International Journal of Engineering & Technology, 7 (2.14) (2018) 5-8

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research Paper

Entity-based Parameterization for Distinguishing Distributed Denial of Service from Flash Events

M A Mohamed[1,*], N Jamil[2], A F Abidin[1], M M Din[2], W N S W Nik[1], A R Mamat[1],

[1]Faculty of Informatics and Computing, Universiti Sultan Zainal Abidin, Besut Campus,Terengganu, Malaysia
[2]Institute of Informatics and Computing in Energy, Universiti Tenaga Nasional, Kajang, Selangor, Malaysia
**Corresponding author E-mail:mafendee@unisza.edu.my*

Abstract

In a perfect condition, there are only normal network traffic and sometimes flash event traffics due to some eye-catching or heart-breaking events. Nevertheless, both events carry legitimate requests and contents to the server. Flash event traffic can be massive and damaging to the availability of the server. However, it can easily be remedied by hardware solutions such as adding extra processing power and memory devices and software solution such as load balancing. In contrast, a collection of illegal traffic requests produced during distributed denial of service (DDoS) attack tries to cause damage to the server and thus is considered as dangerous where prevention, detection and reaction are imminent in case of occurrence. In this paper, the detection of attacks by distinguishing it from legal traffic is of our main concern. Initially, we categorize the parameters involved in the attacks in relation to their entities. Further, we examine different concepts and techniques from information theory and image processing domain that takes the aforementioned parameters as input and in turn decides whether an attack has occurred. In addition to that, we also pointed out the advantages for each technique, as well as any possible weakness for possible future works.

Keywords: *DDoS Attack, Flash Event, Parameter Classification, Packet Entropy, Information Distance.*

1. Introduction

Once connected to a network, a computer system cannot escape from being a target to different logical attacks. Attacks can come from either outside or mostly from the inside. Inside attacks are rather easy and could be mounted by internal employees as a person who already has some access level. Outside attacks can be very challenging, and the occurrence may be due to many different reasons such as rewards, revenge or maybe just for challenges. Computer attacks can cause a different degree of damage from the least of highly severe. The most disastrous would be those that damaging the assets and hence hurting the integrity of the data. Others would be attacking the privacy such that of stealing data. In any cases, the cost can be millions of dollars and loss of reputations.

A group of attacks called denial of service (DoS) that causes the downtime to the system, damaging the availability of the data making it not accessible to users. This attack does not hurt the privacy or integrity, but to service offering entities such as online shopping, online banking and online news, this would be devastating as the service goes unavailable to customers in that not only the transaction is unsuccessful, but in the long term it affects the credibility of the company.

A much more powerful version of DoS attacks has become one of the top security issues affecting networks and disrupting services to legitimate users and recently are mounted by professionals using Botnets of tens of thousands of compromised machines. This resulted to what is referring to as Distributed Denial of Service (DDoS) [1-3]. DDos attack could be mounted from thousands of compromised computers. Attackers defeat the server processing capability and gain control over the server by flooding a lot of messages or requests and make server resources unavailable to legitimate users. Obviously, having a large number of hosts increase the efficiency and effectiveness of DDoS attacks. On the other hand, flash crowds consist of legitimate requests, where the server has the responsibility to handle many requests as possible. Both DDoS and Flash events create abnormal traffic condition. In this regard, discriminating from DDoS and flash crowds are an important topic within the research community.

In this paper, we review various articles on DDoS and FE including those of discriminating them, which were randomly collected from different digital libraries. There are many articles written in this regard proposing different countermeasures to such problem, but unfortunately, most of the proposed countermeasures tend to fail, that is why in our review we summarized the findings/proposed countermeasure from each article and more importantly their gaps. This review will help in providing an insight into the field of DDoS detection, Flash Event (FE) and distinguishing between them. The remainder of this article is structured as follows: Section 2 investigates two topics of interest namely distributed denial of service and flash events, examining their properties and characteristics. Section 3 surveys the advancement in research on detecting the DDoS and discriminating it from FE. Section 4 provides discussion on the current state and possible future research direction in this area. Finally, section 5 concludes findings of this paper.

Copyright © 2018 Authors. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

2. Distributed Denial of Service and Flash Crowds

Distinguishing one thing from the other can be easy in some cases and can be hard to others. Normally, we address this by examining the characteristics of each item and then try to determine which one belongs to each and then grade them from those that are closely identical to those that are extremely contradicting. Since the characteristics itself are volatile in a sense that it can hold a value from a given range, sometimes it can be closely identical, but the other time it can go distantly different, thus making it more difficult to classified. This mean, there is a grey area for which each characteristic belongs to.

Distributed denial of service (DDoS) is an attack, considered the most destructive targeting mass disruption of the service, mounted by the attacker(s) from a large number of computer systems possibly illegally owned, and probably being scattered distant apart across the globe. The objective is to render the victim server into either crashing (complete shutdown) or unavailable state by utilising its scarce resources such as CPU, memory and bandwidth and thus making it unable to fulfil user requests further. The task is to send as many requests to the server from as many different computers as possible when the server becomes exhaustive. System crashing may easily be remedied using costless patches from manufacturer whereas, being left with zero resources available may not have an easy remedy but costly such as adding extra CPU, memory devices or the use of load balancing for those who want to go extra miles into physical solutions.

DDoS attack can be classified from many different perspectives [4-7]. From one angle, there are two types of DDoS attacks, namely high-rate and low-rate. A low-rate DDoS is special in that it is capable of concealing its appearance as it imposes very much similar behaviour to normal traffic. As such, it has a better chance to avoid current anomaly-based DDoS detection systems. A high-rate DDoS attack has been very popular nowadays, reaching illegitimate traffic production of up to 600 Gbps. This volume of fake requests could bring down just any existing websites, and therefore a fast detection and response are certainly imminent. From another, there are two types of DDoS attacks detection techniques namely misuse-based detection which depends on the matching to the known signature, and anomaly-based detection which compare the incoming network behavior to the profiled behavior in real time. In comparison, the misuse-based detection requires to be updated daily with recent most signatures. Meanwhile, the anomaly-based detection needs to monitor and study the network traffic to create a signature with any network traffic that deviates from normal network traffic at some percentage. When this happens, it is considered as an intrusion. In this paper, we focus on anomaly-based detection technique.

Flash crowd is a situation where a server receives numerous requests from legitimate users in a certain period of time due to important or sudden events has taken places such as sports, news, earthquake and tsunami. From the perspective of packet traffic, flash crowd traffic is also legit as that of normal traffic, but it is far from being normal. In general, flash crowd event seems to be identical to that of high-rate DDoS attacks. However, the in-depth investigation may reveal the differences between the two.

If we look at the infrastructure perspective, both events are launched on the same infrastructure, that is via the internet. The scenario consists of a collection of computer system known as the source, a network via with the request is transmitted to the server and the server itself who is responsible for offering the service to users. By studying the source, packets and server characteristics, some classification between DDoS and flash crowd are achievable. We categorized characteristics based on which entities it is attached. In this case, there are three entities. Therefore we have three groups of characteristics.

Obviously, the number of source nodes is going to be large enough so that the rate of production of the request is massive. When we look at the characteristics of the source node, we might look at it individually or collectively in a cluster. Individually, we could investigate the rate of request produced by the same IP address (although IP spoofing is possible), the timing between request, the size, content of each packet, and relation between packets, to name a few. On the other hand, when the characteristics of the source nodes are studied collectively, the physical distribution of source nodes based on the IP, the randomness of packet generations between nodes, the correlation between packets from different nodes are more apparent to be investigated.

Packet traffic on its own comprises of many characteristics. From a transmission point of view, we can study the delays, throughput, sequences. From a content point of view, among packets sent from the same node, we can study the entropy and correlation, whereas for packets sent from different nodes, we can study randomness and probability.

What happen at the server side remains limited to the investigation on the behaviour of resources such as CPU and memory loads during the events. In fact, the server side is the place where we conduct studies on the source node and packet traffic characteristics.

All these direct parameters are manipulated and used singly or in conjunction with others using another comparison concepts, to classify the event under investigation into either DDoS or flash crowd.

3. Classification of DDOS And Flash Crowd

The previous section exposed us to the direct phenomenon that is bounded to the event. These phenomena are defined by parameters that are directly described it. In this section, we review some of the articles that have significant contribution to the advances in the classification of DDoS from the flash crowd. These advances make use of patterns such as randomness and chaotic to manipulate the aforementioned parameters based on the concepts such that found in statistics, soft computing, information theory, and signal processing. Early work on discriminating DDoS from FE was thoroughly studied by [8]. However, the current sophistication of attacks requires more advanced tools and techniques. Some researches were conducted to identify DDoS from normal traffic while other DDoS from flash event.

The first half of this section surveys through research advances in the direction of information theoretical based techniques which in general relies upon two concepts of information entropy and information distance [9].

In order to resolve the inefficiency of fingerprints approach in detecting DDoS and to differentiate it from flash crowd, [10] employed three metrics for information distance which comprises of Jeffery distance, Sibson distance and the Hellinger distance to measure the similarity among traffic flows. Of them three, it was found that the Sibson distance is the most accurate, capable of achieving 65% accuracy in discriminating DDoS flooding attacks from flash crowds.

Some authors [11] proposed two metrics namely generalized entropy and information distance, to detect low-rate DDoS attacks for measuring the normal traffic from the attack traffic. Generalized entropy metric is a generalization of Shannon entropy that can be used to measure the diversification of uncertainty of a sys-

tem. Information distance is based on the Kolmogorow complexity, which is used to measure the different between two objects. In this work, it was shown that the generalized entropy and information distance based technique outperformed those that of traditional Shannon entropy and Kullback-Leibner divergence method at enlarging the adjudication distance. Moreover, the method is capable of lowering the false positive rate.

Another authors [12] proposed an anomaly detection of DDoS using Tsallis Entropy and Lyapunov exponent. Tsallis entropy is one-parameter generalization of Shannon entropy. It is responsible for processing the network traffic of source IPs and destination IPs. These value will be chaotically analyzed using the Lyapunov exponent to determine DDoS detection based on some predefined threshold.

To improve accurateness of previous techniques, authors of [13] proposed an empirical evaluation for both low-rate and high-rate DDoS attacks. For this purpose, the authors make use of various information metrics such as Shannon entropy, Hartley entropy, Renyis entropy, Generalized entropy, Kullback-Leibler divergence. These metrics are used to describe traffic characteristics and to build an effective model for DDoS detection. Based on MIT Lincoln Laboratory and CAIDA DDoS datasets to illustrate the efficiency and effectiveness of each metric for detecting DDoS attacks, the results show that by using an appropriate choice of metrics, there is a chance to clearly segregate between normal traffic from attacks traffics.

Another authors [14] proposed a hybrid detection system for detecting DDoS attacks. In this work, the authors made use of both anomaly-based and signature-based detection methods to enhance the overall detection accuracy. The anomaly part is developed based on multidimensional Gaussian mixture model (GMM), to discriminate normal traffic from the abnormal one. Initially, raw data go through feature extraction process whereby, an expectation maximization algorithm is used for estimating parameters of the GMMs. The probability density function is computed from the input parameters. Furtheron, information distance based on Kullback-Leibler method is used to quantify the distance between two functions and hence decide whether the event is normal or otherwise.

Another version of using information distance was further improved by [15]. In this works, the authors proposed the use of novel information theory-based generalized ø-entropy and ø-divergence metrics by examining the parameters in the packet header, before computing the value of information distant that decide the event type. The method is very sensitive and capable of detecting even a small variation in the network traffic feature. However, this method works well with the assumption that all attack nodes are centralized with the pre-defined identical malicious program.

Further, the second half of this section surveys through techniques for detecting DDoS attack based on those that are used in signal and image processing. With the assumption that the internet traffic is long-range dependent or self-similarity, many research on DDoS detection has been done using Wavelet transform.

Authors of [16] proposed a DDoS detection technique relying upon undecimated discrete wavelet transform and Bayesian analysis. This technique works well in detecting significant changes in variance and frequency in the given time series. However, the capability is limited by its decomposition scale as well as the complexity of the technique itself.

Authors in [17] utilized energy distribution characteristics based on wavelet analysis for detecting the DDoS attack. The fact is, the attack traffic would trigger signification energy distribution deviation as compared to stationary energy distribution for the normal traffic. In this study, the traffic signature is developed from the difference of two successive energy distributions. If this value goes beyond some predefined threshold, a DDoS alarm is raised.

Some authors in [18] exploited the multi-scale characteristics of traffic time-varying signals in that the normal traffic and the attack traffic are different in the range of its frequency band. By depending on the fact that signal can be decomposed into low, middle and high frequencies, by varying the decomposition process adaptively, this technique is capable of detecting DDoS attacks at all frequency ranges. Moreover, by having decomposing threshold, it has an ability to avoid the blindness of wavelet packet decomposition and hence avoid the problem of decomposing scale's self-adaptive selection.

Authors in [19] proposed a technique that decomposes a set of traffic metrics into the time-scale space. All these metrics are represented as a function of time. By correctly choosing the right metrics, and be modelled by Wavelet transform, Lipschitz singularity detection technique can be used to detect an attack. The attack is represented as noise in the signal and wavelet analysis can be used to detect such thing based on decision rule with a predefined threshold value. This technique is advantageous in that it is an efficient and of low computational cost.

Some authors in [20] proposed a new multistage DDoS detection technique that is capable on detecting subtle DDoS attacks at an early stage with high accuracy in real time. Previous works focused on detecting anomaly based on single characteristics which may generate high false negative rate. On the contrary, combined characteristics detection would require high computational and memory costs. At one point in time, this multi-characteristics technique extracts behaviours that significantly effective for detecting DDoS attacks such as the volume of traffic, distributions of characteristics in the headers of packets and average packet length to become a new metric namely network traffic state. At one point in time, a Markov-based state prediction is used to predict the next state and in case of this state is not identical to the real state, the so-called prediction failure has occurred, and one out of many reasons would be due to the DDoS attack. To minimize false negative at this point, a fine-grained singularity detection based on wavelet transform and Lipschitz exponents are used, conceptually similar to that of [19].

Other authors in [21] proposed an advancement in traffic matrix analysis and anomaly detection using two-dimensional Diffusion Wavelet and Principal Components Analysis. Initially, some significant parameters from end-to-end traffic flow are selected into the traffic matrix and are multi-resolution analysed using diffusion wavelet transform to provide meaningful characteristic parameters in different scales. To detect and localize the anomaly, principal components analysis combined with source data similarity techniques. This technique has been demonstrated to successfully do its job in backbone network and it may potentially be used to solve many other network applications.

4. Discussion

In general, discriminating DDoS from flash event is achievable by either showing the two events are sufficiently different or they are insufficiently identical. We can further divide anomaly detection techniques into profiling technique and threshold technique and categorized as to which our technique from information theory and signal processing belongs to.

We observe that every day new technology improvements are introduced just to ease everyone's life. Sometimes these solutions

also come with new vulnerabilities and loopholes that invites attackers for a more challenging task and hence further increases complexities in attacks. Due to this, new techniques to mitigate these risks must also come with the sophistication of which it is evident by this survey.

5. Conclusion

In this paper, we have studied the advancement in the sophistication of techniques used to discriminate DDoS from normal and flash events. The concepts from information theory and image processing were shown to have successfully been implemented for this purpose, although there is still room for improvements and further research into areas such as DDoS in wireless and mobile network.

Acknowledgement

This research is supported partly by the Malaysian Ministry of Higher Education Grant No: FRGS/1/2015/ICT03/UNISZA/02/1[RR141].

References

[1] D. Parwani, A. Dutta, P. K. Shukla. (2017). Prevention Mechanism of DDoS Attacks: A Critical Review. International Journal of Science, Engineering and Technology, 5(3), 99-112.
[2] A. Keshariya, N. Foukia. DDoS Defense Mechanisms: A New Taxonomy. In: Garcia-Alfaro J., Navarro-Arribas G., Cuppens-Boulahia N., Roudier Y. (eds) Data Privacy Management and Autonomous Spontaneous Security. Lecture Notes in Computer Science, vol 5939(2010). Springer, Berlin, Heidelberg.
[3] J. Mirkovic and P. Reiher. (2004). A Taxonomy of DDoS Attack and DDoS Defense Mechanisms. ACM SIGCOMM Computer Communications Review, 34(2), 39-53.
[4] P. Kaur, M. Kumar, A. Bhandari. (2017). A review of detection approaches for distributed denial of service attacks. Systems Science & Control Engineering, 5(1), 301-320.
[5] A. Bhandari, A. L. Sangal and K. Kumar. (2014). Performance Metrics for Defense Framework against Distributed Denial of Service Attacks. International Journal of Network Security, VI, 38-47.
[6] K. Yoohwan, W. C. Lau, M. C. Chuah, H. J. Chao. (2006). PacketScore: a statistics-based packet filtering scheme against distributed denial-of-service attacks. IEEE Transactions on Dependable and Secure Computing, 3(2), 141-155.
[7] A. Hussain, J. Heidemann, C. Papadopoulos. A framework for classifying denial of service attacks. In: Proceedings of the 2003 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications, SIGCOMM 2003 (2003), pp:99-110.
[8] Jung, B. Krishnamurthy, and M. Rabinovich. Flash crowds and denial of service attacks: characterization and implications for CDNs and web sites. Proceedings of the 11th International Conference on World Wide Web (WWW '02), (2002), pp:293–304.
[9] B. Hu, L. Bi, S. Dai. (2017). Information Distances versus Entropy Metric. Entropy, 19(6), 260.
[10] S. Yu, T. Thapngam, J. Liu, S. Wei and W. Zhou. Discriminating DDoS Flows from Flash Crowds Using Information Distance, Third International Conference on Network and System Security, Gold Coast, QLD, (2009), pp:351-356.
[11] Y. Xiang, K. Li and W. Zhou. (2011). Low-Rate DDoS Attacks Detection and Traceback by Using New Information Metrics, IEEE Transactions on Information Forensics and Security, 6(2), 426-437.
[12] X. Ma and Y. Chen. (2014). DDoS Detection Method Based on Chaos Analysis of Network Traffic Entropy, IEEE Communications Letters, 18(1), 114-117.
[13] M. H. Bhuyan, D. K. Bhattacharyya and J. K. Kalita. (2014). Network Anomaly Detection: Methods, Systems and Tools. IEEE Communications Surveys & Tutorials, 16(1), 303-336.
[14] Ö. Cepheli, S. Büyükçorak, G. K. Kurt. (2016). Hybrid Intrusion Detection System for DDoS Attacks. Journal of Electrical and Computer Engineering, Article ID 1075648, 8 pages.
[15] S. Behal, K. Kumar. (2017). Detection of DDoS attacks and flash events using information theory metrics: An empirical investigation. Comput. Commun. 103(C), 18-28.
[16] V. Alarcon-Aquino, J. A. Barria. (2001). Anomaly detection in communication networks using wavelets. IEE Proceedings - Communications, 148(6), 355-362.
[17] L. Li, G. Lee. (2005). DDoS Attack Detection and Wavelets. Telecommun Syst. 28(3-4), 435-451.
[18] J. Gao, G. Hu, X. Yao, R. K. C. Chang. Anomaly Detection of Network Traffic Based on Wavelet Packet. Asia-Pacific Conference on Communications, Busan, (2006), pp:1-5.
[19] M. Hamdi, N. Boudriga. (2007). Detecting Denial-of-Service attacks using the wavelet transform. Computer Communications, 30(16), 3203-3213.
[20] F. Wang, H. Wang, X. Wang, J. Su. (2012). A new multistage approach to detect subtle DDoS attacks. Mathematical and Computer Modelling, 55(1–2), 198-213.
[21] T. Sun, H. Tian, X. Mei. (2015). Anomaly Detection and Localization by Diffusion Wavelet-based Analysis on Traffic Matrix, Computer Science & Information Systems, 12(4), 1361-1374.