A Efficient Approach for Intrusion Detection and Prevention in Sensor Networks

Sai Teja Gudupati
Electrical and Computer Engineering
University of Waterloo
Waterloo, Canada
stgudipa@uwaterloo.ca

Shadman Raihan
Electrical and Computer Engineering
University of Waterloo
Waterloo, Canada
s2raihan@uwaterloo.ca

Somesh Kumar Gupta
Electrical and Computer Engineering
University of Waterloo
Waterloo, Canada
sk2gupta@uwaterloo.ca

I. Literature Review

The Internet of Things (IoT) has seen massive growth in the last two decades. Since the number of connected devices is increasing rapidly, the reliability of the network has become a vital issue. Denial of Service (DoS) attack is regarded as one of the main threats in online service. DoS attacks system infrastructure by making the network resource unavailable to its intended users or flooding the targeted business' network bandwidth. Intrusion Detection Systems (IDSs) analyze the computer network traffic logs to identify these malicious attacks.

KDD CUP 99 [1] is one of the most used data set for benchmarking IDSs. Machine learning methods [2] using this data set have achieved 98.94% in recall and 97.89% in precision while keeping the false alarm rate at 0.92% in case of binary classification. However, only a few IDS solutions are using anomaly detection approaches and experts think that this technology is not mature enough. Mahbod.et.al. [1] pointed out the deficiency in the KDD data set with a novel solution to fix it. The data set contains simulated attacks which falls one of the following four types.

1) **Denial of Service Attack (DoS):** This type of attack is initiated by making computing resources busy which denies users from accessing a service.
2) **User to Root Attack (U2R):** In this class of attack, the attacker gets the access of a legitimate account by sniffing password or social engineering. The attacker then uses the root access to exploit the vulnerability.
3) **Remote to Local Attack (R2L):** This happens when the attacker uses some vulnerability to gain the local access of that machine.
4) **Probing Attack:** The attacker tries to gain information about the victim's computer network avoiding its security controls.

Original KDD data set includes redundant samples in the train set which makes the classifier biased towards more frequent records. This bias was preventing the algorithm from learning more harmful but infrequent records like User to Root Attack (U2R). To solve these problems, the authors proposed a new data set named NSL-KDD which got rid of this bias by discarding redundant records and duplicate data in the train set and test sets respectively. The new NSL-KDD data set is still vulnerable to some of the problems discussed by McHugh [3]. However, it is still used as benchmarking data for various intrusion detection methods.

Jiyeon.et.al. [4] proposed a Convolutional Neural Network (CNN) based system for detecting DoS attacks. The performance of the proposed CNN model was evaluated against a recurrent neural network (RNN) model. Here, they have generated two types of image data set from the KDD dataset. Grayscale set of images has one channel while the RGB set has 3 colour channels. For evaluation, 18 scenarios have been created considering various hyperparameters including image size, kernel size and the number of convolutional layers. In this approach, they have implemented both binary and multi-class classification. But for multi-class classification, they have only considered 3 types of DoS attacks. One possible reason for this kind of choice can be a small sample size of other types of attacks since the performance of CNN based algorithms increases with the increase of sample size. RGB images outperformed Grayscale images in terms of accuracy in the case of both binary and multi-class classifications. They also showed the performance of the model increased when they used more than one convolutional layer in multiclass classification. Hyperparameter tuning was done to find out the best possible model. CNN and RNN model both showed 99% accuracy for binary classification. For multiclass classification, CNN achieved 99% accuracy while RNN showed 100%, 80% and 85% accuracy for Smurf detection, Neptune and Benign respectively.

Anomaly-based detection mechanism for DoS attacks has been classified into statistical and machine learning-based approaches in the study of [5]. In the statistical approach, the normal traffic profile is calculated using the mean and standard deviation of normal traffic [6]. The benefit of the statistical approach is less detection time and low computational cost. In machine learning, classification algorithms and ensemble learning have been used for learning the traffic. Amma.et.al [7] used the Class center-based triangle area vector (CCTAV) method to calculate the class center of the dataset and feature extraction. For each sample data, the triangle area vector (TAV) was generated. A normal traffic profile was created using the mean and standard deviation of Mahalanobis distance (MahD) of all normal traffic. A statistical approach based on MahD was used to identify DoS attacks. In this paper, the authors showed that CCTAV has a computational complexity of $O((e^2)^n)$ which is less compared to existing methods.

Data preprocessing plays a vital role in determining the precision and accuracy of a machine learning model. Nerijus.et.al [8] analyzed the effect of data pre-processing on detection accuracy with the NSL-KDD data set. Suleman.et.al [9] found out some of the potential features of DoS attack by entropy calculation and granulation computing. In this approach, the weight of each feature was calculated using the entropy calculation.

Su.et.al. [10] used BLSTM (Bidirectional Long Short-term memory) and attention mechanism for intrusion detection. The key features for network traffic classification have been obtained by the attention mechanism which screens the network flow vector generated by the BLSTM model. They also employed multiple layers of the convolutional layer to extract the local features of traffic data. This deep neural network does not require any feature engineering skills and can automatically extract the key features of the hierarchy. This algorithm has five components i.e input layer, multiple Layers, BSLTM layer, attention layer and output layer.

The performance of a sensor network largely depends on the computational efficiency of the sensor node. Now, a large number of smart sensor nodes are using artificial intelligence. So, it is of paramount importance that intelligent tools should be used in an energy-efficient way in the wireless sensor network. Various types of machine learning techniques are being introduced to reduce energy consumption and communication cost in WSNs. Subha.et.al [11] analyzed various neural networks based algorithms such as ART, ARTl, FUZZY ART, IVEBF and EBCS and pointed out their performance in improving the lifetime of a wireless sensor network.

The study by Tao.et.al [12] uses a hybrid approach for detecting DoS Cyberthreats. The hybrid approach consists of ensemble learning in which the first step is to find clusters of different labels in the dataset using Spectral clustering. In the second step, it uses Deep Neural Network(DNN) to do multiclass classification. It then predicts whether traffic is normal or the whether an attack is being attempted by specifying which type of attack is being conducted on the network with an accuracy score. However, the hybrid approach mentioned by [12] uses unsupervised learning in the first step to cluster the labels which seems unreasonable as the DoS Cyber attacks is a Supervised learning problem. The clustering technique used increases the computational complexity with increase in the power consumption by the network. This however, might not find large scale applications for detecting Intrusions in WSNs as we know that power consumption in WSNs are limited.

Wireless sensor network has constraints in terms of processing power and energy consumption. A large network such as BLSTM and CNN is more resource-intensive and is not an ideal choice for intrusion detection such as DoS attacks in a Wireless Sensor Network. Our work aims to deploy efficient and low power consuming machine learning algorithms while improving intrusion detection performance in detecting DoS Cyber attacks.

References

[1] Mahbod Tavallaee, Ebrahim Bagheri, Wei Lu, and Ali A Ghorbani. A detailed analysis of the kdd cup 99 data set. In *2009 IEEE symposium on computational intelligence for security and defense applications*, pages 1–6. IEEE, 2009.

[2] Mei-Ling Shyu, Shu-Ching Chen, Kanoksri Sarinnapakorn, and LiWu Chang. A novel anomaly detection scheme based on principal component classifier. Technical report, MIAMI UNIV CORAL GABLES FL DEPT OF ELECTRICAL AND COMPUTER ENGINEERING, 2003.

[3] John McHugh. Testing intrusion detection systems: a critique of the 1998 and 1999 darpa intrusion detection system evaluations as performed by lincoln laboratory. *ACM Transactions on Information and System Security (TISSEC)*, 3(4):262–294, 2000.

[4] Jiyeon Kim, Jiwon Kim, Hyunjung Kim, Minsun Shim, and Eunjung Choi. Cnn-based network intrusion detection against denial-of-service attacks. *Electronics*, 9(6):916, 2020.

[5] Abdelouahid Derhab and Abdelghani Bouras. Multivariate correlation analysis and geometric linear similarity for real-time intrusion detection systems. *Security and Communication Networks*, 8(7):1193–1212, 2015.

[6] David J Weller-Fahy, Brett J Borghetti, and Angela A Sodemann. A survey of distance and similarity measures used within network intrusion anomaly detection. *IEEE Communications Surveys & Tutorials*, 17(1):70–91, 2014.

[7] NG Amma and S Selvakumar. A statistical class center based triangle area vector method for detection of denial of service attacks. *CLUSTER COMPUTING-THE JOURNAL OF NETWORKS SOFTWARE TOOLS AND APPLICATIONS*, 2020.

[8] Nerijus Paulauskas and Juozas Auskalnis. Analysis of data preprocessing influence on intrusion detection using nsl-kdd dataset. In *2017 open conference of electrical, electronic and information sciences (eStream)*, pages 1–5. IEEE, 2017.

[9] Suleman Khan, Abdullah Gani, Ainuddin Wahid Abdul Wahab, and Prem Kumar Singh. Feature selection of denial-of-service attacks using entropy and granular computing. *Arabian Journal for Science and Engineering*, 43(2):499–508, 2018.

[10] Tongtong Su, Huazhi Sun, Jinqi Zhu, Sheng Wang, and Yabo Li. Bat: Deep learning methods on network intrusion detection using nsl-kdd dataset. *IEEE Access*, 8:29575–29585, 2020.

[11] CP Subha, S Malarkan, and K Vaithinathan. A survey on energy efficient neural network based clustering models in wireless sensor networks. In *2013 International Conference on Emerging Trends in VLSI, Embedded System, Nano Electronics and Telecommunication System (ICEVENT)*, pages 1–6. IEEE, 2013.

[12] Tao Ma, Fen Wang, Jianjun Cheng, Yang Yu, and Xiaoyun Chen. A hybrid spectral clustering and deep neural network ensemble algorithm for intrusion detection in sensor networks. *Sensors*, 16(10):1701, 2016.