Graph Embedding for Graph Neural Network in Intrusion Detection System

Dinh-Hau Tran[1], Minho Park[2]
[1]Department of Information Communication Convergence Technology,
Soongsil University, Seoul 156-743, South Korea
[2]School of Electronic Engineering, Soongsil University, Seoul 156-743, South Korea
hautran.0103@gmail.com, mhp@ssu.ac.kr

***Abstract*—Currently, with the rapid expansion of network systems, network security remains a critical concern. Intrusion Detection Systems (IDS) are widely employed to efficiently detect network attacks. Extensive research has focused on applying machine learning models to IDS. Among these models, Graph Neural Network (GNN) is attracting attention as a promising candidate. However, preprocessing network data for the GNN model still poses several challenges. Thus, in this study, we propose an innovative approach to preprocess network flow data before feeding it into the GNN model. Our method involves extracting relevant features from flow data to create nodes and edges for the GNN model. The simulation results indicate that our proposed method significantly enhances the performance of IDS in detecting network attacks.**

***Index Terms*—Intrusion detection system (IDS), graph neural network (GNN), machine learning, flow-based characteristic**

I. Introduction

Today, with the big demand for information technology systems, computers and network devices are rapidly developing [1]. The security of these infrastructures remains a paramount concern in network systems. To address this, intrusion detection systems (IDS) are deployed on most systems to detect attacks on the network system [2]. It plays the role of an effective shield for our system against potential risks. IDS can monitor incoming and outgoing traffic effectively. It exploits various techniques to detect anomalies by comparing signatures from known attacks and scrutinizing unusual behavior in the flow data. However, the attack schemes are becoming increasingly sophisticated and pose significant challenges for IDS. Therefore, the detection of unknown attack patterns is not efficient.

Recently, machine learning (ML) and deep learning (DL) have been used in many fields, such as image processing [3], [4], [5], storage systems [6], and Cybersecurity [7]. Many deep learning techniques, such as convolutional neural networks (CNN), recurrent neural networks (RNN), and traditional multi-layer perceptron (MLP) have also been applied to IDS systems to help improve network monitoring efficiency. Nonetheless, these techniques exhibit limited effectiveness when applied to datasets comprising network flows. This stems from a mismatch between these models and the type of data the IDS is monitoring. Traditional DL models are often only trained with flat data such as vector or grid data. These above DL models are not capable of exploiting complex structures of network flows, while the information in the complex structure is essential in detecting advanced persistent threat (APT) or zero-day attacks. Furthermore, the employed ML techniques focus on analyzing individual network flows, neglecting their inter-dependencies, such as [8] [9].

Among various research techniques in ML, the graph neural network (GNN) models are considered the most suitable for conducting research based on traffic data. GNN is a subclass of Deep Learning techniques designed to operate on graph-structured data including 'nodes' and the relationship between them called 'edges'. Based on the collection of 'nodes' and 'edges', a graph can easily describe the relationship in telephone, social networks, molecules, and so on. Similarly, we can recognize that the traffic flows in computer networks naturally exhibit a graph structure. Therefore, GNN is a suitable model for the flow data of the IDS. In addition, the message-passing mechanism in GNN allows nodes to aggregate and learn information about its neighborhood. It helps GNN to effectively leverage the structure and topology of networks. With the above advantages, GNN can help improve the ability to detect potential attacks within network flows without missing important information about those relationships.

However, GNN-based IDS still have not achieved the desired level of reliability and stability, as the input data for the model has not been optimized before training. To the best of our knowledge, all current research on GNN models for IDS just focuses on representing flow data as only nodes or edges. For instance, [2] presented flow data as a graph format, where network traffic flows are mapped to the graph edges and the endpoints as the nodes in the graph. On the other hand, the paper [10] proposed a method that allows to represent network flows as a graph, where a node includes the flow features as a (IP_src, IP_dst, port, protocol, request, response)-tuple. In another approach way, in paper [11], the authors introduced a heterogeneous graph built based on network flows, and 3 nodes are created for each flow: the source host node, the flow node, and the destination host node. We recognized that the above studies attempted to design graph structures, which are similar to the network topology. However, it has not analyzed the relationship between the elements in the network. Therefore, in this paper, we proposed a method called network flow-based graph convolutional network (FGCN) to design a graph from network flows. In our proposed model, the features of

flow are represented as a node. Simultaneously, to obtain the edges, the relationships between nodes are exploited from the source IPs in the flows. This helps IDS achieve higher accuracy and reduce false alarm rates because the GNN architecture is exploited effectively based on the major elements as nodes and edges. In addition, our proposed model can apply to the whole network with many endpoints instead of only two endpoints in [12]. To improve the detection in [13], our proposed model is developed generally for many types of attack. We implemented the model on two data-sets as CIC-IDS2017 and UNSW-NB15. The experimental results show that the accuracy of the proposed model achieved 94% for CIC-IDS2017 and 96.4% for UNSW-NB15 data-sets.

We summarize our contributions as follows:

- We propose a new approach to represent the network flow data into nodes and edges in graph data.
- We propose the FGCN model, which is a novel GNN-based model, and improves effectively IDS in attack detection.
- The proposed model is used to apply on two standard data-sets and we supplied the simulation results for proofing the effectiveness of the proposed model.

The remainder of this paper is organized as follows. Section II presented our proposed method. The simulation results are discussed in section III. Finally, we supplied conclusions in section IV.

II. Proposed Method

In previous studies, the authors only used nodes or edges to present the features from network traffic flows. Because the GNN model exploits both nodes and edges in graph data, in this paper, we propose a method to extract the features of nodes and edges from the traffic flows. Our proposed model is presented in Fig. 1.

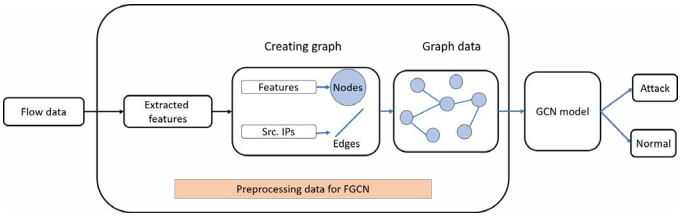

Fig. 1: Diagram of the proposed model.

In the model, the flow data is the message transferred between two computers in the network system. It is specified by the features as Source IP, Source Port, Destination IP, Destination Port, Protocol, ... Among them, we chose and removed the features as Flow ID, Timestamp, and Flag features. Then, the remainder features are kept to create the nodes and edges.

We chose the above features to remove because these features are just statistical information and do not consist of the characteristics of the traffic data. This helps our proposed method can focus on the meaningful information in the flow data and avoid distortions. By preserving almost full features for the training process, the proposed model can obtain higher accuracy compared with the model, which uses a few features from flow data.

Next, we implemented a graph with the extracted features as a node. To create an edge in the graph, we are based on the source IP in each flow data. The flows, which have the same source IP, are linked to create an edge between them. The output of the pre-processing data is the graph data and is used to feed the GNN model.

GNN model has many types and in this paper, we used the GCN model in Fig. 2. Messages Passing layer is always the most important layer in any GNN model. It is a mechanism that allows a node to learn the features of its neighbors, and this is how the GNN model effectively exploits the latent relationship in the network. In the GCN model, this layer is performed by the convolution operation. This process is detailed in the formula (1) and (2), where $h(x)$ is the hidden feature of node x in layer (l), and $m(x)$ is the sum of features from neighboring nodes (node u) of node x in layer $(l-1)$. We applied layers of convolutions and the ReLU activation function to generate the latent representation for each node.

$$m_x^{(l)} = \text{AGGREGATE}^{(l)}(\{h_u^{(l-1)} : u \epsilon \mathbb{N}_{(x)}\}), \quad (1)$$

$$h_x^{(l)} = \text{CONCAT}^{(l)}(h_x^{(l-1)}, m_x^l). \quad (2)$$

The GCN model includes one input layer, h hidden layers, and one output layer. In the input layer, the size of the data is matched with the graph data from the preprocessing. In each hidden layer, we have p perceptrons. The output layer is dependent on the data-sets (this is explained in section III).

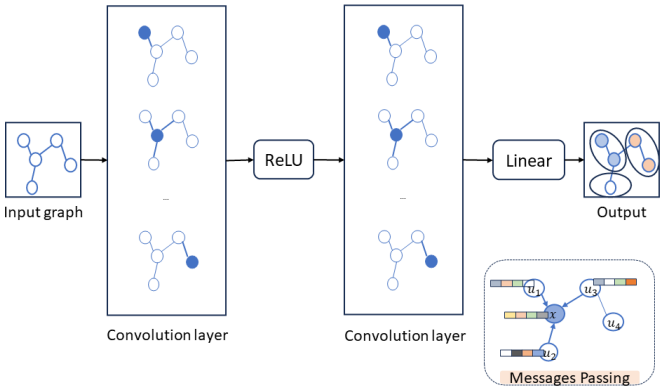

Fig. 2: GCN model

The output of our model is the number of classifications. Depending on the data-sets, the output is the binary vector or boolean type. If data-sets have multi-attack types as the label, the output is a binary vector. Otherwise, if data-sets have just only attack or normal, the output is a boolean type.

III. Results and Discussion

We used the model in Fig. 1 to simulate the network intrusion data-sets as the CIC-IDS2017 and UNSW-NB15. The CIC-IDS2017 contains benign and the most up-to-date common attacks, and in each flow includes 84 features. During

the UNSW-NB15 has a hybrid of the real modern normal and the contemporary synthesized attack activities of the network traffic and it contains 49 features in each flow. Both data-sets are split into 60% for the training set, 20% for the validation set, and 20% for the testing set. By training 200 epochs with ZERO-GRAD optimizer, binary cross-entropy loss function to converge, we obtained the results presented in Figs. 3 and 4. We also compared the accuracy of the proposed model with the previous studies.

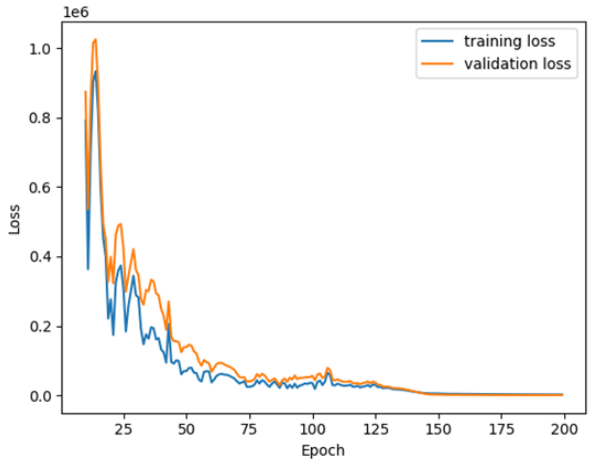

Fig. 3: The loss value of the FGCN model on CIC-IDS2017 data-set

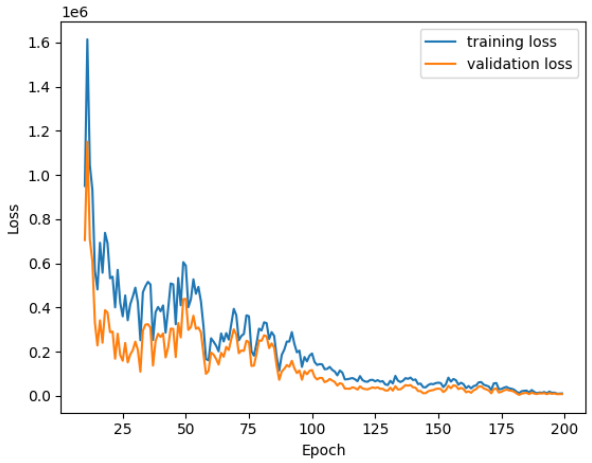

Fig. 4: The loss value of the FGCN model on UNSW-NB15 data-set

The simulation results show that our proposed model achieved high stability when converging and it begins to converge at 125^{th} epoch. The results show that our model gained the peak of 94% and 96.4% of accuracy on CIC-IDS2017 and UNSW-NB15 data-sets, respectively. We can easily recognize that UNSW-NB15 obtained higher accuracy than CIC-IDS2017. Because CIC-IDS2017 needs to detect multi-attack types, the output of CIC-IDS2017 is a binary vector. While UNSW-NB15 just contains normal or attack types. Therefore, our proposed model can improve the accuracy of the short output of UNSW-NB15 compared with the long output of CIC-IDS2017. Finally, our results are better at accuracy and faster in processing time than previous studies.

IV. Conclusion

We have proposed a novel FGCN model using the new approach to represent the network flow data into nodes and edges in graph data. In this model, nodes of the graph represent almost all features of flows, and edges are formed by extracting the relationship among flows based on source IPs. The proposed model achieved high stability and accuracy of 94% and 96.4% for CIC-IDS2017 and UNSW-NB15 data-sets, respectively.

Acknowledgment

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. NRF-2023R1A2C1005461).

References

[1] H. He, X. Sun, H. He, G. Zhao, L. He and J. Ren, "A Novel Multimodal-Sequential Approach Based on Multi-View Features for Network Intrusion Detection," IEEE Access, vol. 7, pp. 183207-183221, 2019.

[2] W. W. Lo, S. Layeghy, M. Sarhan, M. Gallagher and M. Portmann, "E-GraphSAGE: A Graph Neural Network based Intrusion Detection System for IoT," in NOMS 2022-2022 IEEE/IFIP Network Operations and Management Symposium, Budapest, Hungary, 2022, pp. 1-9.

[3] M. -T. Duong and M. -C. Hong, "EBSD-Net: Enhancing Brightness and Suppressing Degradation for Low-light Color Image using Deep Networks," in 2022 IEEE International Conference on Consumer Electronics-Asia (ICCE-Asia), Yeosu, Korea, Republic of, 2022, pp. 1-4.

[4] H. A. Hoang and M. Yoo, "3ONet: 3-D Detector for Occluded Object Under Obstructed Conditions," IEEE Sensors Journal, vol. 23, pp. 18879-18892, 2023.

[5] M. -T. Duong, S. Lee and M. -C. Hong, "DMT-Net: Deep Multiple Networks for Low-Light Image Enhancement Based on Retinex Model," IEEE Access, vol. 11, pp. 132147-132161, 2023.

[6] T. A. Nguyen and J. Lee, "Interference Estimation Using a Recurrent Neural Network Equalizer for Holographic Data Storage Systems," Applied Sciences, vol. 13, pp. 11125, 2023.

[7] C. -N. Nhu, and M. Park, "Two-Phase Deep Learning-Based EDoS Detection System," Applied Sciences, vol. 11, pp. 10249, 2021.

[8] M. Sarhan, S. Layeghy, N. Moustafa, and M. Portmann, "Netflow datasets for machine learning-based network intrusion detection systems," Nov. 2020.

[9] K. Tomar, K. Bisht, K. Joshi and R. Katarya, "Cyber Attack Detection in IoT using Deep Learning Techniques," in 2023 6th International Conference on Information Systems and Computer Networks (ISCON), Mathura, India, 2023, pp. 1-6.

[10] J. Zhao, X. Liu, Q. Yan, B. Li, M. Shao and H. Peng, "Multi-attributed heterogeneous graph convolutional network for bot detection," Inf. Sci., vol. 537, pp. 380-393, 2020.

[11] D. Pujol-Perich, J. Suarez-Varela, A. Cabellos-Aparicio and P. Barlet-Ros, "Unveiling the potential of graph neural networks for robust intrusion detection," ACM SIGMETRICS Perform. Eval. Rev., vol. 49, pp. 111-117, 2022.

[12] B. Pang, et al., "CGNN: Traffic classification with graph neural network," arXiv preprint arXiv:2110.09726, 2021.

[13] J. Zhou, Z. Xu, A. M. Rush, M. Yu, "Automating botnet detection with graph neural networks," arXiv preprint arXiv:2003.06344, 2020.