ADV MATH SCI JOURNAL
Advances in Mathematics: Scientific Journal **9** (2020), no.6, 4103–4114
ISSN: 1857-8365 (printed); 1857-8438 (electronic)
https://doi.org/10.37418/amsj.9.6.92 Spec Issiue on ICAML-2020

VALIDATION OF RATIONALE CHARACTERIZING THE ANOMALIES CAUSING DENIAL OF SERVICE

A. DHINGRA[1] AND M. SACHDEVA

ABSTRACT. The DDoS attack is not just about surging traffic; it includes all the active appliances into its botnet, targeting a specific model of the device. Subsequently, the attacks have become stealthier and sophisticated. Flash event (FE), the legitimate counterpart of DDoS attack, overwhelms the server with requests from legitimate users trying to access information. Both traffic patterns compromise the availability of the target server. This paper highlights the characteristics of DDoS and FE, and evaluates the characteristics discerning DDoS and FE traffic using real-time benchmark datasets. The rationale for both events has been empirically investigated. The paper compares the relative variability of parameters discerning the anomalies by applying the coefficient of variation. The analysis shows that CV of time-interval between request, request sent by each user throughout the event, and request sent by each user in a given time-interval is higher for FE when compared to DDoS attack. From the statistical results obtained, it can be concluded that the FE traffic is more unplanned and non-specific. Thus, though two mentioned events exhibit the same behaviour, the variation in behaviour is observed in quantifying the rationale.

1. INTRODUCTION

Distributed Denial of service attack (DDoS) has become a severe threat to the availability and reliability of the victim server. The attack saturates or somehow blocks the victim server and its infrastructure with spoofed requests using

[1]*corresponding author*

2010 *Mathematics Subject Classification.* 68M12, 94A17.

Key words and phrases. DDoS attacks, Flash Event, network anomaly, IP Address.

the network of compromised systems called bots. Similar deterioration in the performance can be seen when the server experiences a sudden surge in traffic caused by legitimate users trying to access the breaking news or information related to some exciting event. Such a surge in legitimate traffic is termed as Flash Event (FE). It is challenging to decide whether the rise in traffic is an attack or FE as both the anomalies exhibit similar behaviour. The attackers, sometimes, take the advantage of FEs to dupe the defense system and achieve their motives.

A rise in frequency, as well as severity of DDoS and FE, can be attributed to the swift development of the Internet of Things (IoT) in recent years. When these devices get compromised, the intensity of attack increases manifold and, subsequently, affects the functioning of complex network infrastructure. Thus, knowingly or unknowingly, IoT contribute to the volume of network traffic [21].

It is a challenge to discriminate the incoming traffic as DDoS or FE, especially when both show similar traffic behaviour. It becomes even more complicated when the attack is in the disguise of FE. Research has been undertaken to identify and validate the rationale for discriminating the two anomalies [2, 4, 5, 10].

This paper attempts to articulate and conclude that though, DDoS traffic and FE traffic show similar behavior, a few typical characteristics associated with simple attributes, like source IP, can classify the anomalies efficiently. Researchers have investigated some characteristics and used them for discriminating the traffic. Few characteristics, however, are yet to be examined. The paper [1] comprehends the occurrence of DDoS and FE, and studies the behaviour of network traffic during FE, [2] identifies the discerning characteristics of DDoS and FE, [3] empirically validates the additional discerning characteristics using commercial datasets wherever possible, along with the benchmark dataset available online.

The organization of the paper is as follows: Section 2 explains the DDoS attacks and discusses the characteristic features of FE. It examines the work previously done in the field of FE, discrimination of FE from DDoS, and characteristics explored by various researchers. Section 3 explores the characteristics which discern the traffic as legit or illegitimate and validates them. The paper is concluded with the future scope of research in section 4.

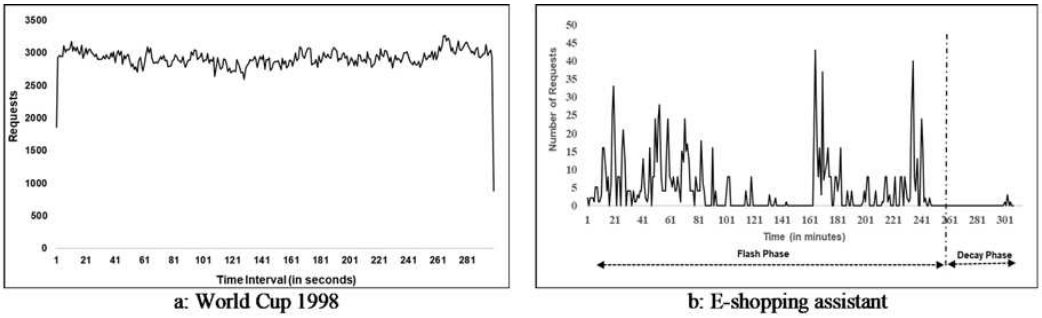
2. DDoS, FE AND RELATED WORK

A DDoS attack uses the network of unsecured, connected devices, including computers, CCTVs, refrigerators, and even cars, to flood the target with server requests. The massive attack of 1.2 Tbps, affecting the internet connections on the US East Coast and Europe, was experienced in October 2016, [12]. The 'Mirai' malware was used to compromise IP-cameras and routers to enhance the load on the network. Two years later, in February 2018, official web-servers of the PyeongChang Winter Olympics were forced to be shut down after the cyber-attack. The website was inaccessible for 12 hours, [16]. The year 2020 saw a number of direct and indirect attacks related to all websites conquering the corona virus pandemic. Online games were hit repeatedly with servers of *Eve Online* being flooded with junk traffic for nine days at a stretch, [24]. As observed, attackers have started using multiple methods of attack, making attacks stealthier than before and easy to dodge the installed defenses causing maximum damage. As per Netscout 14th Annual Worldwide Infrastructure Security Report [**?**], it takes only 5 minutes, on average, for any device to be compromised once plugged online. This scenario has led to the growth in number, as well as the size of botnets. It has enhanced the threat from such jeopardized systems from 17% in 2018 to 21% in 2019 and further increasing in 2020 due to the increased use of the Internet. The frequency of attacks has doubled from 68 per business in 2012 to 130 per business in 2017, [9].

FE is a traffic anomaly exhibiting properties similar to that of DDoS traffic, but involving hundreds and thousands of legit users sending simultaneous requests to the server. Requests may not be able to reach the server or may reach after a notable delay due to several transmission attempts and/or loss in packets causing services to be disrupted entirely or partially, [22, 24]. Thus, FEs are sometimes mistaken to be DDoS attacks. Such an event occurred in March 2020, when Australian authorities reported a DDoS attack on the *MyGov social services* portal. But, it turned out that the site could not cope with the influx of legitimate requests from citizens out of work as a result of the pandemic, [24].

The formation of FE is mainly attributed to the increase in frequency of requests as compared to the increase in the number of users, [13]. It starts and rises gradually during a ramp-up phase, which lasts for a short duration. On reaching a peak, it becomes stable. This is the sustained-phase, during which

load balancing is initiated. When the event is over, the ramp-down phase begins. The traffic starts decreasing slowly and reaches a normal level, [1]. This pattern can be observed in the WorldCup-1998 dataset analyzed for FE, as shown in Figure 1(a). However, [6] proposed the model consisting of two-phases- the Flash Phase and Decay Phase. The authors assert that as the traffic increases, the response time increases, and the users switch over to other websites for obtaining information. Thus, the sustained-phase is short-lived as compared to the other two phases and can be ignored. Figure 1(b) shows the two phases of the FE, as observed in the e-shopping-assistant trace. It shows the surge in activity separated by a period of inactivity, indicating the increase in the number of users and/ or increase in the frequency of requests.



a: World Cup 1998 b: E-shopping assistant

FIGURE 1. Request rate of two different flash events

Thus, the need of the hour is to detect these traffic irregularities at the earliest and characterize them as benign or malicious so that appropriate measures can be taken and the collateral damage be avoided.

Jung et al. in [10] suggested that the FE and DDoS traffic can be discriminated based on the traffic pattern, client characteristics, and file-reference characteristics. A number of previously seen clients are observed during FE, and their distribution among clusters is not uniform when compared to the client pattern observed in DDoS. Kandula et al. in [11], detected the anomaly using puzzles, called CAPTCHA's, and proposed a probabilistic authentication mechanism to differentiate DDoS from FE. Park et al. in [14] proposed a simple yet effective technique, FDD (FE and DDoS Distinguisher) to discriminate the traffic using a randomness check of cluster distribution of clients and dependencies between requests and client distribution in the consecutive time-intervals. Bhatia et al. in [5] proposed the three parameters that can effectively apprehend the traffic

behavior and pattern: Change in the rate of incoming traffic, Change in rate of new source IP address, and distribution of requests among the source IP address. Yu et al. in [17] stated that FE is an unexpected but legitimate rise in traffic that does not have any self-similarity in the flow. Flow-correlation coefficient has been used as a similarity metric for detecting and discriminating malicious and genuine traffic.

Sachdeva et al. in [18] exploited the entropy of traffic clusters to distinguish between DDoS attack and FE. It was observed that in the case of FE, source entropy increases abruptly in comparison to traffic cluster entropy, which shows a marginal increase. Behal et al. in [3] suggested ø-entropy and ø-Divergence metrics for detecting anomalies. The entropy difference between traffic flows discriminates the traffic. Singh et al. in [19], generated application-layer DDoS attacks and identified certain metrics like packet delivery ratio, end-to-end delay, average queuing delay, and average packet rate to identify the higher layer attacks. Patil et al. in [15] proposed a Hadoop based detection methodology called E-Had. Information distance metric, entropy, along with the number of packets for any mapper, is used to detect anomalies and discriminate FE from DDoS. Bhandari et al. in [2] analyzed the nature of FE, the generation of FE traffic, its geographical distribution, the network from which originating, and the duration of FE to formulate its taxonomy. The authors suggested certain characteristics of the traffic, which can be effectively used to discriminate the traffic as a DDoS attack or FE. These characteristics include a change in request rate, the number of different sources IP address, page-access behavior, flow similarity along with the request rate.

3. Investigating characteristics discerning anomalies

It has been observed that in the case of DDoS, the request is fabricated to block the network (infrastructure) intentionally, whereas, in the case of FE, the request is to retrieve the information which unintentionally blocks the traffic. Differentiating the traffic on the basis of underlying motive is, thus, difficult due to the abstract nature of intent. Another difference could be based on the targeted layer of an OSI model. The protocol attacks exploit the weakness of the network layer and/or the transport layer. These attacks consume the processing

capacity of the victim or bring down critical resources like the firewall to disrupt the services.

For DDoS, users are distributed across the globe, depending upon the position of bots, but for FE, traffic can be local, regional, or global depending on the scope of the event. Few of the FE are predictable, like the World Cup traffic or traffic accessing the University results. Sometime FE is unpredictable, just like DDoS attacks, thus cannot be provisioned for in advance. The only way to handle it is by timely detection, discrimination, and then dropping off the malicious requests or by load balancing. Jung et al. in [10] speculated that during a FE, there is an increase in the number of clients. Each client sends a few requests, which is in contrast to a DDoS attack. During the attack, the number of clients and requests per client becomes steady after a few time-intervals. This is due to the involvement of bots. It can be concluded that studying the behavior of traffic assists in identifying unique characteristics of DDoS attack and FE, thus, differentiating the traffic. This section defines and evaluates prominent comparable characteristics.

Rationality and accuracy of the proposed characteristics have been validated using the following datasets chosen appropriately from available sources. 1) Centre for Applied Internet Data Analysis (DoS attack dataset), 2007 for DDoS traffic (henceforth, 'CAIDA'), [7]. 2) WorldCup98 for FE traffic (henceforth 'WorldCup-1998')- The fraction of traffic received on June 24, 1998, day 60, and June 30, 1998, day 66 of the World Cup, has been extracted to represent FE [8]. 3) Real-time dataset of the small Bangalore based start-up, an e-shopping assistance (henceforth, 'e-shopping assistant')- The website experienced FE on February 10, 2015, and the trace of the traffic for this day has been analysed.

3.1. **Request-rate of traffic.** In case of DDoS, all bots involved, send the requests at the same time and repeat after the specified interval. Therefore, request-rate increases/ decreases sharply within a short period. The pattern and features of traffic determine the type of DDoS launched, [20]. Figure 2(a)-2(d) graphically represents a constant-rate attack, increasing rate attack, pulsating attack, and low rate attack, respectively, as observed during analysis of the CAIDA dataset. During FE, request-rate either shows an abrupt increase as

in Figure 1(a) or displays pulsating behavior, as in Figure 1(b), depending on the type of FE.

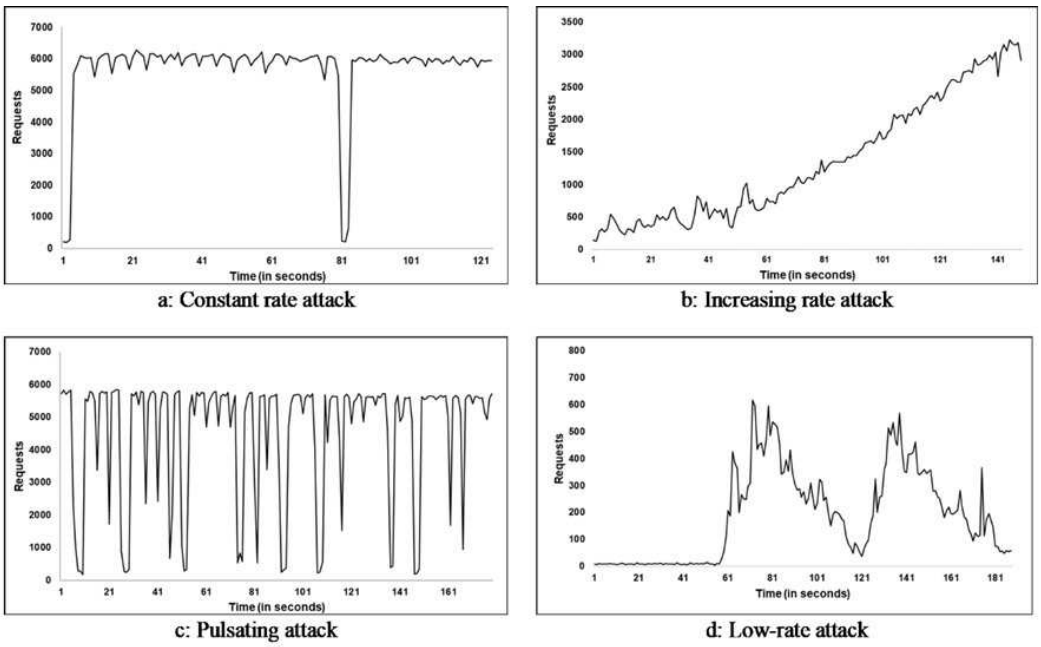


FIGURE 2. Types of DDoS attacks

3.2. **Request per client per time-interval.** In this paper, average and coefficient of variation are used, to compare the two traffic scenarios for similarity in case of a number of requests sent by the client within a given time interval. The average number of requests sent by any client within a given time-interval is shown in Figure 3. In the case of DDoS, average requests per client are uniformly distributed across time-intervals. For FE, however, there is a fluctuating pattern, highlighting that users send requests as per their individual preference, and there is no defined strategy for the same. The statistics provided in Table 1, shows that the CV in the case of DDoS is 0.01, which is much less when compared to that of the e-shopping-assistant, which stands at 0.715, indicating the similarity in the number of requests sent by each client in DDoS traffic.

3.3. **Request per client throughout the event.** The number of requests received from a particular client throughout the course of an event varies for DDoS and FE. Statistics in Table 1 suggest that CV for DDoS is as low as 0.31, and for

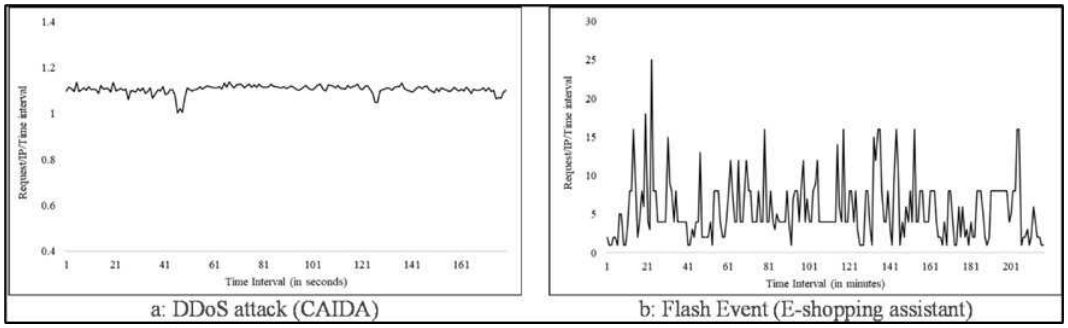


FIGURE 3. Request per client per time-interval

TABLE 1. Statistics for discerning characteristics

Characteristic		**DDoS**	**FE**	**E-shopping assistant**
Type of Packets		ICMP	HTTP-GET	HTTP-GET
Observation Time		187 sec	304 sec	451 min
Req/IP/time-interval	Average	1.106	1.353	5.622
	Coefficient of variation	0.017	0.486	0.715
Req/IP through event	Average	160.695	73.706	25.950
	Coefficient of variation	0.312	2.173	1.843
Time Interval bet req.	Average	1.667	11.182	58.734
	Coefficient of variation	1.319	1.992	2.562
Fresh IPs	First 10 time-intervals	6533	4195	-
	After 10 time-intervals	484	8864	-
	% upto 1st 10 time-int.	93.102%	32.114%	-
	% after 10 time-int.	6.89%	67.89%	-
Time Duration	No. of active IPs (20 time-intervals)	443	6505	33
	No. of IPs active (till last 20 intervals)	3588	262	1
	Total number of IPs	6971	14390	45
	% for 20 time-interval	6.3548%	45.205%	73.33%
	% for more than 20 time-intervals	51.4703%	1.8207%	2.22%

FE, it is 2.17 and 1.84. The lower value of DDoS attacks confirms the presence of pre-programmed bots that send the request in a monotonous pattern throughout the event. For FE, the number of requests per client varies and hence is more random as indicated by the higher value of CV.

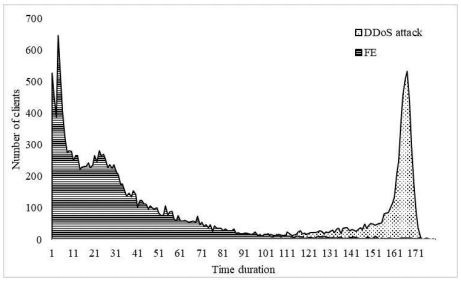


FIGURE 4. Duration of a client during the event

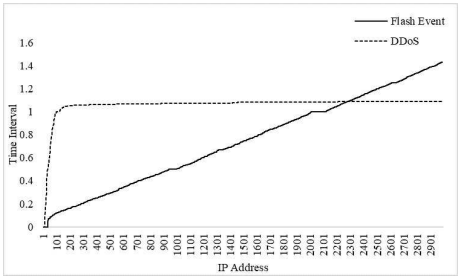


FIGURE 5. Time-interval between requests

3.4. **Duration of a client throughout the event.** Figure 4 illustrates that during a DDoS attack, the same compromised machines remain active throughout the attack period and keep sending request packets. Table 1 indicates that 51.47% of IPs are seen until the last 20 time-intervals. In the case of FE, however, it is the reverse. The cluster at the beginning of the event, as in Figure 4, indicates that very few IPs remain active for a longer duration. Most of the client requests are short-lived. The client stops sending the requests in the case of performance degradation. The data in Table 1 shows that for the two instances of FE-World Cup 1998 and e-shopping-assistant, 45% and 73.3% of IP, respectively, are active only for 20 time-intervals.

3.5. **Time-interval between consecutive requests for a client.** Figure 5 illustrates that the majority of IP addresses (bots) involved in DDoS attacks keep sending requests at a consistent interval of time. Only a few of the IPs show erratic behavior. These could be the legit users trying to access the webserver. In the case of FE, however, the time-interval between requests is inconsistent as per user convenience. Table 1 statistics show that the e-shopping-assistant has the CV of 2.56, indicating a considerable variation in the interval for FE as compared to DDoS, which stands at 1.3.

3.6. **Change in rate of fresh-IP.** It has been suggested by researchers that change in rate of fresh-IP can be used for detection of DDoS attack and differentiating the traffic as legitimate or illegitimate, [2, 23]. During DDoS, the rate of change of fresh IPs is initially high and then declines as the attack proceeds. As the attack proceeds, hardly any fresh IPs join the bandwagon. Table 1, the percentage of fresh IPs observed in the first 10-time intervals in the case of

DDoS is 93.1%, and after that, only 6.89% of fresh IPs send requests to the victim server. For world-cup 1998, however, only 32.1% of Fresh IPs are observed during the first 10-time intervals. 67.89% join through the course of an event with the spread of news.

4. Conclusion

The DDoS attacks and FE resemble in behaviour with each other and hence need to be detected and discriminated at the earliest in order to reduce the collateral damage to the infrastructure involved. The situation becomes challenging when DDoS attack traffic, IoT traffic, and FE traffic are mixed up, and the attacker uses FE to launch an attack. This makes it crucial to understand the features of DDoS and FE and classify them accordingly so that the attack can be detected before it causes any damage to infrastructure. The paper has attempted to understand DDoS attacks and characteristics of different types of FE traffic. After careful study of related literature, characteristics that discern the DDoS attacks from FE have been deduced. The statistical analysis of available datasets, along with the graphical representation, has been presented. This will help researchers to be cognizant of the parameters which can distinguish two anomalies and how they behave during an event. It needs to be understood that certain characteristics of traffic like source IP, are fundamental and stable over time. Thus, the distribution of such attributes for consecutive time intervals can help understand the behavior of traffic.

Future work would be undertaken to use the defined characteristics for the detection of traffic anomalies and discriminating the legitimate and illegitimate traffic. Source IP and Destination IP seem to be the most promising features to be exploited for further research in the detection of anomalies.

ACKNOWLEDGMENT

The authors would like to acknowledge I. K. Gujral Punjab Technical University, Kapurthala, India, for the great support and assistance rendered to carry out this research work. The authors also thank Mr. Ashish Parnami, CTO, e-shopping-assistant, for providing the relevant dataset for validation.

REFERENCES

[1] I. ARI, B. HONG, E. L. MILLER, D. D. E. LONG: *Managing flash crowds on the internet*, 11th IEEE ACM Intl. Symp. on modelling, analysis, and simulation of computer telecommunications systems, IEEE, USA, (2003), 246–249.

[2] A. BHANDARI, A. L. SANGAL, K. KUMAR:*Characterizing flash events and distributed denial of service attacks: an empirical investigation*, Security Commun. Netw., **9**(13) (2016), 2222–2239.

[3] S. BEHAL, K. KUMAR: *Detection of DDoS attacks and flash events using novel information theory metrics*, Comput. Netw., **116** (2017), 96–110.

[4] S. BEHAL, K. KUMAR, M. SACHDEVA: *Characterizing DDoS attacks and flash events: review, research gaps, and future directions*, Comput. Sci. Review., **25** (2017), 101–114.

[5] S. BHATIA, D. SCHMIDT, G. MOHAY: *Ensemble-based DDoS detection and mitigation model*, Proceedings of Fifth International Conference on Security of Information and Networks (SIN '12), ACM, USA, (2012), 79–86.

[6] S. BHATIA, G. MOHAY, D. SCHMIDT, A. TICKLE: *Modelling web-server flash events*, Proceedings of 11th IEEE International Symposium on Network Computing and Applications (NCA), IEEE, (2012), 79–86.

[7] UCSD DDOS ATTACK: *Available online*,(2007), http://www.caida.org/data/passive/ddos-20070804_dataset.xml.

[8] M. ARLITT, T. JIN: *1998 World Cup Web Site Access Logs*, 1998.

[9] GLOBAL RISKS REPORT: *Available online*, (2018), http://www3.weforum.org/docs/WEF GRR18 Report.pdf.

[10] J. JUNG, B. KRISHNAMURTHY, B. RABINOVICH: *Flash crowds and denial of service attacks: characterization and implications for CDNs and web sites*, Proceedings of 11th International Conference on World Wide Web (WWW '02), ACM, USA, (2002), 293–304.

[11] S. KANDULA, D. KATABI, M. JACOB, A. BERGER: *Botz-4-sale: surviving organized DDoS attacks that mimic flash crowds*, Proceedings of 2nd Conference on Symposium on Networked Systems Design and Implementation, USA, (2005), 287–300.

[12] MIRAI ATTACK: *Available online*, (2016), https://www.theguardian.com/technology/2016/oct/26/ddos-attack-dyn-mirai botnet.

[13] J. PAN, H. HU, Y. LIU: *Human behaviour during flash crowd in web surfing*, Physica A: Statistical Mechanics and its Applications, **413**(C) (2014), 212–219.

[14] H. PARK, P. LI, D. GAO, H. LEE, W. ZHOU: *Distinguishing between FE and DDoS using randomness check*, Proceedings of 11th Information Security Conference, LNCS 5222, (2008), 131–145.

[15] N. V. PATIL, C. RAMAKRISHNA, K. KUMAR, S. BEHAL: *E-Had: A distributed and collaborative detection framework for early detection of DDoS attacks*, J. of King Saud Univ.-Comput. and Inform. Sci., (In Press) (2019).

[16] PYEONGCHANG WINTER OLYMPICS: *Available online*, (2018), https://www.independent.co.uk/sport/Olympics/ winterolympics/winter-olympics-pyeongchang-2018-cyber-attack-opening-ceremony- 8204056.html.

[17] S.YU, W. ZHOU, W. JIA, S. GUO, Y. XIANG, F. TANG: *Discriminating DDoS attacks from flash crowds using flow correlation coefficient*, IEEE Trans. on Parallel and Distrib. Sys., **23**(6) (2012), 1073–1080.

[18] M. SACHDEVA, K. KUMAR: *A comprehensive approach to discriminate DDoS attacks from flash events*, J. of Inform. Security and Applications, **26** (2016), 8–22.

[19] K. SINGH, K. KUMAR, P. SINGH: *Impact analysis of application layer DDoS attacks on web services: a simulation study*, Intl. J. of Intel. Engineering Informatics, **5**(1) (2017), 80–100.

[20] X. ZANG, J. GONG, X. HU: *An adaptive profile-based approach for detecting anomalous traffic in backbone*, IEEE Access, **7** (2019), 56920–56934.

[21] C. KOLIAS, G. KAMBOURAKIS, A. STARVOU, J. VOS: *DDoS in the IoT: mirai and other botnets*, Computer, IEEE Computer Society, **50**(7) (2017), 80–84.

[22] A. DHINGRA, M. SACHDEVA: *DDoS detection and discrimination from flash events: a compendious review*, Proceedings of First International Conference on Secure Cyber Computing and Communication (ICSCCC), IEEEXplore, Jalandhar, India, (2018), 518–524.

[23] E. AHMED, G. MOHAY, A. TICKLE, S. BHATIA: *Use of IP addresses for high rate flooding attack detection*, Security and Privacy – Silver Linings in the Cloud, SEC 2010, IFIP Advances in Information and Communication Technology, **330** (2010), 124–135.

[24] O. KUPREEV, E. BADOVSKAYA, A. GUTNIKOV: *DDoS attacks in Q1 2020*, SecureList, 2020.

RESEARCH SCHOLAR
DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
I. K., GUJRAL PUNJAB TECHNICAL UNIVERSITY
KAPURTHALA -144603, INDIA
Email address: avndhingra@gmail.com

ASSOCIATE PROFESSOR
DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
I. K., GUJRAL PUNJAB TECHNICAL UNIVERSITY, KAPURTHALA-144603, INDIA
Email address: monika@ptu.ac.in