A technical review and comparative analysis of machine learning techniques for intrusion detection systems in MANET

Safaa Laqtib¹, Khalid El Yassini², Moulay Lahcen Hasnaoui³

¹²Informatics and Applications Laboratory (IA), Department of Mathematics and Computer Science, Faculty of Sciences, Moulay Ismail University, Morocco
³ISIC ESTM, L2MI Laboratory, ENSAM, Moulay Ismail University, Morocco

ABSTRACT

Machine learning techniques are being widely used to develop an intrusion detection system (IDS) for detecting and classifying cyber attacks at the network-level and the host-level in a timely and automatic manner. However, Traditional Intrusion Detection Systems (IDS), based on traditional machine learning methods, lacks reliability and accuracy. Instead of the traditional machine learning used in previous researches, we think deep learning has the potential to perform better in extracting features of massive data considering the massive cyber traffic in real life. Generally Mobile Ad Hoc Networks have given the low physical security for mobile devices, because of the properties such as node mobility, lack of centralized management and limited bandwidth. To tackle these security issues, traditional cryptography schemes can not completely safeguard MANETs in terms of novel threats and vulnerabilities, thus by applying Deep learning methods techniques in IDS are capable of adapting the dynamic environments of MANETs and enables the system to make decisions on intrusion while continuing to learn about their mobile environment. An IDS in MANET is a sensing mechanism that monitors nodes and network activities in order to detect malicious actions and malicious attempt performed by Intruders. Recently, multiple deep learning approaches have been proposed to enhance the performance of intrusion detection system. In this paper, we made a systematic comparison of three models, Inception architecture convolutional neural network (Inception-CNN), Bidirectional long short-term memory (BLSTM) and deep belief network (DBN) on the deep learning-based intrusion detection systems, using the NSL-KDD dataset containing information about intrusion and regular network connections, the goal is to provide basic guidance on the choice of deep learning models in MANET.

Keywords:
Attack
BLSTM
DBN
Deep learning
Intrusion detection system IDS
Inception-CNN
MANET

1. INTRODUCTION

A Mobile ad hoc Network (MANET) is generally defined as a network that has many free or autonomous nodes [1], often composed of mobile devices or other mobile nodes that can arrange themselves in various ways and operate without strict top-down network administration. There are many different types of setups that could be called MANET and the potential for this sort of network is still being studied. Inherent vulnerability of Mobile ad hoc Networks introduces new security problems, which mostly address...
the network and data link layer of the protocol stack. Because each packet must be passed through intermediate nodes quickly, that packet has to travel from the source to the destination. Malicious routing attacks may target the routing detection or maintenance process by failing to follow the specifications of the routing protocol [2, 3]. This increases the possibility of attacks such as eavesdropping, spoofing, denial of service, and impersonation. Compared to fixed networks, Mobile ad hoc Network security is taken into account from various points such as availability, privacy, reliability, encryption, authentication, access control, and usage control. Due to the prominent characteristics of Mobile ad hoc Networks, security methods used to secure fixed networks are not feasible for MANET [4]. New threats such as attacks from internal malicious nodes, Byzantine, and wormhole attacks are difficult to defend. An Intrusion Detection System (IDS) is an effective way to identify when an attack occurs in a MANET.

For the above reasons, it is very important to deploy in MANET as a second line of defense an intrusion detection system [5]. Intrusion detection systems (IDS) are a mechanism for monitoring and investigating events occurring in a computer system. An IDS incorporates methods for modeling and discovering abnormal behaviors and complex techniques. They try to determine whether or not the network is going through any malicious activity. This is typically accomplished by gathering data automatically from a variety of systems and network sources and then analyzing the information for potential security issues.

Current intrusion detection and prevention methods, such as firewalls, access control protocols and authentication, have several drawbacks in defending networks and devices from ever more advanced attacks, such as a denial of service [6]. However, most systems based on such techniques are suffering from high false positive and false negative detection rates and lack of continuous adaptation to evolving malicious behaviors. Deep learning, therefore, allows to quickly perform data analysis and visualization, the aim is to enable the detection of device vulnerabilities and flaws by security professionals. Several Deep Learning (DL) approaches have been applied to the issue of intrusion detection to increase detection rates and adaptability. Such techniques are often used to establish the attacks existing and detailed knowledge base.

The remainder of this paper is organized as follows. Section 2 is dedicated to discuss Security attacks in MANET. Then, Section 3 describes the intrusion detection system architectures in MANET. Section 4 provides a deep learning models for intrusion detection system in MANET. In section 5 the experiments and resultas. In section 6 we conclude by conclusion.

2. SECURITY ATTACKS IN MANET

Compared to wired infrastructure networks, Mobile ad hoc Networks (MANET) [7] are more vulnerable to attacks. MANET face more security threats than centralized networks due to their dynamic topology and the lack of centralized network administration. In the Mobile ad hoc networks, several characteristics could be used to classify attacks. Examples would include looking at the behavior of the attacks (passive vs. active), the source of the attacks (external vs. internal).

In active attack, the attacker is actively involved in the network operations and tries to change the messages being transmitted. By disrupting the entire network process, the attacker may modify, insert, forge, drop data. The frequency of this attack is high because the whole network can be brought down [8]. They are easy to detect as the network performance degrades significantly. In passive attack, the attacker does not corrupt the shared data but listens to it. They are attempting to gain confidential information and analyze the traffic patterns transmitted. They are difficult to detect because they do not interrupt or modify the information sent or received. It is also possible to classify the attacks into two categories depending on the domain of the attacks, namely external attacks, and internal attacks. Internal attacks are carried out by nodes that are not part of the domain of the network. External attacks are triggered by nodes that are already part of the network. External attacks are more severe than internal attacks [9].

2.1. Denial of service attacks (DoS)

A denial of service attack is an attempt to make a machine or network resource unavailable to its intended users [10, 11], such as to temporarily or indefinitely interrupt or suspend services of a host connected to the Internet. This attack can be launched at different layers: the physical layer, network layer, transport layer.

2.2. Remote to user attacks (R2L)

Occurs when an attacker does not have an account on the victim machine and attempts to gain access by sending packets to a machine over a network in order to generate some vulnerability on that machine that allows him/her to gain local access as a user of that machine.
2.3. User to root attacks (U2R)

This attack occurs when normal system user illegally gains access to either root's or super user's privileges such as Perl, xterm [12].

2.4. Probing

Probing occurs when an attacker scans a network in order to gather information or find known vulnerabilities that allow him/her to hack the entire network. Usually, this method is used in information mining such as saint, port sweep, Mscan, Nmap, etc [13].

3. INTRUSION DETECTION SYSTEM ARCHITECTURES IN MANET

Many intrusion detection systems (IDS) have been developed for MANET to detect various types of attacks. IDS plays an important role in MANET to detect any type of attacks [14, 15]. An IDS is a software system used to analyzes misbehavior and violation of policy, and then generate a report based on it. Basically, intrusion detection system is classified into following three basic categories according to their operational structure.

3.1. The standalone architecture

In this system, to determine intrusion, the intrusion detection system runs independently on the individual node. All decisions made about a particular activity depend solely on information collected at its own node, as there is no collaboration between nodes in the network [16]. Therefore, there is no transfer of information. Even, as no alert information is transferred, a node in the same network does not have any information about the other nodes in the network. Because of its limitations, this model is not efficient, it can be used effectively in a network where all nodes already have an IDS installed. Compared to multi-layered network infrastructure, this system is also suitable for single layer network. Since the information available on any single node is not sufficient to detect intrusions, this system has not been chosen as MANET IDS.

3.2. The distributed and collaborative architecture

In this architecture, the intrusion detection engine is installed on each node in this architecture, which monitors local audit data and detects intrusion. We also participate in the cooperative detection and/or response process by sharing audit information and/or detection results with neighboring nodes to solve the problem. When the intrusion is captured, either a local response (e.g. alerting the local user) or a global response may be issued by an IDS agent. Each node is involved in the method and response of intrusion detection as having an IDS agent running on it [17]. An IDS agent is responsible for detecting and collecting local information and data in order to identify any attack if an attack occurs in the network, as well as taking an independent response. However, when the evidence is non-conclusive, neighboring IDS agents also cooperate in global intrusion detection. This system, like standalone IDS, is also more suitable for flat network systems, not multi-layer systems.

3.3. The hierarchical IDS intrusion detection architecture

Hierarchical IDS system Expand the distributed and cooperative IDS system functions and have been implemented for multi-layer network infrastructures where the network is divided into various small networks known as clusters. Usually, each cluster head has more functionality than other cluster members, such as transmitting data packets to other clusters. We can therefore say that these cluster heads work in some way as central points similar to wired network control devices such as routers, switches or gateways. The multi-layering concept applies to intrusion detection systems where there is a proposal for hierarchical IDS. Each IDS agent runs on a specific member node and is responsible for its node, i.e. monitoring and deciding on intrusions detected locally. A cluster head is responsible for their node locally as well as globally for their cluster, such as monitoring network traffic and announcing a global response when detecting network intrusion [18].

4. DEEP LEARNING MODELS FOR INTRUSION DETECTION SYSTEM IN MANET

Deep learning is a class of Machine learning algorithms that uses multiple layers to progressively extract higher level features from the raw input. The aim is to make machines like computers think and understand how humans think by imitating the grid of the human brain connection, Deep learning architectures such as deep neural networks, deep belief networks, recurrent neural networks and convolutional neural networks have been applied to fields including computer vision, speech recognition, natural language processing, audio recognition, social network filtering, machine translation, bioinformatics,
drug design, medical image analysis, material inspection and board game programs, where they have produced results comparable to and in some cases superior to human experts [19]. Supervised learning algorithm is applied to a dataset that has features and each of those features associated with a label. However, deep learning algorithms come under unsupervised learning algorithms which are applied to a dataset which has many features in order to learn useful properties from the structure of the dataset [20].

The Security applications of deep learning models like Intrusion Detection System (IDS), malware detection, spam-filtering have become essentials in designing tasks for data protection, classification, and prediction. These different types of tasks depending on the intelligence to build a model that usually classifies and discriminates between samples of "benign" and "malign," such as attacks and benign packets [21]. The complexity of attack techniques tools is increased with the rapid increase with the use of Deep learning models. There are lots of popular variants of Deep learning models like CNNs and RNNs. To solve the hardness of training, BLSTM is proposed to alleviate some limitations of the basic RNN, Inception convolutional neural network (CNN) which a variant of basic CNN and deep belief network (DBN). This section gives a brief introduction of the models we have used in our experiments: the inception architecture CNN, BLSTM and DBN [22].

4.1. The inception architecture CNN

Szegedy et al [23] suggest the Inception architecture CNN to solve the problem of a large number of parameters and speed up the learning of CNN. An Inception network is typically a network consisting of modules of the above type stacked on each other, with occasional max-pooling layers with phase two to halve grid resolution. It seemed useful to start using Inception modules only at higher layers for technical reasons (memory capacity during training) while retaining the lower layers in traditional convolutional fashion.

As we can see in Figure 1, the 1x1 convolutions are used to compute reductions before the expensive 3x3 and 5x5 convolutions. Besides being used as reductions, they also include the use of rectified linear activation which makes them dual-purpose [24]. One of the main benefits of this architecture is that it makes it possible to dramatically increase the number of units at each stage without an uncontrolled blow-up in computational complexity. Another practically useful aspect of this design is that it is in line with the principle that visual information should be processed on different scales and then aggregated so that the next stage can simultaneously abstract features from different scales. The improved use of computational resources allows for increasing both the width of each stage as well as the number of stages without getting into computational difficulties. Another way to utilize the inception architecture is to create slightly inferior, but computationally cheaper versions of it.

![Inception architecture CNN](image)

**Figure 1. Architecture of the inception module with dimensions reductions [25]**

4.2. Bidirectional long short-term memory (BLSTM)

Instead of running an RNN only in the forward mode starting from the first symbol, we start another one from the last symbol running from back to front. Bidirectional recurrent neural networks introduce a hidden layer to more robust processes by passing information in a reverse direction. Figure 2 illustrates the architecture of a bidirectional recurrent neural network. In fact, this is not too dissimilar to the forward and backward recursion we encountered above. The main distinction is that in the previous case these equations had a specific statistical meaning. Now they are devoid of such easily accessible interpretation and
we can just treat them as generic functions. This transition epitomizes many of the principles guiding the design of modern deep networks: first, use the type of functional dependencies of classical statistical models, and then use the models in a generic form [26].

![Figure 2. Architecture of BLSTM [27]](image)

### 4.3. Deep belief network (DBN)

To overcome the overfitting problem in Multi-Layer Perceptron (MLP), we can set up a DBN, do unsupervised pretraining to get a decent set of feature representations for the inputs, then finetune on the training set to actually get predictions from the network. While weights of an MLP are initialized randomly, a DBN uses a greedy layer-by-layer pretraining algorithm to initialize the network weights through probabilistic generative models composed of a visible layer and multiple layers of stochastic, latent variables, which are called hidden units or feature detectors. Restricted Boltzmann Machines (RBM) in the DBN are stacked, forming an undirected probabilistic graphical model similar to Markov Random Fields (MRF): the two layers are composed of visible neurons and then hidden neurons. The top two layers in a stacked RBM have undirected, symmetric connections between them and form an associative memory, whereas lower layers receive top-down, directed connections from the layer above. A hybrid model is established by stacking up RBMs, as illustrated in Figure 3. The top two layers form the RBM and the lower layers form a directed belief net [28]. This hybrid model is called a deep belief network (DBN). The deep-belief-network is a simple, clean, fast Python implementation of deep belief networks based on binary Restricted Boltzmann Machines (RBM). In our case, it was based on NumPy and TensorFlow libraries to take advantage of GPU computation.

![Figure 3. Hybrid model of the DBN after greedy layer-wise learning. The top two layers form the RBM and the bottom layers form a directed belief network [29]](image)
5. EXPERIMENT AND RESULTS

5.1. Data preprocessing

The methodology discussed in this paper is applied on the entire NSL-KDD dataset. The NSL-KDD dataset was proposed to deal with inherent problems of the KDD Cup 1999 dataset which contain too many redundant records. An example of dataset record is ‘0 tcp ftp_data SF 491 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 2 2 0 0 0 0 1 0 0 1 0 0 1 0 0 1 0 25 0.17 0.03 0.17 0 0 0 0.05 0 normal’. As you can see, data contains some text values also. Pre-processing of original NSL-KDD dataset is necessary to make it a suitable input for learning models. We need to transform the nominal features to numeric values. Only column number 2(Protocol_type), 3(Services), 4(Flag) and 42(Attack or Normal) contains nominal values [30].

5.2. Evaluation metrics

For evaluation purposes, Accuracy (ACC), Precision (P), Recall (R) metrics are used. These metrics are calculated by using four different measures, true positive (TP), true negative (TN), false positive (FP) and false negative (FN). Accuracy is the percentage of the records number classified correctly over total the records. Precision means the percentage of your results which are relevant. On the other hand, recall refers to the percentage of total relevant results correctly classified by your algorithm [31].

True Positive Rate (TPR): also known as Detection Rate (DR) is the percentage of the anomaly records number correctly flagged as anomaly over the total number of anomaly records in (4):

\[
\text{Accuracy} = \frac{TP+TN}{TP+FP+FN+TN}
\]

Precision = \[
\frac{TP}{TP+FP}
\]

Recall = \[
\frac{TP}{TP+FN}
\]

\[
\text{DR} = \frac{TP}{TP+FN}
\]

False Positive Rate (FPR): the percentage of the normal records number wrongly flagged as anomaly is divided by the total number of normal records in (5) [32].

\[
\text{FPR} = \frac{FP}{FP+TN}
\]

5.3. Comparative study of three deep learning models-based intrusion detection system

The research on security issues relating to IDS exists since the birth of computer architectures. In recent days, applying deep learning to IDS is of prime interest among security researchers and specialists. A comparative study of three deep learning models, Inception convolutional neural network (CNN), Bidirectional long short-term memory (BLSTM) and deep belief network (DBN) applied to IDS.

Tables 1-3 illustrate the accuracy, precision, and recall of our three deep learning models in KDD+ and KDD-21, as we can see that the inception CNN and BLSTM exceed the DBN. Figures 4-7 provide a comparison of the experimental results of Tables 1-3. From the results, we could find that the inception architecture CNN got the highest overall accuracy (ACC). Besides, the BLSTM model surpassed the DBN model on both the overall precision and overall recall rate in KDD+ and KDD-21. Although the DBN model performed worse than the other three models on ACC, precision, recall rate in KDD+ and KDD-21, it obviously failed on information on the attack. The proper explanation that BLSTM tried a lot on the whole sequence comprehension and Inception-CNN could extract the key information more quickly. Inception-CNN and BLSTM perform better than DBN. In theory, DBN should be the best model but it is very hard to estimate joint probabilities accurately at the moment.

We found that all three models had good performance on the Normal data and DoS data and Probe. Table 4 shows the DR of KDDTest+; Table 5 shows the FPR of KDDTest+; Table 6 shows the DR of KDDTest-21; Table 7 shows the FPR of KDDTest-21. Results in Tables 4-7 compare DR and FPR of KDDTest+ and KDDTest-21 we can conclude that the Inception-CNN and BLSTM provide better results in DR and FPR compared to DBN, we can find that results of R2L and U2R are relatively quite small for all models, which might because of the insufficiency of their records in the dataset. However, the models still can detect some of them. Bidirectional long short-term memory (BLSTM) and Inception-CNN, are used to detect anomalies in sequence. The results show that the Inception-CNN and BLSTM showed superiority over the other DBN model. We tie that to the fact that Inception-CNN and BLSTM has the ability to define normal behavior from large datasets and can be used to detect a new unseen threat. It can be concluded that if one only wants to classify the network traffic as normal or attack, Inception CNN or BLSTM they will be a better choice.
A technical review and comparative analysis of machine learning techniques... (Safaa Laqtib)
6. CONCLUSION

Recently, Deep learning for intrusion detection system has received much deliberation. In any IDS, audit data samples are analyzed to set detection rules in highly mobile node network to protect against number of novel attacks. The primary advantage of using Deep learning based intrusion detection systems is that it is highly accurate and able to detect or categorize attacks without any environmental influence. Different Deep learning based IDS approaches have their own benefits and disadvantages. Therefore, considering the MANET scenarios, it is important to choose a precise method for implementing IDS.

This paper is motivated by the need to develop good training algorithms for deep architectures-based IDS, since these can be much more representationally efficient than shallow ones such as SVMs and one-hiddenlayer neural nets, which may be proved important for selecting the appropriate methods on bases of the situation in MANET. In this work, the practical problems of existing IDS have been addressed and different Deep Learning models (Inception-CNN, BLSTM and Deep Belief model) are compared to solve these problems. The models have been implemented and tested on NSL-KDD dataset. The reason behind the superiority of Inception-CNN in general that, it’s the ability to define normal behavior from a large dataset and can be used to detect a new unseen threat. This work can be extended in two directions: first, implement other deep models and create hybrid models and voting systems across different models to detect and recognize low false alarm threats. Second, provide existing systems with real-world data for multiple networks in such a way that the model can increase its accuracy by adapting the definition of normal activities through different un-calibrated datasets.

REFERENCES

[1] E. Amiri, H. Keshavarz, H. Heidari, E. Mohamadi, and H. Moradzadeh, “Intrusion Detection Systems in MANET: A Review,” Procedia - Social and Behavioral Sciences, vol. 129, pp. 453–459, 2014.
[2] S. Laqtib, K. E. Yassini, M. Houmer, M. D. E. Ouadghiri, and M. L. Hasnaoui, “Impact of mobility models on Optimized Link State Routing Protocol in MANET,” 2016 International Conference on Wireless Networks and Mobile Communications (WINCOM), 2016.
[3] S. Laqtib, K. E. Yassini, and M. L. Hasnaoui, “Performance Evaluation of Multicast Routing Protocols in MANET,” in International Conference on Advanced Intelligent Systems for Sustainable Development, Springer, Cham, 2018.
[4] S. Laqtib, K. E. Yassini, and M. L. Hasnaoui, “Link-state QoS routing protocol under various mobility models,” Indonesian Journal of Electrical Engineering and Computer Science (IJECS), vol. 16, no. 2, pp. 906-916, Nov. 2019.
[5] I. Butun, S. Morgera and R. Sankar, “A Survey of Intrusion Detection Systems in Wireless Sensor Networks,” IEEE Communications Surveys & Tutorials, vol. 16, no. 1, pp. 266-282, 2014.
[6] S. Iqbal, M. L. M. Kiah, B. Dhaghighi, M. Hussain, S. Khan, M. K. Khan, and K.-K. R. Choo, “On cloud security attacks: A taxonomy and intrusion detection and prevention as a service,” Journal of Network and Computer Applications, vol. 74, pp. 98–120, 2016.
[7] A. Fudholi and K. Sopian, “Review on Solar Collector for Agricultural Produce,” International Journal of Power Electronics and Drive Systems (IJPEDS), vol. 9, no. 1, pp. 414-419, Jan. 2018.
[8] T. Laagoubi, M. Bouzi, and M. Benchagra, “MPPT and Power Factor Control for Grid Connected PV Systems with Fuzzy Logic Controllers,” International Journal of Power Electronics and Drive Systems (IJPEDS), vol. 9, no. 1, pp. 105-113, Jan. 2018.
A technical review and comparative analysis of machine learning techniques... (Safaa Laqtib)