Detection of Smurf Flooding Attacks Using Kullback-Leibler-based Scheme

Benamar Bouyeddou[a], Fouzi Harrou[b], *Member, IEEE*
[a]STIC Lab., Department of Telecommunications,
Abou Bekr Belkaid University, Tlemcen, Algeria
Email: bouben81@yahoo.fr, benamarkadri@yahoo.fr

Ying Sun[b], Benamar Kadri[a]
[b]King Abdullah University of Science and Technology,
CEMSE Division, Thuwal, 23955-6900, Saudi Arabia
Email: fouzi.harrou@kaust.edu.sa

***Abstract*—Reliable and timely detection of cyber attacks become indispensable to protect networks and systems. Internet control message protocol (ICMP) flood attacks are still one of the most challenging threats in both IPv4 and IPv6 networks. This paper proposed an approach based on Kullback-Leibler divergence (KLD) to detect ICMP-based Denial Of service (DOS) and Distributed Denial Of Service (DDOS) flooding attacks. This is motivated by the high capacity of KLD to quantitatively discriminate between two distributions. Here, the three-sigma rule is applied to the KLD distances for anomaly detection. We evaluated the effectiveness of this scheme by using the 1999 DARPA Intrusion Detection Evaluation Datasets.**

***Keywords*—ICMP Flood, cyber-attack, KL distance, anomaly detection, DARPA99 dataset.**

I. Introduction

Although internet and IP networks improvement are an indispensable component of our society, until now there exist several attacks on networks that can dramatically affect the security of computers and network systems. Guaranteeing an appropriate level of security and confidentiality is becoming increasingly more important than before. Of course, it is primordial to develop efficient intrusion detection systems for detecting cyber-attacks and protecting cyber systems from unauthorized accesses.

The internet control message protocol (ICMP) is an integral part of the TCP/IP protocols stack which is used for reporting both error messages and control messages [1], [2]. Networks devices, such as routers and destination hosts, send ICMP error messages with delivering IP packets to signal the encountered problems (e.g., unknown network, destination unreachable, and time to live exceeded). The ICMP control messages are usually used by administrators for diagnosis and administration tasks, such as traceroute to find routes and ping for testing the connectivity between two devices [3]. In addition, ICMPv6 play a central role in IPv6 networks. It performs new functionality, such as neighbor discovery, fragmentation, addresses auto-configuration and substitutes other IPv4 protocols (e.g., ARP and IGMP) [4].

Although the increasing development of information technology and sophisticated network's devices, networks system are still vulnerable to malicious and illegitimate users that can generate ICMP attacks. By exploiting this weakness, several attacks against IPv4 and IPv6 networks are based on both versions of ICMP including the scanning attack, the redirection attack, the operating system fingerprinting attack and a lot of DOS/DDOS attacks [2], [5], [6].

Due to its importance, the detection of cyber attacks problem has received much attention from researchers recently. Xiang et al. [7] proposed two approaches based on the generalized entropy and the information distance metric for detecting low-rate DDoS attacks. However, the implementation of these approaches requires the collaboration of all routers in the networks to identify low-rate DDoS, which make them difficult to achieve in practice. In [8], François et al. (2012) proposed an intrusion prevention system (IPS) called FireCol to protect the legitimate users by forming rings composed of IPSs of different internet service providers. This approach is within fully distributed collaborative intrusion detection systems [9]. However, it is challenging to provide the same detection performance as centralised systems [9]. Bhatia (2016) [10] introduced an Ensemble-based model with an Exponentially Weighted Moving Average (EWMA) scheme to detect DDOS attacks in the network traffic and isolate flash events. This approach uses traffic characteristics (i.e., packets and IP address), servers load (CPU and memory) and the correlation between them. AlEroud and Alsmadi (2017) [11] introduced an inference-based intrusion detection by integrating K-Nearest Neighbor and graph theory to discriminate DDoS attacks and benign flows in SDNs. Nezhad et al. (2017) [12] used an autoregressive integrated moving average (ARIMA) time series model to detect DoS and DDoS attacks for IP networks. This approach monitors at each time point the ration between the number of packets and the number of IP sources. In intrusion detection, machine learning turn out to play an important role [13], [14]. Bhaya and EbadyManaa (2017) [13] used data mining to design an unsupervised clustering algorithm EM-CURE to identify DDoS attacks in network flow. This approach has been evaluated uing three datasets, DARPA2000, CAIDA2007 and CAIDA2008 datasets. Zekri et al. (2017) [15] proposed an approach based on C45 machine learning algorithm with signature detection approaches to detect DDOS attacks in cloud . This approach focuses on monitoring layer 3 and layer 4 in the OSI 7-layer model. Recently several approaches have been proposed to detect cyber intrusion based on deep learning framework [14], [16]. For instance, in [17], Niyaz et al. (2016) proposed deep learning approach based on a stacked autoencoder to detect DDoS in a software defined network. However, deep learning-based intrusion detection approaches are computationally costly and require huge training data.

The ICMP-based DOS/DDOS attacks are becoming one of the most security issues, particularly, in IPv6 and next-generation of networks [6]. These attacks, which can be easily launched by non-flooding or by flooding attacks, can generate a serious damage to cyber systems [18], [19]. The non-flooding attacks, such as ping of death, some scenarios of blind connection-reset, blind throughput-reduction, and blind performance degrading [5], use ICMP messages to exploit some protocol's vulnerability; one message is, generally, enough to crash the targeted victim. The flooding attacks attempt to saturate the targeted victim with a large traffic of ICMP messages (e.g., ping flood, router advertisement, neighbor solicitation and smurf) [18].

Numerous defense mechanisms against ICMP-based attacks have been developed [20]. Indeed, Most of the ICMPv4 messages, such as ping and broadcast are commonly blocked (filtered) in IPv4 networks. However, this strategy cannot be applied to IPv6 networks because it is strictly indispensable to include the ICMPv6 protocol in an IPv6 network to guarantee correct functionalities. Most core functionalities, such as neighbor discovery and router discovery are associated with ICMPv6 [4]. Internet Engineering Task Force (IETF) proposed the Secure Neighbor Discovery (SeND) to protect ICMPv6 used in the Neighbor Discovery Protocol operations [21].

Vinothkumar et al. used the Halting Anomalies with Weighted ChoKing (HAWK) algorithm to extend the FireCol system for preventing low rate DDOS ICMP flooding attacks. However, this algorithm is based on the assumption that the attackers had a regular behavior, which is not always the case in the practice [22]. Bellovin et al. propose a new ICMP message named the ICMP traceback message, which is generated arbitrarily by routers to track the source of DDOS attacks [23]. But with the high overhead caused by this type of messages, most of attackers cannot be accurately located. Furthermore, hackers can reproduce these messages themselves to perform new kind of DDOS attacks. Beck et al. [24]implemented the Neighbor Discovery Protocol Monitoring (NDPMon), an enhanced version of the IPv4 Adress Resolution Protocol watch (ARPwatch), to assure the network security and detect attacks related to NDP in IPv6 networks. Unfortunately, experimental results showed that this tool has many limitations, particularly; it takes a long time in the training phase. Zulkiflee et al. [25] applied support vector machine (SVM) to select the most relevant features of traffic network (e.g., time of arrival, IP addresses, ports numbers, and protocols) to make difference between the normal traffic and DDOS attacks. Nevertheless, the performance of this method depends on the selected features and can lead to false classifications if the input features are not properly chosen.

This paper proposes an efficient approach based on Kullback-Leibler Divergence (KLD) to detect ICMP flooding DOS/DDOS attacks. The KLD metric has been extensively used in broad applications including classification, speech and image recognition [26], [27], telecommunication [28], industry [29], [30], and medicine [31]. KLD is commonly used to compute the dissimilarity between two probability distributions, which make it very useful as an anomaly detection metric. The KLD distance between the distribution of the abnormal samples and the distribution of normal samples (training) is much larger than the distance between the distribution of normal samples and the reference distribution of the training data. In other words, KL distance becomes closer to zero under the anomaly-free data, whereas a larger KL distance values are obtained in the presence of anomalies. Due to its high sensitivity, KLD can be used to reveal even small deviations from the normal behavior of the monitored system. Hence, ICMP flooding attacks can be done by checking the distance between distributions of new samples and training observations. To do so, we introduce the KLD-Shewhart monitoring chart by exploiting the sensitivity of KLD to changes and the popular three-sigma rule to fix the decision threshold. The proposed scheme has been evaluated by using the 1999 DARPA Intrusion Detection Evaluation Dataset [1].

The remainder of this paper is organized as follows. Section II briefly reviews the ICMP Smurf attacks. Section III introduces the proposed KLD-Shewhart chart. Section IV investigates the KLD-Shewhart chart's effectiveness in detecting ICMP Smurf attacks using DARPA99 dataset. Finally, conclusions and discussions are presented in Section V.

II. ICMP SMURF ATTACKS

Nowadays, DOS and DDOS attacks are becoming a major risk that seriously affects internet service availability around the world. The main objective behind DOS and DDOS attacks is to exhaust the targeted victim's resources in such a way to render them inaccessible to their legitimate users [32]. Figure 1 shows an example of DDOS attacks.

The ICMP amplification attack, which is commonly named Smurf attack, is one of the most common DDOS attacks in IPv4 and still exists in IPv6 networks [6]. The Smurf attack exploits both pings' messages ECHO-REQUEST and ECHO-REPLY which are usually used by administrators for diagnostic purpose. Figure 2 shows the general format of an ICMP message. In ICMPv4, ECHO-REQUEST and ECHO-REPLY messages are identified by types 8 and 0 respectively, and in ICMPv6 their types are 128 and 129 respectively [18].

In similar manner to many other types of DDOS amplifications attacks, the Smurf attack is performed by using numerous hosts to send simultaneously a large traffic of messages, which are in this case the ICMP ECHO-REPLY messages, to overwhelm the targeted victim.

In Smurf attack, first attacker spoofs the IP address of the victim and then sends through the broadcast servers a great

[1] https://www.ll.mit.edu/ideval/data/1999data.html

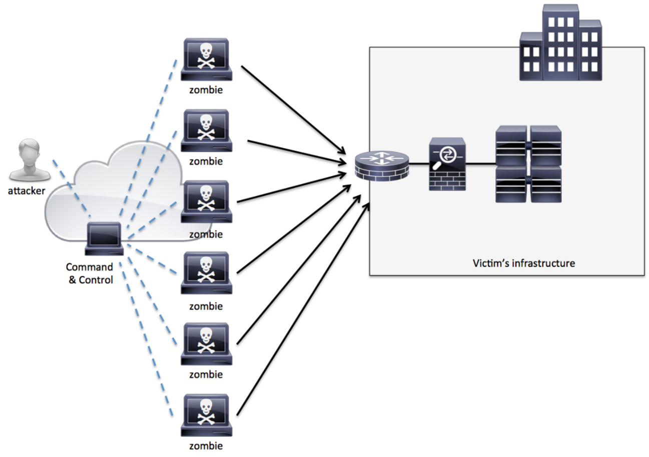
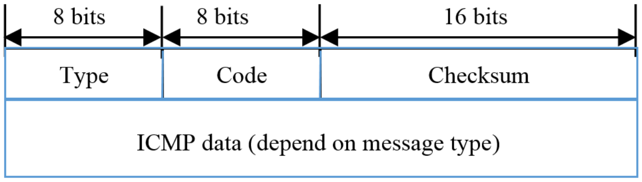


Figure 1. Example of a DDOS attack.



8 bits	8 bits	16 bits
Type	Code	Checksum
ICMP data (depend on message type)		

Figure 2. General format of both ICMPv4 and ICMPv6 messages.

number of ICMP ECHO-REQUEST messages to the all hosts within the broadcast domain (Figure 3). Most of these hosts will accept the ECHO-REQUESTs and reply with the ICMP ECHO-REPLY messages to victim's IP address. There could potentially be hundreds of machines in a broadcast network, which will increases the traffic in function of the number of responding hosts. Therefore, a whole bunch of ECHO-REPLY messages which are simultaneously received by the victim, resulting in its saturation and causing network degradation or a total denial of service [5].

III. KLD-BASED ANOMALY DETECTION

In this section, we present an efficient anomaly detection scheme using the Kullback-Leibler Divergence (KLD) metric. The KLD is a well-known metric to quantitatively measure the distance between two probability distributions. This metric has been extensively used in several fields including information theory, statistical inference, and data mining. The KLD between two probability density functions $p_1(x)$ and $p_2(x)$ is defined as:

$$\begin{aligned} KLD(p_1\backslash\backslash p_2) &= \int_{\mathbb{R}^{d_x}} p_1(x)\log\left[\frac{p_1(x)}{p_2(x)}\right]dx \\ &= \mathbb{E}_{p_1}\left[\log\left(\frac{p_1(x)}{p_2(x)}\right)\right], \end{aligned} \tag{1}$$

Let us consider the particular case when $p_1(x)$ and $p_2(x)$ are Gaussian, $p_1{\sim}\mathcal{N}(\mu_0,\sigma_0)$ and $p_2{\sim}\mathcal{N}(\mu_1,\sigma_1)$, where μ_0 and μ_1 are the means and σ_0^2, σ_1^2 are the variances for p_1 and p_2, the closed form of KLD is [33],

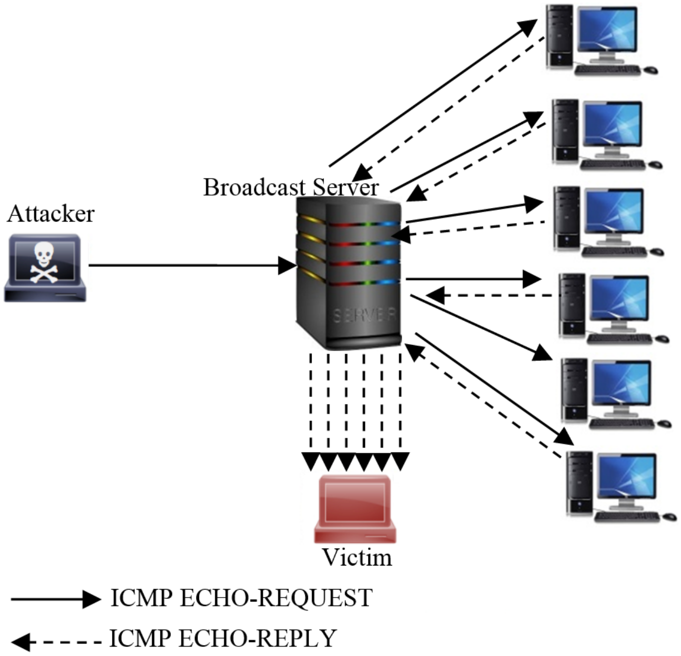


Figure 3. ICMP Smurf attack.

$$\begin{aligned} KLD(p_1\backslash\backslash p_2) &= \frac{1}{\sigma_0\sqrt{2\pi}}\int \exp\left(\frac{(x-\mu_0)^2}{2\sigma_0^2}\right)\left[\log\frac{\sigma_1}{\sigma_0}\right. \qquad (2) \\ &\quad \left. -\frac{(x-\mu_0)^2}{2\sigma_0^2}+\frac{(x-\mu_1)^2}{2\sigma_1^2}\right]dx \\ &= \frac{(\mu_1-\mu_0)^2}{2\sigma_1^2}+\frac{1}{2}\big(\log\frac{\sigma_1^2}{\sigma_0^2}+\frac{\sigma_0^2}{\sigma_1^2}-1\big). \qquad (3) \end{aligned}$$

When $p_1(x)$ and $p_2(x)$ are similar, KLD is close to zero due measurement noise and uncertainty. Otherwise, large values of KLD reflect a great deviation between the two distributions. This makes KLD very useful in anomaly detection framework. Here, we used KLD as an anomaly indicator. In absence of anomalies, KLD becomes closer to zero, whereas a larger KLD value is obtained under the presence of anomalies that lead to the presence of a potential attack. To fix the decision threshold, we used Shewhart chart, also known as the three-sigma rule. The upper control limit is determined as,

$$UCL_{KLD} = \mu_0^{KLD} + 3\sigma_0^{KLD}, \tag{4}$$

where μ_0^{KLD} and σ_0^{KLD} are the mean and standard deviation of KLD measurements under anomaly-free case. When the ith KLD value is beyond the upper threshold, UCL_{KLD}, then we claim that there is an anomaly. Otherwise, the supervised network is considered in control.

IV. EXPERIMENTAL RESULTS

In this section, we assess the ability of the KLD-based Shewhart monitoring scheme to detect ICMP-based Smurf attacks and compared its efficiency with the conventional Shewhart scheme. To do so, we used the IP network traffic of the DARPA99 dataset.

A. *DARPA99 dataset:*

The DARPA 99 dataset is a commonly used intrusion detection benchmark for validating approaches and mechanisms in cyber attacks detection. It has been designed by Lincoln Laboratory at MIT in cooperation with Defense Advanced Research Projects Agency (DARPA) [2]. This dataset consists of 8Gbytes of inside and outside raw tcpdump data, five weeks of traffic captured in a real network which simulates the real network of a military Air Force base. The trace includes, among much other class of attacks, some attacks based on the ICMP protocol, such as the DDOS Smurf attack.

B. *Detection of ICMP Smurf attack:*

From DARPA99's trace that provides the raw traffic of all protocols and their own messages, we have extracted a new dataset contains only the ECHO-REPLY messages (DARPA99/ECHO-REPLY). Specifically, the extracted data comprises three weeks of training data with normal ICMP traffic generated by network's devices when reporting error messages or when launching some diagnostic operations. Also, the extracted data contains two weeks of testing data with anomalous ICMP traffic. To carry out more experiments, we also simulated ICMP Smurf attacks with low intensity and high intensity. The flowing steps recapitulate our adopted procedure to detect the Smurf attacks using KLD-Shewhart and Shewhart charts.

(1) Training data DARPA99/ECHO-REPLY are used to compute the upper control limit.

(2) We compute the decision statistics of KLD-Shewhart scheme.

(3) For new sample, if the decision statistic exceeds the control limit, then ICMP Smurf attack is declared.

1) Case study (A)- ICMP Smurf attacks with low intensity: In this case study, we investigate the capability of KLD-Shewhart scheme to detect low rate DDOS attacks. To do so, we introduced simulated Smurf attacks to the attack-free DARPA 99 traffic data. We injected two minutes of low intensity ICMP Smurf attacks each three hours; the victim receives about 20 ECHO-REPLY messages in the observation time. Monitoring results of the two scharts are shown in Figure 4(a-b). The results show that KLD-Shewhart correctly detected all introduced ICMP Smurf attacks. Shewhart chart can detect these attacks but with few false alarms.

2) Case study (B)- DARPA 99 ICMP Smurf attacks (DARPA (w5,d1)): In this case study, we evaluate the performance of KLD-Shewhart and Shewhart in detecting ICMP Smurf attacks occured on week 5, day 1 of DARPA 99 dataset. There are two attacks occurred in week 5, day 1. The first starts at 09h33mn00s against Pascal (@:172.16.112.50) with a duration of 2mn. The second was at 13h18mn12s against the Marx (@IP: 172.16.114.50) for 1s.

Monitoring results are shown in Figure 5(a-b). As it is expected, the two charts can detect these attacks which are characterized by their high intensity.

[2] https://www.ll.mit.edu/ideval/data/1999data.html

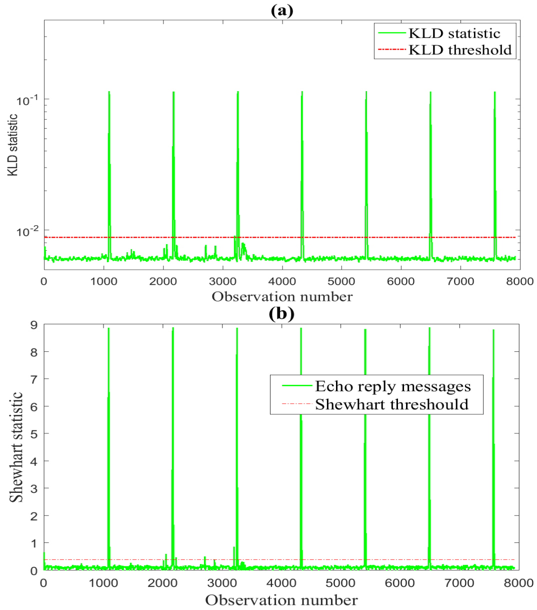


Figure 4. Detection results of (a) KLD-Shewhart scheme and (b) Shewhart scheme in the presence of ICMP Smurf attack with low intensity, case (A).

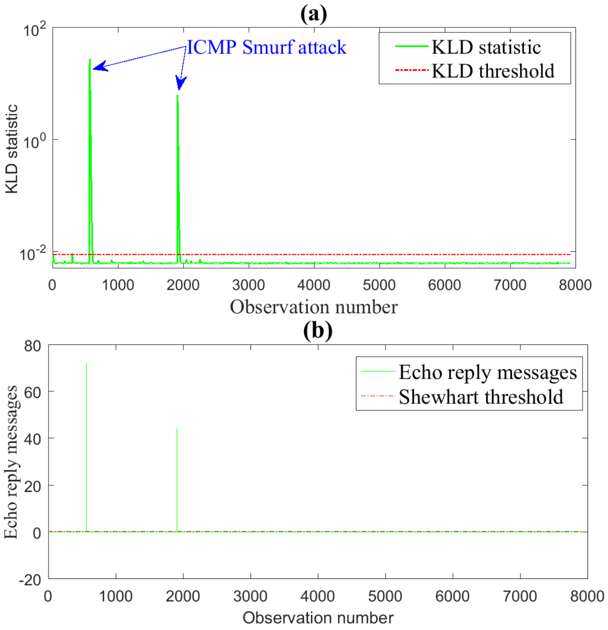


Figure 5. Detection results of (a) KLD-Shewhart scheme and (b) Shewhart scheme in the presence of ICMP Smurf attack in DARPA 99, case (B).

V. Conclusion

A methodology to detect ICMP Smurf distributed denial of service attacks based on Kullback-Leibler divergence (KLD) is proposed in this paper. Here, detection of ICMP Smurf attacks is addressed as an anomaly detection problem. We exploited the capacity and sensitivity of KLD metric in quantifying the dissimilarity between two distributions to detect Smurf attacks. Specifically, we proposed KLD-Shewhart scheme by applying Shewhart scheme to the KLD metric between training data and new testing samples. The proposed method is easy to implement and to understand. We evaluated the efficiency of KLD-Shewhart using the DARPA99 dataset. The proposed method KLD-Shewhart showed better performance compared to Shewhart chart. In future work, we plan to extend this approach in detecting other types of cyber-attacks using updated and realistic traffic network.

Acknowledgement

The research reported in this publication was supported by funding from King Abdullah University of Science and Technology (KAUST) Office of Sponsored Research (OSR) under Award No: OSR-2015-CRG4-2582. The authors (Benamar Bouyeddou and Benamar Kadri) would like to thank the STIC Lab, Department of Telecommunications, Abou Bekr Belkaid University for the continued support during the research.

References

[1] A. Ingle and M. Awade, "Intrusion detection for icmp–flood attack," *Int J Comput Sci Inf Technol*, vol. 1, no. 1, pp. 1–4, 2013.

[2] B. Sieklik, R. Macfarlane, and W. J. Buchanan, "Evaluation of TFTP DDoS amplification attack," *computers & security*, vol. 57, pp. 67–92, 2016.

[3] J. Postel, "Internet control message protocol specifics tion," *HFG792*, 1981.

[4] A. Conta and M. Gupta, "Internet control message protocol (icmpv6) for the internet protocol version 6 (ipv6) specification," 2006.

[5] F. Gont, "ICMP attacks against TCP," 2010.

[6] O. E. Elejla, M. Anbar, and B. Belaton, "ICMPv6-based DoS and DDoS attacks and defense mechanisms," *IETE Technical Review*, vol. 34, no. 4, pp. 390–407, 2017.

[7] Y. Xiang, K. Li, and W. Zhou, "Low-rate ddos attacks detection and traceback by using new information metrics," *IEEE transactions on information forensics and security*, vol. 6, no. 2, pp. 426–437, 2011.

[8] J. François, I. Aib, and R. Boutaba, "Firecol: a collaborative protection network for the detection of flooding ddos attacks," *IEEE/ACM Transactions on Networking (TON)*, vol. 20, no. 6, pp. 1828–1841, 2012.

[9] X.-F. Chen and S.-Z. Yu, "CIPA: A collaborative intrusion prevention architecture for programmable network and SDN," *Computers & Security*, vol. 58, pp. 1–19, 2016.

[10] S. Bhatia, "Ensemble-based model for ddos attack detection and flash event separation," in *Future Technologies Conference (FTC)*. IEEE, 2016, pp. 958–967.

[11] A. AlEroud and I. Alsmadi, "Identifying cyber-attacks on software defined networks: An inference-based intrusion detection approach," *Journal of Network and Computer Applications*, vol. 80, pp. 152–164, 2017.

[12] S. M. T. Nezhad, M. Nazari, and E. A. Gharavol, "A novel DoS and DDoS attacks detection algorithm using ARIMA time series model and chaotic system in computer networks," *IEEE Communications Letters*, vol. 20, no. 4, pp. 700–703, 2016.

[13] W. Bhaya and M. EbadyManaa, "DDoS attack detection approach using an efficient cluster analysis in large data scale," in *New Trends in Information & Communications Technology Applications (NTICT), 2017 Annual Conference on*. IEEE, 2017, pp. 168–173.

[14] N. Shone, T. N. Ngoc, V. D. Phai, and Q. Shi, "A deep learning approach to network intrusion detection," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 2, no. 1, pp. 41–50, 2018.

[15] M. Zekri, S. El Kafhali, N. Aboutabit, and Y. Saadi, "DDoS attack detection using machine learning techniques in cloud computing environments," in *Cloud Computing Technologies and Applications (CloudTech), 2017 3rd International Conference of*. IEEE, 2017, pp. 1–7.

[16] T. Tang, S. A. R. Zaidi, D. McLernon, L. Mhamdi, and M. Ghogho, "Deep recurrent neural network for intrusion detection in sdn-based networks," in *2018 IEEE International Conference on Network Softwarization (NetSoft 2018)*. IEEE, 2018.

[17] Q. Niyaz, W. Sun, and A. Y. Javaid, "A deep learning based DDoS detection system in software-defined networking (SDN)," *arXiv preprint arXiv:1611.07400*, 2016.

[18] K. R. Fall and W. R. Stevens, *TCP/IP illustrated, volume 1: The protocols*. addison-Wesley, 2011.

[19] A. Garg and A. N. Reddy, "Mitigation of dos attacks through QoS regulation," *Microprocessors and Microsystems*, vol. 28, no. 10, pp. 521–530, 2004.

[20] C. Douligeris and A. Mitrokotsa, "Ddos attacks and defense mechanisms: classification and state-of-the-art," *Computer Networks*, vol. 44, no. 5, pp. 643–666, 2004.

[21] J. Arkko, J. Kempf, B. Zill, and P. Nikander, "Secure neighbor discovery (send)," Tech. Rep., 2005.

[22] M. V. Kumar and R. Udayakumar, "Identifying and Blocking High and Low Rate DDOS ICMP Flooding," *Indian Journal of Science and Technology*, vol. 8, no. 32, 2015.

[23] S. M. Bellovin, M. Leech, and T. Taylor, "Icmp traceback messages," 2003.

[24] F. Beck, T. Cholez, O. Festor, and I. Chrisment, "Monitoring the neighbor discovery protocol," in *Computing in the Global Information Technology, 2007. ICCGI 2007. International Multi-Conference on*. IEEE, 2007, pp. 57–57.

[25] M. Zulkiflee, M. Azmi, S. Ahmad, S. Sahib, and M. Ghani, "A framework of features selection for ipv6 network attacks detection," *WSEAS Trans Commun*, vol. 14, no. 46, pp. 399–408, 2015.

[26] N. Singh and R. Agrawal, "Combination of kullback–leibler divergence and manhattan distance measures to detect salient objects," *Signal, Image and Video Processing*, vol. 9, no. 2, pp. 427–435, 2015.

[27] A. Karine, A. Toumi, A. Khenchaf, and M. El Hassouni, "Target recognition in radar images using weighted statistical dictionary-based sparse representation," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 12, pp. 2403–2407, 2017.

[28] D. Olszewski, "Fraud detection in telecommunications using kullback-leibler divergence and latent dirichlet allocation," in *International Conference on Adaptive and Natural Computing Algorithms*. Springer, 2011, pp. 71–80.

[29] F. Harrou, Y. Sun, and M. Madakyaru, "Kullback-leibler distance-based enhanced detection of incipient anomalies," *Journal of Loss Prevention in the Process Industries*, vol. 44, pp. 73–87, 2016.

[30] F. Harrou and Y. Sun, "Enhanced anomaly detection via PLS regression models and information entropy theory," in *Computational Intelligence, 2015 IEEE Symposium Series on*. IEEE, 2015, pp. 383–388.

[31] A. S. Leonard, D. B. Weissman, B. Greenbaum, E. Ghedin, and K. Koelle, "Transmission bottleneck size estimation from pathogen deep-sequencing data, with an application to human influenza a virus," *Journal of virology*, vol. 91, no. 14, pp. e00 171–17, 2017.

[32] B. Bouyeddou, F. Harrou, Y. Sun, and B. Kadri, "Detecting syn flood attacks via statistical monitoring charts: A comparative study," in *Electrical Engineering-Boumerdes (ICEE-B), 2017 5th International Conference on*. IEEE, 2017, pp. 1–5.

[33] L. Pardo, *Statistical inference based on divergence measures*. CRC press, 2005.