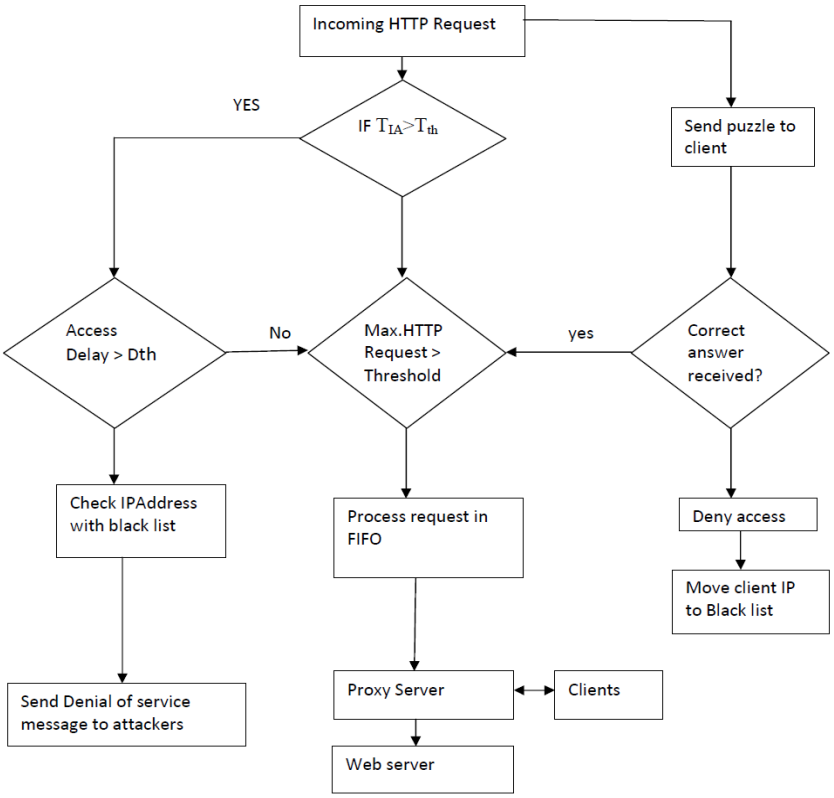

Fig. 5: Flow chart of the Proposed system

IV.Implementation Issues And Results:

We have implemented our approach successfully . The inter arrival time of two successive HTTP requests from an individual node is measured and compared against the threshold. The threshold value for TIA is fixed after studying the behavior of user for a period of time. In our experimental work we have fixed the value for TIA as 500 ms .The filtering of DDoS attackers by using threshold for inter-arrival time of requests helps us to differentiate machine generated request attacks from flash crowds generated due to human users.

Our scheme also fixes a threshold value for the number of hits on a particular web page.This threshold value varies depending on the popularity of web page and web site. If a particular web page or web site is frequently visited by users then the threshold is fixed as a large value In our experimental work we fixed the threshold for critical web pages as a low value and for other web pages as a high value

Our scheme implements page filters for each page and when the requested number of hits from an individual client exceeds the threshold value the IP address of that particular client is moved to blacklist. When the client requests for a web document the size of the requesting document is also accounted to aid in detecting attackers. Figure 6 shows the distribution of clients and attackers on a network

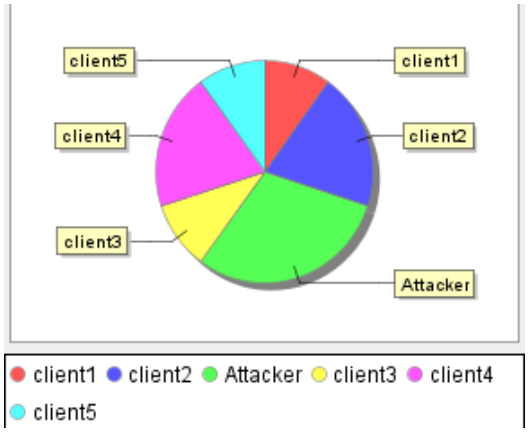

Fig. 6: Pie Chart showing distribution of attackers and clients

Figure 7 shows the plot of webpage versus hitcount. The maximum threshold value for hit count was set as 20. A small value of threshold for hit count may annoy the user if he is an legitimate client.Hence the threshold is to be fixed in a real time Internet after studying the popularity of web page.

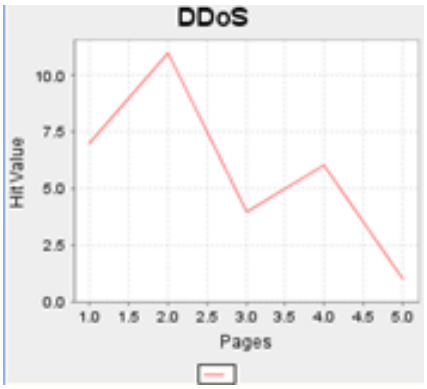

Fig. 7: Web pages Vs Hit count

Figure 8 shows the comparison between web pages in terms of hit count and the total value of hits is shown by the yellow bar and figure 9 shows the hit chart.

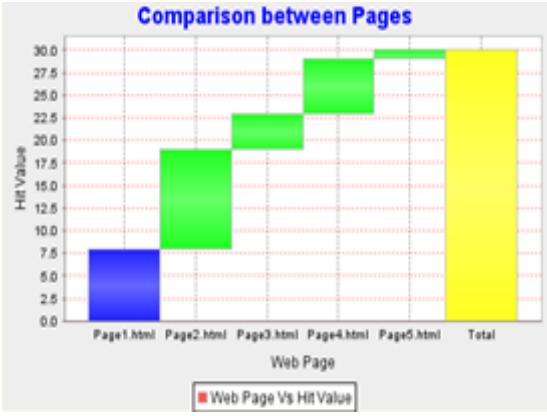

Fig. 8: Comparisonof usage between web pages

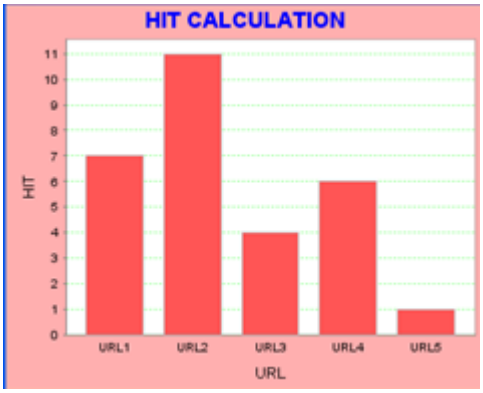

Fig. 9: Hit chart

Pearson's Correlation Coefficient is proposed to detect the repeated features of Packet's Arrival. The Time Stamp and Changing Rate of Port Numbers are considered for the feature extraction from the successive requests from the clients [Legitimate/Illegitimate].The generated pattern is directed to the proxy server. Consequently, the decision is left to the proxy server whether the requested client can be denied or provided the service. This experiment is carried out with several datasets.

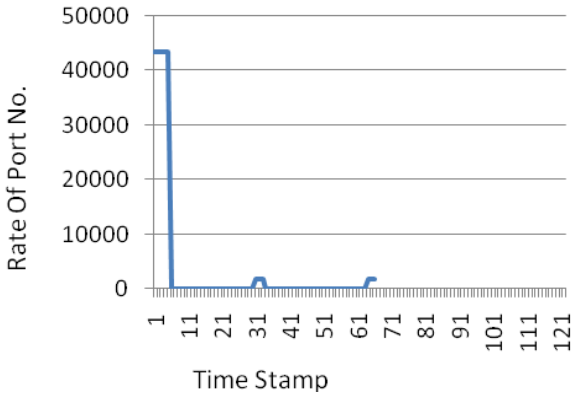

Fig. 10: Experiment on Dataset Darpa 1999-Feature Extraction for DDoS

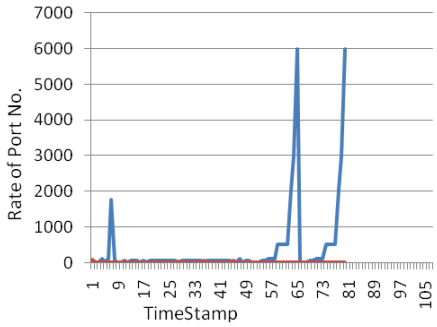

Fig. 10: Experiment on Dataset Darpa 1998-. Feature Extraction for Flash Crowds

Conclusion And Future Work:

Wehave proposed a three level defense against flash-crowd attacks which differentiates DDoS from flash crowds. Our implementation work shows that the proposed defense scheme is successful against DDoS attackers. The results are very much promising.Our future work involves implementing this technique in the current Internet topology to overcome application layer DDoS attacks. Further the load on web servers and proxy servers have to be investigated individually and with the help of human users we can evaluate the likelihood of false positives. Above all the success of our work depends on the ethical behavior of the people inthis Internet era.

REFERENCES

Abraham Yaar, Adrian Perrig, Dawn Song, 2006. "StackPi: New Packet Marking and Filtering Mechanisms for DDoS and IP Spoofing Defense," IEEE J. Selected Areas in Comm., 24: 10.

Chen, Hwang, and ku, 2006. "Collaborative Detection of DDoS attacks over multiple network domains",IEEE Transactions on parallel and distributed systems, tpds-0228-0806 pp: 1-14.

Gavrilis, D., I. Chatzis and E. Dermatas, 2007. "Flash crowd detection using decoy hyperlinks," 2007 IEEE International Conference on Networking,Sensing and Control, pp: 466-470.

http:// www.wikipedia.org

http://www-ece.rice.edu/networks/papers/dos-sched.pdf

JelenaMirkovic, AlefiyaHussain, Sonia Fahmy, Peter Reiher and Roshan K. Thomas, 2009."Accurately Measuring Denial of Service in Simulation and TestbedExperiments"*IEEE Transactions On Dependable And Secure Computing*, 6(2): 81-94.

Jian Yuan and Kevin Mills, 2005. "Monitoring the Macroscopic Effect of DDoS Flooding Attacks" IEEE Transactions On Dependable And Secure Computing, 2(4): 324-335.

Jung, J., B. Krishnamurthy and M. Rabinovich, 2002."Flash crowds and denial of service attacks: Characterization and implications for CDNs and web sites," in *Proc. 11th IEEE Int. World Wide Web Conf.*, pp: 252-262.

Kandula, S., D. Katabi, M. Jacob, and A. Berger, 2005. "Surviving Organized DDos Attacks That Mimic Flash Crowds," in USENIX Symposium on Network Systems Design and Implementation.

Mirkovic, J., G. Prier and P. Reiher, 2002."Attacking DDoS at the source," in *Proc. Int. Conf. Network Protocols*, pp: 312-321.

Park, K., V. Pai, K. Lee and S. Cal, 2006. "Securing Web Service by Automatic Robot Detection," in proceedings of Usenix Annual Technical Conference.

Ranjan, S., R. Swaminathan, M. Uysal and E. Knightly, 2006. "DDoS-resilient scheduling to counter application layer attacks under imperfect detection," in *Proc. IEEE INFOCOM*, Apr. [Online]. Available:

von Ahn, L., M. Blum and J. Langford, 2004. "Telling Humans and Computers Apart Automatically," Communications of the ACM, 47(2): 57-60.

Xie, Y. and S. Yu, 2005. "A detection approach of user behaviors based on HsMM," in Proc. 19th Int. Teletraffic Congress (ITC19), Beijing, China, 2: 451-460.

Xie, Y. and S. Yu, 2006. "A novel model for detecting application layer DDoS attacks," in Proc. 1st IEEE Int. Multi-Symp.Comput.Computat. Sci. (IMSCCS|06), Hangzhou, China, 2: 56-63.

Yen, W. and M.-F. Lee, 2005."Defending application DDoS with constraint random request attacks," in *Proc. Asia-Pacific Conf. Commun.*, Perth, Western Australia, pp: 620-624.

Yi Xie and Shun –Zheng Yu, 2009. "Monitoring the Application –Layer DDoS attacks for Popular Websites" *IEEE/ACM Transactions onNetworking*,17(1)1: 5-25.