International Journal of Advances in Engineering and Management (IJAEM)
Volume 2, Issue 1, pp: 01-05www.ijaem.net ISSN: 2395-5252

A Multi-Domain Base Cyber Intrusion Detection Using Deep Autoencoder Model

Maryam Abdullahi Musa[1], Abdulsalam Ya'u Gital[2], Kabiru Ibrahim Musa[3], Emmanuel Nannin Ramson[4]
[1,2,3,4]*Abubakar Tafawa Balewa University, Bauchi, Nigeria*
Corresponding Author: Emmanuel Ramson Nanin

Date of Submission: 01-12-2023 Date of Acceptance: 02-12-2023

ABSTRACT: TThe importance of the internet across the globe cannot be over-emphasized as such network security is essential to curb future attack occurrence. Cyber-attacks like DDoS and Ransom-ware yielded a lot of damages to connected devices by endangering and accessing them, not withstanding these damages are air marked to be on the rise. To overcome these issues, machine learning has been used in different computing aspects such as cyber–Intrusion Detection. Recently, deep learning, extreme learning and deep extreme learning networks have superseded machine learning in this context due to their iterative hidden layers that can manipulate complex features of cyber intrusion data. However, almost all of these intrusion detection techniques use machine learning algorithms, that need to be trained from scratch. This suggests the lack of exploration of transfer learning approach in this context. Additionally, the few existing transfer learning approach proposed in the literature focuses on single domain task learning for a particular attack, thus can only transfer knowledge from one source domain, how to transfer knowledge from multiple source domains needs further study. This research work attempts to distinguish the performance of two datasets; wind power time series and ISCX2012 using transfer learning approach. Results obtained shows better performances in the use of transfer learning.

KEYWORDS: Machine learning, Autoencoder, Intrusion Detection and Cyber-attack.

I. INTRODUCTION

Cybersecurity is a set of technologies and processes designed to protect computers, networks, programs and data from attack, damage, or unauthorized access. An active attack happens when a network attacker accesses the interface settings and disconnects certain services of IoT devices may be attacked in various ways, including interruption, interventions, and changes in active attacks. describes active attacks, e.g., DoS, middle-hand attacks, Sybil attacks, spoofing, hole attacks, jamming, selective Forwarding, malicious inputs, data tampering, etc. (Tahsien, Karimipou & Spachos, 2020).

DoS attacks are primarily responsible for disabling system services by generating many repetitive demands, as shown in Fig. 1 As a result, the user cannot navigate and connect to the IoT device, making informed decision-making impossible. Furthermore, DoS attacks keep IoT devices turned on all the time, reducing battery life (Haji & Ameen, 2021). A distributed denial of service (DDoS) attack occurs as several attacks are initiated from different IP addresses to generate various requests to hold the server busy. This makes distinguishing between natural and malicious traffic impossible. In recent years, a specific IoT botnet virus known as Mirai has been responsible for initiating disruptive DDoS attacks, causing thousands of IoT computers to malfunction due to interferences (Tahsien et al., 2020)

Data tampering is a severe threat not just to corporations but also to people's lives and property. As a result, companies must take precautions to avoid such assaults and reduce whatever damage they may inflict (Haji & Ameen, 2021).

User to root attack access normal user account, later gain access to the root by exploiting the vulnerabilities of the system. Root to local is one type of computer network attacks in which an intruder sends set of packets to another computer or server over a network where he/she does not have permission to access as a local user. Probing attacks are an invasive method for by passing security measures by observing the physical silicon implementation of a chip. As an invasive attack, one directly accesses the internal wires and connections of a targeted device and extracts sensitive information. Distributed denial of service involves multiple online devices collectively known as botnet, which are used to overwhelm a target website with fake traffic, making severs unavailable to legitimate users. In normal attack, an intrusion detection system identifies whether the network traffic behavior is normal or abnormal.

In recent days, cybersecurity is undergoing massive shifts in technology and its operations in the context of computing, and data science (DS) is driving the change, where machine learning (ML), a core part of "Artificial Intelligence" (AI) can play a vital role to discover the insights from data. Machine learning can significantly change the cybersecurity landscape and data science is leading a new scientific paradigm (Tolle, Tansley, and Hey, (2011).

The continuous development and extensive usage of Internet benefit numerous network users from a quantity of aspects. Meanwhile, network security becomes much more important with wide usage of network. Network security is closely related to computers, networks, programs, various data, and so forth, where the purpose of defense is to prevent unauthorized access and modification (Wu, Wei, & Feng, 2020). However, the growing number of internet-connected systems in finance, E-commerce, and military makes them become targets of network attacks, resulting in large quantity of risk and damage. Essentially, it is necessary to provide effective strategies to detect and defend attacks and maintain network security.

Furthermore, different kinds of attacks are usually required to be processed in different ways. How to identify different kinds of network attacks thus becomes the main challenge in domain of network security to be solved, especially those attacks never seen before(Wu *et al.*, 2020). Fig. 1 depict the taxonomy of IoT attacks

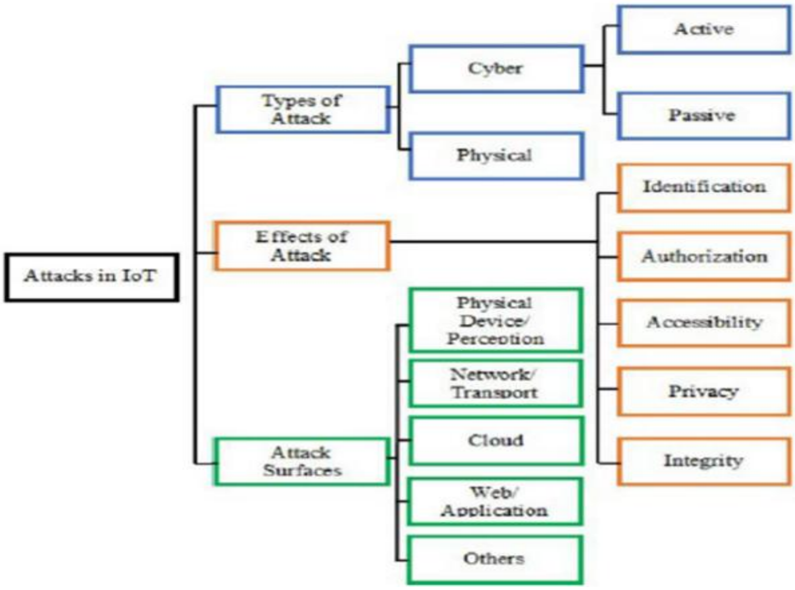

Fig 1: IOT Attacks

II. STATE OF THE ART

Machine learning has been used in different computing aspects such as cyber–Intrusion Detection. Recently, deep learning, extreme learning and deep extreme learning networks have superseded machine learning in this context due to their iterative hidden layers that can manipulate complex features of cyber intrusion data. However, almost all of these intrusion detection techniques use machine learning algorithms, that need to be trained from scratch. This suggests the lack of exploration of transfer learning approach in this context. Additionally, the few existing transfer learning approach proposed in the literature focuses on single domain task learning for a particular attack, thus can only transfer knowledge from one source domain, how to transfer knowledge from multiple source domains needs further study.

The work of (Qureshi *et al.*, 2020) developed an intrusion detection detection system with a wind power time series datasets, results obtained shows promising performances. However, from literature (Subasi *et al.*, 2005) data used not in the problem context gives lower performances than one within the conformity of a problem domain. We therefore use a network intrusion dataset and the wind power datasets used by (Qureshi *et al.*, 2020).

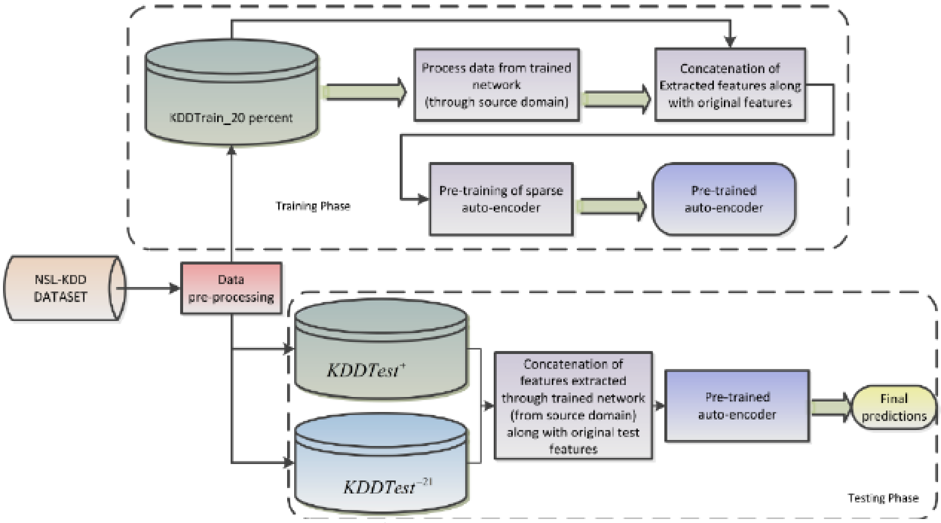
KDDTrain_20 percent
Process data from trained network (through source domain)
Concatenation of Extracted features along with original features
Training Phase
Pre-training of sparse auto-encoder
Pre-trained auto-encoder
NSL-KDD DATASET
Data pre-processing
$KDDTest^+$
$KDDTest^{-21}$
Concatenation of features extracted through trained network (from source domain) along with original test features
Pre-trained auto-encoder
Final predictions
Testing Phase

Fig 2: Block Diagram of Existing Framework (Qureshi *et al.*, 2020).

The study first exploited the concept of self-taught learning to train Deep Neural Networks for reliable network intrusion detection. First, a pre-trained network on regression related task is used to extract features from NSL-KDD dataset. Original features along with extracted features from the pre-trained network are then provided as an input to the sparse auto-encoder. Self-taught learning based extracted features, when concatenated with the original features of NSL-KDD dataset, enhances the performance of the sparse auto-encoder.

III. METHODOLOGY

In this research, a new approach called improved deep transfer auto-encoder is proposed for intelligent detection of cyber-attack using multi-wavelet as activation function for effectively learning useful features hidden in the non-linear cyber-attack data. additionally, correntropy is used to modify the cost function to enhance the reconstruction quality and then pre-train the improved deep auto-encoder using sufficient data from different related source domain of cyber-attacks, and transfer its parameters to the target model. Finally, the improved deep transfer model is fine-tuned by training samples in the target domain to adapt to the characteristics of the rest testing data.

A. Data Description

This research make used of the datasets used in the benchmark paper (Qureshi *et al.*, 2020). As part of our objective was to extend the detection to other multiple cyber-attack. To perform the experiments, we choose a recent NIDS dataset (ISCX2012) that contains multiple attack categories for pretraining as source domain and a target dataset NSL-KDD (containing a new attack type not present in the source dataset). The distribution of attack types in 10% KDD dataset is shown in Table 1.

Table 1: Distribution of attack types in 10% KDD dataset

Attacks	Train dataset	Testing dataset
Normal	97278 (19.48%)	60593 (19.48%)
DOS	391458 (79.24%)	229853 (73.90%)
Probing	4107 (0.83%)	4166 (1.34%)
R2L	1126 (0.22%)	16198 (5.20%)
U2R	52 (0.01%)	228 (0.073%)

This ISCX2012 dataset contains seven days of raw network traffic data containing various attack categories, including normal traffic and four types. Some researchers have noticed that the types of attacks considered in KDD99 are now obsolete. In contrast, the ISCX2012 attack types are more modern and closer to reality. This Distribution of attack types in ISCX2012 dataset is presented in Table 2.

Table 2: Distribution of attack types in ISCX2012 dataset

B. Choice of Metrics

We adopt the same metrics used in the benchmark paper (Qureshi *et al.*, 2020). To evaluate the proposed technique, detection rate, False alarm rate, precision and accuracy are used as evaluation measures. Mathematically, the measures are defined in equations 1 to 5.

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+FN+TN} \quad (1)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (2)$$

$$\text{Recall} = \text{Sensitivity}= \frac{TP}{TP+FN} \quad (3)$$

$$\text{Detection rate}= = \frac{TP}{TP+FN} \quad (4)$$

$$\text{False alarm rate} = \frac{FP}{TN+FP} \quad (5)$$

Where TP (True Positive) = Are the cases when the actual class of the data point was 1 (true) and the predicted is also 1 (true). TN (True Negative) = Are the cases when the actual class of the data point was 0 (false) and the predicted is also 0 (false). FP (False alarm) = Are the cases when the actual class of the data point was 0 (false) and the predicted is 1 (true). FN (False Negative) = Are the cases when the actual class of the data point was 1 (true) and the predicted is 0 (false).

Table 3 shows the parameter setting of the proposed deep auto-encoder, in which hidden layers are pre-trained on source domain task. Whereas, Table 4 illustrates the parameters setting during the training phase of the improved auto-encoder using original as well as extracted features from the source domain ISCX2012 or using only original features of NSL-KDD data transfer learning approach.

IV. EXPERIMENTAL RESULTS

In this subsection, we discussed the experimental results of the proposed approach against the state-of-the-art approaches.

C. Performance of the improve DAE on NSL-KDD pretrained with ISCX2012

In this subsection, we evaluate the performance of the proposed transfer approach when pretrained on ISCX2012 datasets and fine tune on extracted features of NSL-KDD target class. As stated earlier, accuracy, False alarm rate and detection rate, and precision were used to assess and determine the best performing model. Fig. 3 report the performance of the proposed approach when pretrained with a similar and related datasets (ISCX2012) and fine tune on the target NSL-KDD datasets.

From Fig.3, it can be noticed that the proposed model performed well in detecting five (5) different new attacks with high scores. In each case of the evaluation metric from column 2 through column 5 (Accuracy, False alarm, Detection Rate and Precision). The proposed model attains the best score in detecting Normal attack.

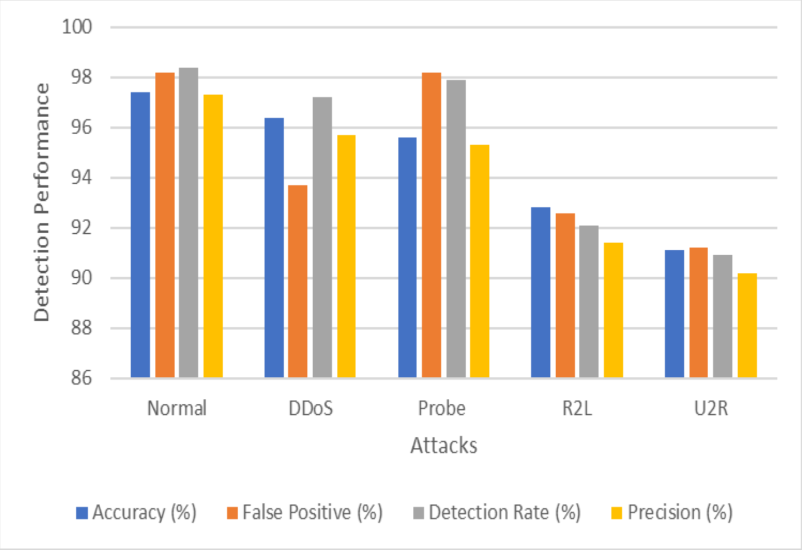

Fig. 3 Performance of the improve DAE on NSL-KDD pretrained with ISCX2012

From Fig.3 The higher the detection performance in percent, the better the performance and the lower the detection performance in percent, the lesser the performance of the model. Clearly, from Fig. 3. The proposed model attained performance score for all the metrics in the case normal of attack with an Accuracy of 97.4 %, False alarm rate of 98.2%, Detection Rate of 98.4% and Precision of 97.3% respectively. However, for the case of probe attack, the proposed model also demonstrates competitive performance in terms of False alarm rate with 98.2%.

Thus, as shown in Fig. 3, similar to traditional machine learning algorithms, the detection rate of the proposed algorithm is higher for Normal, DOS and Probe attack types with sufficient training samples, while the detection rate of U2R and R2L attack types with few samples is relatively low. In terms of False alarm rate, for attack types with more training samples, the False alarm rate is higher. In general, this shows that higher training sample are associated to higher detection performance. However, the transfer of knowledge from one pretrain domain into a new related task has assisted the model in reducing the error of the low training samples thereby improving on the general detection performance of the U2R attack with accuracy of 91.6 % and R2L attack with accuracy of 92.8%. This better when compare to the conventional training approaches which is discussed later in the proceeding sections.

D. Performance of the improve DAE on NSL-KDD pretrained with wind power time series data

For fair evaluation, in this sub section, we evaluate the performance of the proposed approach when pretrain with time series datasets obtain from wind power plant. Fig.4 show the performance of the proposed approach when pretrained with wind power time series data and fine tune to detect five different attacks in the NSL-KDD datasets.

From Fig.4, In each case of the evaluation metric from column 2 through column 5 (Accuracy, False alarm, Detection Rate and Precision). The proposed model attains an improved score in detecting Normal and DDoS attacks respectively. However, similar with the previous analysis. The R2L and U2R where the least detected attacks by the model.

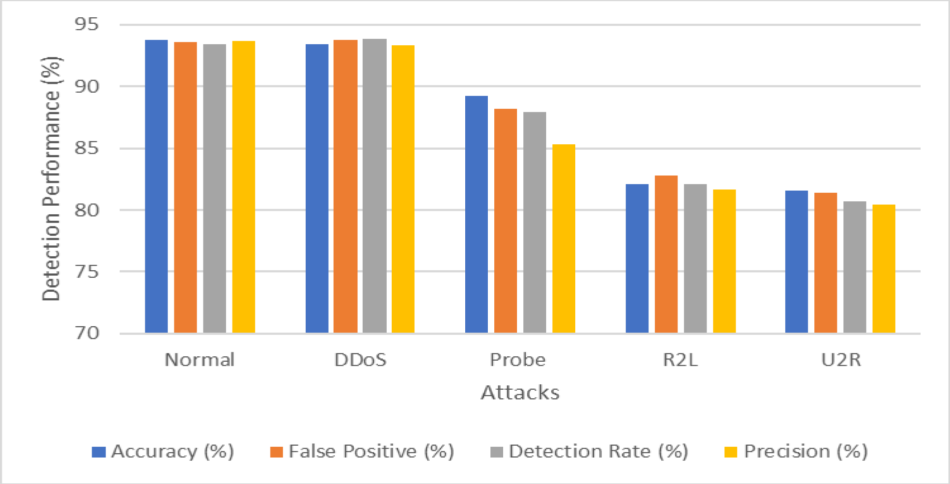

Fig.4 Performance of NSL-KDD pretrained with wind power time series.

From Fig.4. The higher the detection performance in percent, the better the performance and the lower the detection performance in percent, the lesser the performance of the model. Clearly, from Fig.4. The proposed model attained better performance score for all the metrics in the case normal with an Accuracy of 93.8% and Precision of 93.7%, and the DDoS attacks with False alarm rate of 93.6%, Detection Rate of 93.4% and respectively. The reduction in detection rate by the proposed model can attributed with to the use of wind power timeseries datasets as against a more similar and related datasets (ISCX2012).

V. Conclusion

To perform the experiments, we choose a recent NIDS dataset (ISCX2012) that contains multiple attack categories and the wind power time series as used by (Qureshi *et al.*, 2020). From the results obtained better performance score for all the metrics the detection performance was higher when the network is pretrain using the ISCX2012 datasets than when pretrain with wind power datasets.As seen earlier in the literature data that have high similarity with the samples in the target domain is conducive in solving the negative transfer in transfer learning thereby making transfer of knowledge from source domain with better accuracy. This is consistence with our analysis.

Acknowledgement

We wish to thank research our supervisors Prof. Abdulsalam Ya'u Gital, K.I Ibrahim and Prof. Haruna Chiroma for their immense contribution and academic guidance towards the success of this research work.

Reference

Haji, S. H., & Ameen, S. Y. (2021). Attack and anomaly detection in iot networks using machine learning techniques: A review. *Asian Journal of Research in Computer Science*, 30-46.

Qureshi, A. S., Khan, A., Shamim, N., & Durad, M. H. (2020). Intrusion detection using deep sparse auto-encoder and self-taught learning. *Neural Computing and Applications*, *32*(8), 3135-3147.

Subasi, A., Alkan, A., Koklukaya, E., & Kiymik, M. K. (2005). Wavelet neural network classification of EEG signals by using AR model with MLE preprocessing. *Neural Networks*, *18*(7), 985-997.

Tahsien, S. M., Karimipour, H., & Spachos, P. (2020). Machine learning based solutions for security of Internet of Things (IoT): A survey. *Journal of Network and Computer Applications*, *161*, 102630.

Tolle, K. M., Tansley, D. S. W., & Hey, A. J. (2011). The fourth paradigm: data-intensive scientific discovery [point of view]. *Proceedings of the IEEE*, *99*(8), 1334-1337.

Wu, Y., Wei, D., & Feng, J. (2020). Network attacks detection methods based on deep learning techniques: a survey. *Security and Communication Networks*, *2020*.

Wu, Z., Chen, S., Rincon, D., & Christofides, P. D. (2020). Post cyber-attack state reconstruction for nonlinear processes using machine learning. *Chemical Engineering Research and Design*, *159*, 248-261. doi: 10.1016/j.cherd.2020.04.018