DDOS attacks detection based on attention-deep learning and local outlier factor

Abdelkader Dairi
Computer Science Department
University of Science and Technology
Mohamed Boudiaf Oran , Algeria
Email: abdelkader.dairi@univ-usto.dz

Belkacem Khaldi
Ecole Superieure en Informatique
Sidi Bel Abbes
Email: b.khaldi@esi-sba.dz

Fouzi Harrou, Ying Sun
King Abdullah University of Science
and Technology
CEMSE Division
Thuwal,Saudi Arabia
Email: fouzi.harrou.2016@ieee.org

***Abstract*—One of the most significant security concerns confronting network technology is the detection of distributed denial of service (DDOS). This paper introduces a semi-supervised data-driven approach to the detection of DDOS attacks. The proposed method employs normal events data without labeling to train the detection model. Specifically, this approach introduces an improved autoencoder (AE) model by incorporating a Gated Recurrent Unit (GRU) based on the attention mechanism (AM) at the encoder and decoder sides of the AE model. GRU enhances the AE's ability to learn temporal dependencies, and the AM enables the selection of relevant features. For DDOS attacks detection, the local outlier factor (LOF) anomaly detection algorithm is applied to extracted features from the improved AE model. The performance of the proposed approach has been verified using DDOS publically available datasets.**

***Index Terms*—Cybersecurity, Distributed Denial of Service, autoencoder, self-attention, recurrent neural network**

I. INTRODUCTION

To address the drawbacks of traditional networks, SDN (Software-Defined Networking) has arisen in recent years as a new paradigm that decouples control planes from data planes. Despite this, SDN remains vulnerable to certain vulnerabilities and attacks that plague traditional network systems [1]. Furthermore, a slew of security flaws and new vulnerabilities that attackers might use to conduct a variety of malicious attacks could compromise the entire SDN system [1], [2].

DDoS (Distributed Denial of Service) is among the common dangerous attacks targeting the security of both classical networking and SDN systems [3]. It aims at flooding the target networks with a huge volume of malicious traffic, preventing end-users from accessing network services. Furthermore, with the advancement of IoT, many devices are now connected to the internet. As a result, many types of DDoS threats may be exploited through bots devices, making it even harder to detect [1], [4]. These attacks deplete network resources quickly and can overwhelm different SDN layers, such as the between-controller-application layer channel. In addition, the fact that SDN has a focal point of failure, any DDoS attack targeting the SDN might bring the entire network down at the same time. This could result in a significant financial loss for some companies and organizations and a loss of client trust. As a result, intrusion detection systems capable of detecting intruders against these types of attacks become critical for ensuring network security [1], [3], [4], [5], [6].

There has been an increasing amount of works in the literature in recent years to provide secure solutions for the aforementioned types of attacks. The majority of them are data-driven machine learning (ML) solutions. Shallow learning approaches such as k-Nearest Neighbor [7], Support Vector Machine [8], Self-Organizing Map [9], and others were used in previous ML solutions. These have a number of drawbacks, including a considerable reliance on feature engineering and feature selection, a high rate of false alarms, and a limited ability to deal with the large-scale intrusion classification data in the current complex network environments.

Meanwhile, Deep Learning (DL) has become a popular topic among researchers in recent years as artificial intelligence advances, demonstrating its immense potential in dealing with anomalies and intrusion detection in general [10], [11], [12], as well as safeguarding networks and SDNs systems. For instance, Ma et al.[13] proposed a network intrusion detection framework called SCDNN, which combines the spectral clustering (SC) technique with the deep neural network (DNN) model. Yin et al.[14] implemented RNN-IDS, a recurrent neural networks-based network intrusion detection scheme that employs binary and multi-class classification. Following the same tendency, Li et al.[15] combined LSTM (Long Short-Term Memory) with GRU (Gated Recurrent Unit) to provide an accurate network intrusion detection system. Auto-encoders based solutions have also been used recently. For instance, Lopez-Martin et al.[16] developed ID-CVAE, a conditional variational auto-encoder-based intrusion detection system capable of recovering missing features from incomplete training data sets. Shone et al.[17] proposed S-NDAE, a powerful network detection system that employs a stacked nonsymmetric deep auto-encoder coupled with a random forest classifier. Khan et al.[18] suggested TSDL, a two-stage deep learning-based network intrusion detection scheme that mainly utilizes stacked auto-encoders combined with a softmax classifier. Yang et al.[4] proposed SAVAER-DNN, a network intrusion detection framework based on Supervised Adversarial Variational Auto-Encoder With Regularization technique. With respect to the satisfactory results achieved by the aforementioned deep learning frameworks, they still face

the issue of low detection rates of unknown and infrequent attacks.

This paper presents an effective deep learning-based detection approach for DDoS attack detection. The key advantage of this approach is its ability to detect DDoS based on unlabeled data. It belongs to the semisupervised methods and uses only normal data in training, making it more appropriate than supervised learning techniques. Specifically, this approach incorporates a Gated Recurrent Unit (GRU) based on the attention mechanism (AM) at the encoder and decoder sides of the autoencoder (AE) model. This enables enhancing the AE model to describe temporal dependencies via the GRU and selecting relevant features based on the attention mechanism. Furthermore, the local outlier factor (LOF) anomaly detection scheme is applied to extract features from the improved AE model for DDOS attack detection. We assessed the performance of this approach using DDOS publically available datasets.

II. Methodology

The present study uses a data drive approach based on unsupervised learning to deal with Distributed Denial of Service, where the network traffic is analyzed to detect malicious packets (Figure 1). The DoS attack is known to be hard to detect because of its normal contents, meaning that if we analyze one packet in the attack, it looks like any normal packet. However, the source, destination, and especially its high dynamic frequency can be used to identify such attacks from normal traffic. The attack frequency $\mathcal{F}$ represents the number of the same packets from a given source to a given destination. The hackers change the $\mathcal{F}$ value to prevent its detection. This study proposes an effective unsupervised deep learning hybrid framework to deal with this problem. We evaluated different timesteps to model the frequency changes, where data is reshaped as data sequence. We start using a small sequence length and check whether the proposed approach can detect the DDoS attack. The main idea behind the data reshaping is to push the model to focus on a given sequence of data and be able to distinguish between normal and attack traffic. Therefore, the normal traffic is used to train the proposed approach, which is composed of two main parts an unsupervised deep recurrent autoencoder that uses a self-attention mechanism that we called SAE-GRU-A (Figure 2), and an anomaly detection method feed from the outputs of the first part in order to detect the abnormal communications (See Figure 3). The deep recurrent autoencoders are used to learn time dependencies within the data sequence for several sequence lengths (timesteps), where unsupervised learning is performed to extract relevant features and generate a compact continuous representation of the normal traffic (data sequence). In order to improve the features extraction and representation, we used a self-attention mechanism that helped the models focus more on pertinent elements in a given sequence. It was initially used successfully in many areas requiring alignment for text translation, image captioning, and many other areas. The training procedure of the autoencoder is totally unsupervised and consists of the encoder and the decoder. The encoder objective is produced through an encoding function, a compact continuous representation keeping key information (features) from the input via a dimensionality reduction procedure based on gated recurrent units combined with an attention technic. The encoder generates a representative latent space by minimizing an objective function $\mathcal{L}$, which is the cross-entropy that is based on Kullback Leibler (KL) divergence similarity measurement function. KL divergence measures the reconstructed error of the normal traffic input. More specifically, it measures the distance between the probability distributions learned by the deep recurrent autoencoder and the original input.

$$\mathcal{L} = -(y \log(p) + (1-y)\log(1-p)) \tag{1}$$

$$KL(\hat{y}||y) = \sum_{c=1}^{M} \hat{y}_c \log \frac{\hat{y}_c}{y_c} \tag{2}$$

Moreover, the sequences of normal traffic after the training

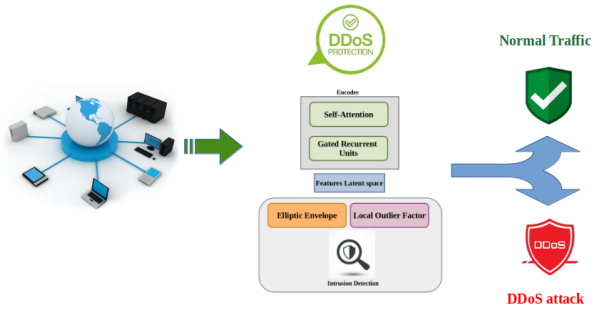

Fig. 1: Proposed approach flowchart.

are represented by a compact latent space that the encoder has learned to encode with a small error. However, the error will be greater if the encoder's input is an abnormal data sequence of a DDoS attack. Here we use the anomaly detection methods to detect this divergence; this is achieved through training in an unsupervised manner with two methods: Local Outlier Factor and Elliptic Envelope. Furthermore, the latent space generated after the training of the autoencoder using normal traffic is used to train the adopted anomaly detection methods.

III. Results and Discussion

A. DDoS attack detection

We evaluate the performance of the proposed strategy using the DDoS Evaluation Dataset (CIC-DDoS2019). This section presents the main experiment's result. The proposed approach is trained using different values of timesteps (from 1 to 42), the sequence length. In order to evaluate the detection performance of the proposed approach using different settings, we adopt the True Positive Rate (TPR), False Positive Rate (FPR), Accuracy, and Precision metrics. The experiments conducted consist of using the deep recurrent autoencoder

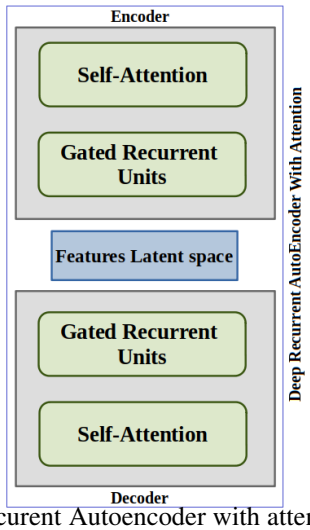

Fig. 2: Deep Recurent Autoencoder with attention mechanism.

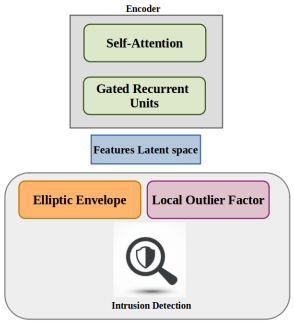

Fig. 3: Intrusion detection using anomaly detection methods.

with attention (SAE-RNN-A) combined with two anomaly detection methods, namely: Elliptic Envelope [19], and Local Outlier Factor [20].

Table I reports the DDoS attack detection performance using the SAE-RNN-A based on Elliptic Envelope [19] anomaly detection approach. The obtained results show that the proposed approach is able to detect the DDoS attack with high Accuracy and Precision of 0.9898 and 0.9947, respectively, for Timesteps equal to 1. The detection performance is improved for all remaining adopted Timesteps except for 6, where the Accuracy and Precision weren't good due to a higher false positive rate.

Figure 4 illustrates the obtained F1-score performance detection for DDoS attack using the SAE-RNN-A based on Elliptic Envelope. It can be seen that the timestep 36 is the highest with a full score of 100% followed by timestep 16 with a score of 99.46% and timestep 1 with a score of 99%.

Figure 5 demonstrates the effectiveness of the proposed DDoS attack detection using the SAE-RNN-A based on the Elliptic Envelope using the AUC metric. It can be seen again that the timesteps 36 and 16 have recorded the highest score.

TABLE I: DDoS attack detection performance using the SAE-RNN-A based on Elliptic Envelope.

Timesteps	TPR	FPR	Accuracy	Precision
1	0.9858	0.0059	0.9898	0.9947
2	0.9589	0	0.9783	1
3	0.9692	0	0.9839	1
4	0.9534	0	0.9754	1
5	0.9362	0	0.9663	1
6	0.9949	0.3871	0.8142	0.7412
9	0.9192	0	0.9571	1
12	0.9658	0	0.9819	1
16	0.9929	0.0038	0.9944	0.9964
18	0.95	0	0.9729	1
24	0.9111	0	0.9529	1
36	1	0	1	1
42	0.95	0	0.975	1

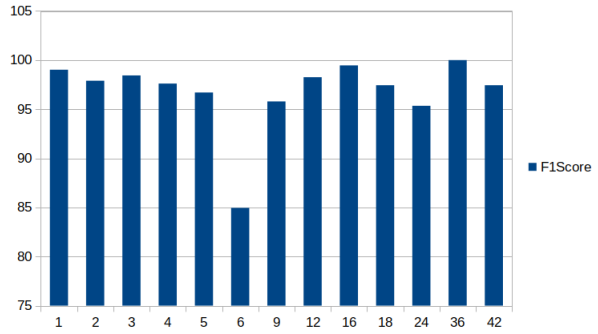

Fig. 4: Recorded F1-score performance detection for DDoS attack using the SAE-RNN-A based on Elliptic Envelope.

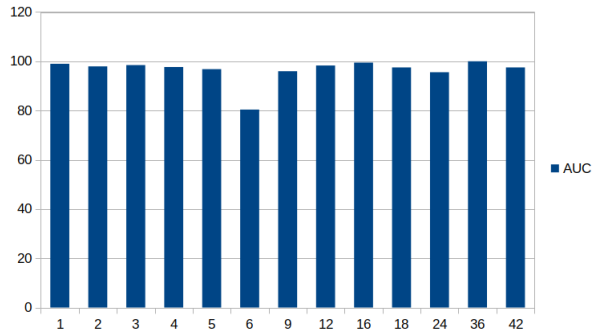

Fig. 5: Recorded AUC for DDoS attack detection using the SAE-RNN-A based on the Elliptic Envelope.

IV. Conclusion

Most of the existing machine learning-driven detection techniques for DDOS cyber-attacks detection are based on supervised learning requiring labeled data in training, which is time-consuming and may not be available in practice. This paper aims to develop flexible and efficient semi-supervised deep learning-driven methodologies to enhance the detection of DDOS attacks. To this end, we amalgamate the advantages of the LOF detector with the feature extraction abilities of deep learning models. Here, we proposed a hybrid AE model that uses GRU with a self-attention mechanism on both the decoder and decoder sides. Detection results based on a publically

TABLE II: DDoS attack detection performance using the SAE-RNN-A based on LOF.

Timesteps	TPR	FPR	Accuracy	Precision
1	0.6316	0.0681	0.7741	0.9113
2	0.8161	0.0085	0.8991	0.9907
3	0.7013	0	0.8436	1
4	0.7814	0.0255	0.8728	0.9715
5	0.8532	0.0036	0.9208	0.9963
6	0.8628	0.0014	0.927	0.9985
9	0.9135	0.0174	0.9459	0.9834
12	0.9737	0.2029	0.8903	0.8428
16	0.9107	0	0.9537	1
18	0.9385	0.0318	0.9521	0.9721
24	0.9278	0	0.9618	1
36	0.975	0.01	0.9818	0.9915
42	0.92	0	0.96	1

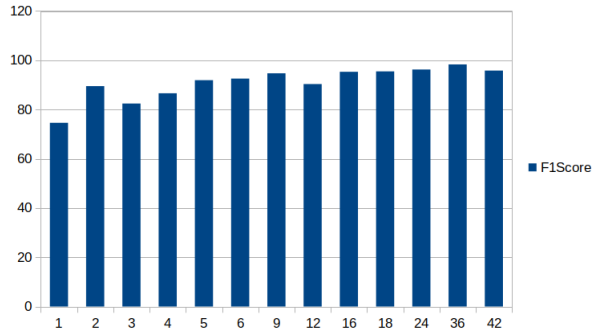

Fig. 6: Recorded F1-SCORE for DDoS attack detection using the SAE-RNN-A based on LOF.

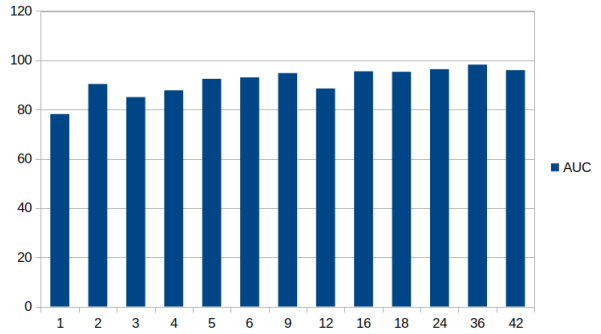

Fig. 7: Recorded AUC DDoS for attack detection using the SAE-RNN-A based on LOF.

available database demonstrated the detection capacity of the proposed AE-GRU-A-driven LOF approach. In future work, we plan to investigate the performance of this approach on other types of attacks.

REFERENCES

[1] J. D. Gadze, A. A. Bamfo-Asante, J. O. Agyemang, H. Nunoo-Mensah, and K. A.-B. Opare, "An investigation into the application of deep learning in the detection and mitigation of ddos attack on sdn controllers," *Technologies*, vol. 9, no. 1, p. 14, 2021.

[2] L. F. Eliyan and R. Di Pietro, "Dos and ddos attacks in software defined networks: A survey of existing solutions and research challenges," *Future Generation Computer Systems*, vol. 122, pp. 149–171, 2021.

[3] J. Li, M. Liu, Z. Xue, X. Fan, and X. He, "Rtvd: A real-time volumetric detection scheme for ddos in the internet of things," *IEEE Access*, vol. 8, pp. 36 191–36 201, 2020.

[4] Y. Yang, K. Zheng, B. Wu, Y. Yang, and X. Wang, "Network intrusion detection based on supervised adversarial variational auto-encoder with regularization," *IEEE Access*, vol. 8, pp. 42 169–42 184, 2020.

[5] F. Harrou, B. Bouyeddou, Y. Sun, and B. Kadri, "A method to detect dos and ddos attacks based on generalized likelihood ratio test," in *2018 International Conference on Applied Smart Systems (ICASS)*. IEEE, 2018, pp. 1–6.

[6] B. Bouyeddou, B. Kadri, F. Harrou, and Y. Sun, "Ddos-attacks detection using an efficient measurement-based statistical mechanism," *Engineering Science and Technology, an International Journal*, vol. 23, no. 4, pp. 870–878, 2020.

[7] F. Chen, Z. Ye, C. Wang, L. Yan, and R. Wang, "A feature selection approach for network intrusion detection based on tree-seed algorithm and k-nearest neighbor," in *2018 IEEE 4th international symposium on wireless systems within the international conferences on intelligent data acquisition and advanced computing systems (IDAACS-SWS)*. IEEE, 2018, pp. 68–72.

[8] I. Ahmad, M. Basheri, M. J. Iqbal, and A. Rahim, "Performance comparison of support vector machine, random forest, and extreme learning machine for intrusion detection," *IEEE access*, vol. 6, pp. 33 789–33 795, 2018.

[9] X. Qu, L. Yang, K. Guo, L. Ma, M. Sun, M. Ke, and M. Li, "A survey on the development of self-organizing maps for unsupervised intrusion detection," *Mobile networks and applications*, vol. 26, no. 2, pp. 808–829, 2021.

[10] F. Harrou, M. M. Hittawe, Y. Sun, and O. Beya, "Malicious attacks detection in crowded areas using deep learning-based approach," *IEEE Instrumentation & Measurement Magazine*, vol. 23, no. 5, pp. 57–62, 2020.

[11] W. Wang, F. Harrou, B. Bouyeddou, S.-M. Senouci, and Y. Sun, "A stacked deep learning approach to cyber-attacks detection in industrial systems: application to power system and gas pipeline systems," *Cluster Computing*, vol. 25, no. 1, pp. 561–578, 2022.

[12] F. Harrou, Y. Sun, A. S. Hering, M. Madakyaru *et al.*, *Statistical process monitoring using advanced data-driven and deep learning approaches: theory and practical applications*. Elsevier, 2020.

[13] T. Ma, F. Wang, J. Cheng, Y. Yu, and X. Chen, "A hybrid spectral clustering and deep neural network ensemble algorithm for intrusion detection in sensor networks," *Sensors*, vol. 16, no. 10, p. 1701, 2016.

[14] C. Yin, Y. Zhu, J. Fei, and X. He, "A deep learning approach for intrusion detection using recurrent neural networks," *Ieee Access*, vol. 5, pp. 21 954–21 961, 2017.

[15] Z. Li, A. L. G. Rios, G. Xu, and L. Trajković, "Machine learning techniques for classifying network anomalies and intrusions," in *2019 IEEE international symposium on circuits and systems (ISCAS)*. IEEE, 2019, pp. 1–5.

[16] M. Lopez-Martin, B. Carro, A. Sanchez-Esguevillas, and J. Lloret, "Conditional variational autoencoder for prediction and feature recovery applied to intrusion detection in iot," *Sensors*, vol. 17, no. 9, p. 1967, 2017.

[17] 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.

[18] F. A. Khan, A. Gumaei, A. Derhab, and A. Hussain, "A novel two-stage deep learning model for efficient network intrusion detection," *IEEE Access*, vol. 7, pp. 30 373–30 385, 2019.

[19] P. J. Rousseeuw and K. V. Driessen, "A fast algorithm for the minimum covariance determinant estimator," *Technometrics*, vol. 41, no. 3, pp. 212–223, 1999.

[20] M. M. Breunig, H.-P. Kriegel, R. T. Ng, and J. Sander, "Lof: identifying density-based local outliers," in *Proceedings of the 2000 ACM SIGMOD international conference on Management of data*, 2000, pp. 93–104.