DDoS attacks are one of the most deadly threats rising in internet [5].

In DDoS attacks the attacker use various means to exhaust the resources of a desired server/system so that the other requests cannot be processed and hence bring the services down. The amount of DDoS attack has been increasing drastically in recent years [7].

In this section we are presenting the different methods which are previously used for differentiation and also discussing some advantages and limitations of these systems.

- In paper [1] author used information distance technique to distinguish DDoS from flash crowd .Both these attacks are motivated different methods to measure the similarity among flows such as Abstract distance metrics, Jeffrey distance, Sibson distance, Hellinger distance. After comparison among these four metrics, it is found that the Sibson distance is the most suitable method. By applying an algorithm to the real datasets, an accuracy around 65% and it is very efficient to improve an accuracy of the flow based discrimination strategy.
- In [2] DDoS is distinguished from flash crowd by using probability metrics. They proposed main contributions to distinguish DDoS attacks from Flash crowds as hybrid metric and the Bhattacharyya metric. The hybrid metric can reduce the false positive rate greatly. But the limitation of this method is that it is not applied in the real network situation, and so cannot find out more recognizable characteristics of IP packets.
- Paper [3] presented a packet arrival pattern for distinguishing DDoS from flash event. In this paper, two methods are used; first Behavior based detection which can discriminate DDoS attack traffic from traffic generated by real users and second Pearson's correlation coefficient which can extract the repeatable features of the packet arrivals. The major limitation is two methods are not tested with different packet information such as packet delay and changing rate of port number so that it can test with the real scenarios in real time. So there is no confirmation of the performance from the predictability test.
- In [4] discrimination of DDoS from flash crowd is done with the help of flow correlation coefficient, used as a similarity metric among suspicious flows. Limitations of this method are, the detection rate of differencing DDoS from flash crowd is less, tracing of the sources of the DDoS attack is not given and it is very hard to identify DDoS attack flows at sources since the traffic is not so aggregate using world cup dataset.
- In [7] authors proposed a survey of botnet technology and defense system. They described different kinds of networks that have access to different types of visibility and this has a strong impact on the effectiveness of any botnet detection mechanism. They surveyed that botnet behaviour is undiscoverable and these are moving targets.
- In paper [10] characterization and implications of flash crowd and DDoS for content distribution networks CDNs and Web Sites is presented. This method cannot used to obtain larger flash crowd logs from diverse places and experiment against instrumented servers

We surveyed on different techniques to differentiate DDOS attack from flash crowd such as Information distance, Probability Metrics, Packet Arrival Patterns and Flow Correlation Coefficient [1]-[4]. Among these techniques "Flow Correlation Coefficient" shows the better results compared to another three techniques. But after a detail study of this technique, we found some drawbacks i.e. detection rate of differencing DDoS from flash crowd is less, tracing of the sources of the DDoS attack is not given and it is very hard to identify DDoS attack flows at sources since the traffic is not so aggregate [5].

So, the novel proposed system increases the rate of differentiating DDoS attack from Flash Crowd by using flow correlation coefficient, increases accuracy and also traces the sources of the DDoS attack.

3. Proposed System

As there are some drawbacks of existing system of discriminating DDoS attack from Flash Crowd. All these drawbacks are overcome in proposed system. After analyzing the characteristics of DDoS attacks and the existing Algorithms to detect DDoS attacks, this report proposes a novel detecting algorithm for DDoS attacks based on flow correlation coefficient

3.1 Block Diagram of Proposed System

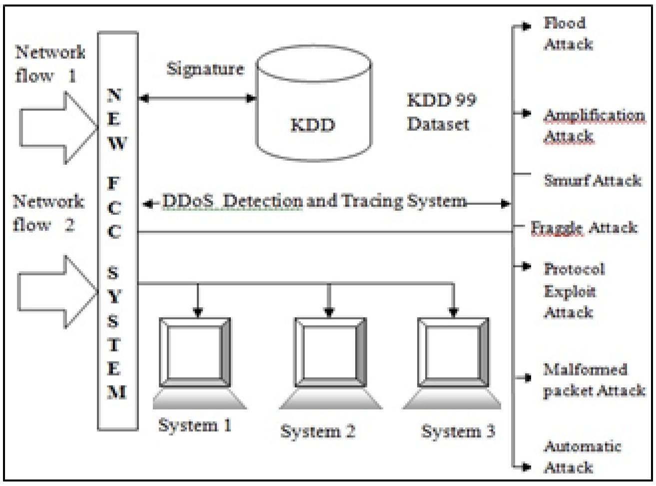

Figure 1: Block diagram of Proposed FCC System and DDOS detection and Tracing System

Two network flows with the same length are given in above Fig.1.Detection algorithm using flow correlation coefficient is used to indicate similarity between two flows. It is sometimes the case that two similar flows may have a phase difference which will decrease the correlation coefficient. So it is easy to deal with because we can shift one flow to match the other and take the maximum value of the correlation coefficients to represent the similarity of two flows. The new FCC system is used to increase the rate of differentiating DDoS from flash crowd.

The proposed FCC system combines parameters from KDD CUP 99 dataset such as time, Duration, Protocol, service flag, src_bytes and dest_bytes of flows at each router to distinguish DDoS from flash crowd. In this way, our novel approach aims to improve the global security level and is the best solution to DDOS attacks in theory.

Also our novel tracing system is used to trace the sources of the DDoS attack .It will detect the subtypes of DDoS attack such as flood attack, amplification attack, smurf ,fraggle attack etc[6].

4. Detection and Tracing Mechanisms

The detection mechanism contains two algorithms to differentiate the DDoS from flash crowd.

1. Network Packet Tracing Algorithm for DDoS using Fuzzy Logic Rules
2. Detection Algorithm using Fuzzy Logic Classifier.

These two phenomenons reduce the workload over the network, detection time, and storage space required for routers and increases performance and scalability.

A sample community network with flows can is given in fig.2 [4]. In the sample community network, R1, R2 and R3 are three routers where R2 and R3 are the edge routers, and we try to protect to the server that is potential victim. Consider, Pi and Pj are two incoming flows observed at R3 and R2, respectively.

The two network flows merge at router R1 and both are directed to the potential victim, and enter the community network through different paths. We collect the number of packets for a given network flow with a specific time interval.

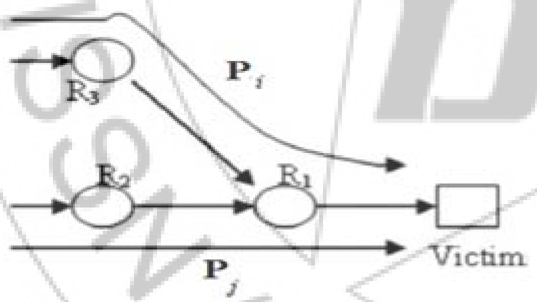

Figure 2: A sample community network diagram

4.1 Network Packet Tracing Algorithm

This algorithm monitors the flow at each router in the network. With the help of this algorithm, each router in the network records the entire flow rate that comes either from client or attacker during non-attack, attack and flash crowd period. In this novel packet tracing algorithm, we are combining Parameters from KDD CUP 99 dataset such as time, duration, protocol, service, flag, source bytes, destination bytes of flows at each router to differentiate DDoS from flash crowd [5]. For network packet tracing purpose, we analyze the four different techniques which are depends on four theorems given below.

1. Network Flow
In a local network or a community network for a given router, we collect the network packets that have the same destination address as one network flow [4].

$$Pi = (pi\,[1], pi\,[2] \ldots\ldots\ldots. pi\,[N]) \quad (1)$$

Here, P*i* represents N number of packets. According to our definition of flow, a router may have many network flows at any given point in time.

2. Flow Strength
For a network flow Pi, consider the length of the network flow be N *(pi[N] >=1)* .
We define the expectation of the flow as the flow strength of Pi. Flow strength represents the average packet rate of a network flow. If pi is a DDoS attack flow, then we also call $A[pi]$ as attack strength [9].

$$A[pi] = 1/N \sum_{i=0}^{N} pi\,[n] \quad (2)\#$$

3. Flow Fingerprint
For a given network flow Pi with length N used to represent the fingerprint as unified representation of Pi, which describes the similarities of different flows [4], [5].

$pi = \{pi'[1], pi'[2] \ldots\ldots pi'[N]\}$

$pi = \{pi'[1]/N*A[pi]*\ pi'[2]/N*A[pi]\ldots\ldots\ pi'[N]/N*A[pi]\}$

$pi = N * *A[pi] *\ pi$

$RES\ pipj = 1/N \sum_{n=0}^{N} pi\,[n] * pj\,[n]$

$$RES\ pipj[K] = 1/N \sum_{n=0}^{N} pi\,[n] * pj\,[n+k] \quad (3)$$

4. Flow Correlation Coefficient
Let pi and pj and (***i # j***) be two network flows with the same length N. We define the correlation coefficient as,

$$FCCpipj[K] = RESpipj[K]/1/N[\sum_{n=1}^{N-1} pi2[n] * \sum_{n=1}^{N-1} pj2[n]]1/2 \quad (4)$$

4.2 Detection Algorithm for similarities in network flow

For this algorithm, we used trapezoidal shape to measure a probability of being an attack identified by each attribute. The fuzzy logic is encoded into four parameters which are, b, c and d. The probability is calculated as shown in Fig.3. And its meaning is described below.

Step 1.Consider, we have to calculate probability of being an attack from condition below, We encode a fuzzy logic for each attribute and normalize the value of each attribute to be in the range of 0.0 to 7.0. The fuzzy logic is encoded in below figure 3.

010	011	100	101
a	b	c	d

Fig.3.Fuzzy Logic

Step 2.Fuzzy encoding for each attribute is done as follows. The parameters a, b, c and d are in range between 0-7 (example: a=2, b=3, c=4, d=5). Each rule considers having n attributes and a class name at the end of string as shown in below Fig.4.

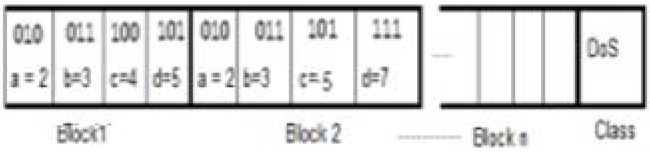

Figure 4: String Encoding

Step 3.Each rule will be mapped to each record in the testing dataset as shown in Fig.5.

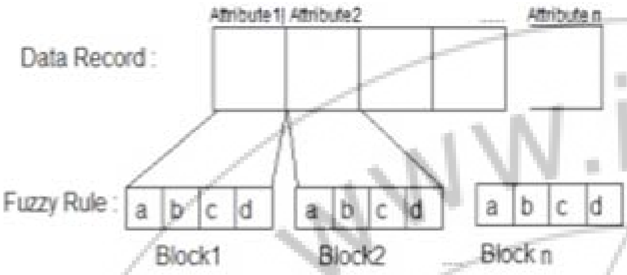

Figure 5: Fuzzy rule with data record

5. Result Analysis

5.1 Implementation

The experiments of this novel proposed system are performed by using weka tool system and KDD CUP 99 dataset. It contains four modules. At the beginning we have to capture packets from different networks and store it on any text file. Then we have to generate the rules based on KDD CUP 99 dataset. After that we have to trace the DDoS attack by using fuzzy logic rules based on similarities. Then we have to calculate detection ratio based on flow correlation coefficient based on fuzzy logic classifier so that we can classify DDoS attack in different types. At the end we will analyse the result expecting increase in rate of detection of DDoS from given flash event.

5.2 Dataset

The experiments are performed by using International Knowledge Discovery Dataset. The KDD CUP 99 dataset is publicly available and considered as a benchmark dataset for testing of various detection algorithms [5]. By using KDD CUP 99 dataset, rather than inserting the attack packets into the normal traces, the labeled attack samples which are obtained by passive monitoring [6]. The KDD CUP 99 datasets consist of two types of dataset: training dataset and testing dataset. Each record of the training data is labeled as either anomalous or normal, which denotes a specific kind of attack. The training dataset contains a total 22 types of attacks and in the testing dataset, 395 dataset has contain additional 15 types of attacks [7].

As we are detecting, sources of DDoS attacks (Smurf, fraggle, Neptune, Teardrop and Ping of Death). After elaborating labeled dataset, it has been found that total number of 41 attributes provides the specifications of the received packets. For this experiment, by using different attributes of packet flows such as time, duration, protocol, service, flag, source bytes, destination bytes at each router to differentiate DDoS from flash crowd.

5.3 Expected Result Set

We are increasing the rate of differentiating DDoS from flash event using flow correlation coefficient compared to existing system. The performance of the network is evaluated in terms of the some metrics trace back time, DDoS detection Ratio based on flow correlation coefficient and Throughput based on flow correlation coefficient. The expected graph of comparing performance of the conventional system and proposed system is given below. As the rate of detection of DDoS from flash event will be increased in the proposed approach of FCC system.

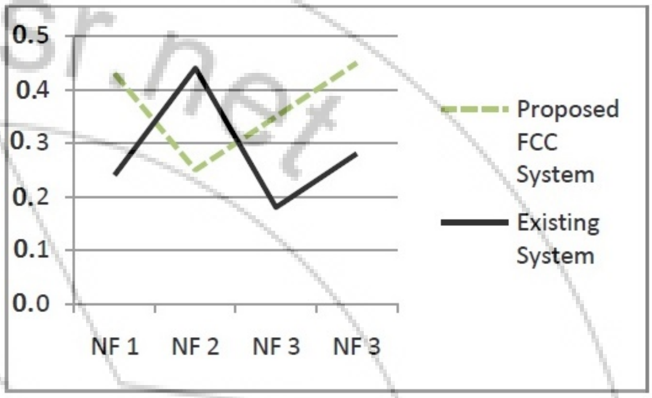

Figure 6: Flow correlation coefficient against network flow

6. Conclusion and Future scope

In this paper, we proposed an enhanced version of FCC System which is an effective and efficient detection and tracing mechanism based on flow correlation coefficient. The proposed method does not need any marking on packets and also any updating of routing software; hence it acts as an independent software module. It also reduces the problem of differentiating the flash crowd i.e. legitimate flow from DDoS attack. From this mechanism, it is proved that by combining parameters from KDD CUP 99 dataset such as time, duration, protocol, service, flag, source bytes, destination bytes of flows at each router, the DDoS attack (malicious flow) can be distinguished from flash crowd, so that there is no probability of rising false alarm. Also, the proposed system can easily detect the actual sources of attack in time and increases effectiveness. In future, it is efficient to apply genetic algorithm to detect all network attacks globally to make network more secure.

References

[1] Shui Yu; Thapngam, T.; Jianwen Liu; Su Wei; Wanlei Zhou, "Discriminating DDoS Flows from Flash Crowds Using Information Distance," Network and System Security, 2009. NSS '09. Third International Conference on 19-21 Oct. 2009

[2] Ke Li; Wanlei Zhou; Ping Li; Jing Hai; Jianwen Liu, "Distinguishing DDoS Attacks from Flash Crowds Using Probability Metrics," Network and System Security, 2009. NSS '09. Third International Conference on 19-21 Oct. 2009.

[3] Thapngam, T.; Shui Yu; Wanlei Zhou; Beliakov, G., "Discriminating DDoS attack traffic from flash crowd

through packet arrival patterns," Computer Communications Workshops (INFOCOM WKSHPS), 2011 IEEE Conference on 10-15 April 2011

[4] Shui Yu; Wanlei Zhou; Weijia Jia; Song Guo; Yong Xiang; Feilong Tang, "Discriminating DDoS Attacks from Flash Crowds Using Flow Correlation Coefficient," Parallel and Distributed Systems, IEEE Transactions on June 2012

[5] Arbor, "IP Flow-Based Technology," http://www.arbornetworks. com, 2011.

[6] Kaur, G.; Varma, S.; Jain, A., "A novel statistical technique for detection of DDoS attacks in KDD dataset," *Contemporary Computing (IC3), 2013 Sixth International Conference on* , vol., no., pp.393,398, 8-10 Aug. 2013

[7] M. Bailey, E. Cooke, F. Jahanian, Y. Xu, and M. Karir, "A Survey of Botnet Technology and Defenses," Proc. Cybersecurity Applications and Technology Conf. for Homeland Security, 2009.

[8] B. Stone-Gross, M. Cova, L. Cavallaro, B. Gilbert, M. Szydlowski, R. Kemmerer, C. Kruegel, and G. Vigna, "Your Botnet Is My Botnet: Analysis of a Botnet Takeover," Proc. ACM Conf. Computer Comm. Security, 2009.

[9] G. Oikonomou and J. Mirkovic, "Modeling Human Behavior for Defense against Flash-Crowd Attacks," Proc. IEEE Int'l Conf. Comm., 2009.

[10] J. Jung, B. Krishnamurthy, and M. Rabinovich, "Flash Crowds and Denial of Service Attacks: Characterization and Implications for CDNs and Web Sites," Proc. 11th Int'l Conf. World Wide Web (WWW), pp. 252-262, 2002

[11] V.L.L. Thing, M. Sloman, and N. Dulay, "A Survey of Bots Used for Distributed Denial of Service Attacks," Proc. SEC, pp. 229-240, 2007.C. Riess and E. Angelopoulou, "Scene illumination as an indicator of image manipulation," Inf. Hiding, vol. 6387, pp. 66–80, 2010.

[12] Srikanth Kandula, Dina Katabi, Matthias Jacob, Arthur Berger., "Botz4sale:Surviving Organized Ddos Attacks That Mimic Flash Crowds.," IEEE Trans. Dependable Secure Computing, vol. 4, no. 1, pp. 56-70, Jan.-Mar. 2007.

[13] G. Carl, G. Kesidis, R. Brooks, and S. Rai, "Denial-of-Service Attack-Detection Techniques," IEEE Internet Computing, vol. 10, no. 1, pp. 82-89, Jan./Feb. 2006.

[14] Y. Chen and K. Hwang, "Collaborative Detection and Filtering of Shrew DDoS Attacks Using Spectral Analysis," J. Parallel Distributed Computing, vol. 66, no. 9, pp. 1137-1151, 2006.

[15] C. Patrikakis, M. Masikos, and O. Zouraraki, "Distributed Denial of Service Attacks," The Internet Protocol J., vol. 7, no. 4, pp. 13-35, 2004.

Author Profile

Dipali B. Pawar is working as Lecturer at ZES's Dnyanganga Polytechnic, Pune from Sept.2011. She has graduated with distinction in Computer Science and Engineering from Sholapur University, Maharashtra, India in 2011 and has worked as Software Developer at Sprance Technologies Pvt. Ltd, Pune on the sponsored project ―Data Acquisition System on SBC‖ for her academic year 2010 to 2011. She has attended workshop on Ethical Hacking at STES's Sinhgad Institute of Technology, Pune for 3 days and completed Microsoft Certification on it. She has worked as Coordinator of Continuous Entertrainership Program (CEP) for Computer Fundamentals. She has participated in National level conference on Computer Networking and Security at Gharda Institute of Technology, Lavel, Ratnagiri, Maharashtra and also in cPGCON-2013 at PICT, Pune. Recently, she is Website Coordinator in her working college. Her research interests are Computer Hardware and Maintenance, Data Structures, Network Security and Web page designing.

Sachin D. Babar is ISTE Life Member. He is graduated in Computer Engineering from Pune University, Maharashtra, India in 2002 and received Master in Computer Engineering from Pune University, Maharashtra, India in 2006. From 2002 to 2003, he was working as lecturer in D.Y. Patil College of Engineering, Pune, India. From 2003 to 2004, he was working as lecturer in Bharati Vidyapeeth College of Engineering, Pune, India. From 2005 to 2006, he was working as lecturer in Rajarshi Shahu College of Engineering, Pune, India. From July 2006, he has been working as an Assistant Professor in Department of Information Technology, STES's Sinhgad Institute of Technology, Lonavala, India. Currently he is pursuing his Ph.D. in Wireless Communication at Center for TeleInFrastruktur (CTIF), Aalborg University, Denmark. He has published 20 papers at national and international level. He has authored two books on subjects like Software Engineering and Analysis of Algorithm & Design. He has received the Cambridge International Certificate for Teachers and Trainers at Professional level under MISSION10X Program. He is IBM DB2 certified professional. His research interests are Data Structures, Algorithms, Theory of Computer Science, IoT and Security.