Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 4, April : 2023

INVESTIGATING ROLE OF LSTM APPROACH FOR IDS DETECTION

Sahil Sehrawat[1], Dr. Dinesh Singh[2] Computer Science and Engineering Department, DCRUST, Murthal, Sonepat, Haryana

Abstract: *This research has presented the intrusion detection system with its working. There are several existing researches in field of IDS, but the major issue is the accuracy. However various ML approach are used to train the IDS in order to detect the intrusion in future but the accuracy remains low. This paper has considered the existing mechanism of IDS and this review would be capable to set the platform for further enhancement in field of IDS system. Objective of research is to consider existing research in area of IDS and DL. Works is focusing on the accuracy and performance issues in conventional research work. In future Hybrid LSTM model with integration of PSO would be proposed in order to get best solution to improve accuracy to evaluate the accuracy of proposed model developed for IDS model.*

Keywords : Intrusion detection system, LSTM, Accuracy, Deep learning, Performance

1. INTRUSION DETECTION SYSTEM

The term IDS is often used to refer to a system that keeps tabs on data transmissions inside a network in search of malicious behaviour. Additionally, it notifies you once suspicious behavior is uncovered. What we're talking about here is network-testing software. It makes sure the system isn't being used in a risky manner and that no rules have been broken. Multiple parts make up an intrusion detection system. Sensors used to generate security events are one such element. The system has detected an intrusion and is responding. The console is an additional part.

When conducting routine tasks, intrusion detection systems look for indicators of known assaults or abnormalities. Such deviations or abnormalities are sent to the protocol and application layers for verification.

The possibility of using an IDS as a safety measure has been investigated. It is working mainly at the network layer of IoT system. IDS, which has been designed for IoT dependent intelligent environments need to work in stringent situation of less processing capability. It should be capable to provide fast response. This is supposed to provide high-volume data processing.

IDS exists either in the form of hardware or in the form of application of computer program. It observes malicious activity happed in any network or systems. Its contribution in the assurance of data safety is significant. As so, it joins the ranks of cutting-edge tools capable of accurately detecting various forms of network assault. Activities which are related to a network related such as quantity of traffic, internet protocol address, service ports, and protocol are analysed by the systems which are derived on the basis of network.

2. LITERATURE REVIEW

Many researches says of accuracy and security. As Security is paramount in IDS .IDS security is an area of concern in order to safeguard the networks in the system. Therefore, the system's precision may be enhanced. Also there is a need to classify attack on base of their name so the network can adjust according to the attack.

Based on advancements in DBNs and a genetic algorithm, Peisong Li and Ying Zhang [1] introduced intrusion detection architecture. (GA). Iterative evolution is used to generate optimal network architectures of DBN for various assaults, such as low-frequency attacks and others. To aid in intrusion detection, it was recommended to use DBNs, in which the network topology was optimised. When genetic algorithm is used it becomes possible to generate maximum number of hidden layers in a normal manner. In a similar way, neurons of hidden layer are also generated. It minimizes system complications up to the possible greatest extent and provides the speed of detection. It becomes possible that to improve the efficiency of IDS with the help of this method. Rate of detection also rises when this method is implemented.

CHUANLONG YIN [2] submitted the model of identification system and method of its usage on the basis of periodic neural networks. In addition to this, they evaluate the effectiveness of design in dual and multiple class organization. They also evaluate neuron

quantity and influence of various learning speed on precision. Dataset used is NSL-KDD. It has been come out of practical consequences that with the help of RNN-IDS classification model can be modelled in a very suitable way. The efficiency and precision of this classification model is far better in comparison to the usual organization methods of automatic learning in dual and multiple class organization. This design makes the precision of intrusion detection better. It offers latest research method in support of intrusion detection.

Bo Dong, Xue Wang [3] considered those methods which were utilized for the purpose of network traffic classification. It was decided by them that they implement number of methods on free information package and carried out practical in the company of these methods. Out of these practical, they discovered the optimal way of intrusion detection. Presently, deep learning is consistent best because of its predicting capabilities in support of automatic learning. Because of this particular factor, methods of deep learning are already in sectors like structural identification or organization. Information about the state of a network may be gleaned by intrusion detection analysis by keeping an eye on security events. Number of usual methods of automatic learning are already been presented to intrusion detection, but improvement in precision and efficiency becomes compulsory.

Imtiaz Ullah(B) and Qusay H. Mahmoud [4] have obtained a dataset IoTID20 also have used different methods of automatic learning such as SVM, DT, RF and various others. Latest information package of IoTID20 provides a base in support of latest intrusion detection methods development within internet of things networks.

Sara A. Althubiti [5] implemented the system of intrusion detection. For this purpose researcher used the information package of Coburg Intrusion Detection. It addition to this research worker also applied method of a deeply structured learning and LSTM. An accuracy of near about point eighty five was achieved through this work. This accuracy was considered reasonable. In order to fulfill our evaluation purpose their LSTM outputs were compared in the company of most elegant methods. For this purpose they used different type of metrics like authenticity, versaty and some other.

W. Li, P. Yi, Y. Wu, L. Pan, and J. Li [6] have researched in field of latest intrusion detection system. This system was derived on the basis of algorithmic program which are organized by KNN . Mechanism has been developed for wireless sensor network.

In 2016, A. L. Buczak and E. Guven [7] did survey of those methods which are used for the purpose of information extraction and automatic learning. They focused on preventative methods of intrusion detection. The integrated information extraction and automatic learning methods are used for performing detection in case of any intrusion.

In 2016, A. Javaid [8] proposed the method of deep learning. Researches were also carried out by them in support of intrusion detection system.

In 2016, T.A. Tang [9] proposed deep learning approach. The objective of approach is to perform network intrusion detection. The area of research was software defined networking.

In 2012, M. Sheikhan [10] presented intrusion detection system. They used reduced size RNN to perform this operation. The mechanism was based on feature grouping.

In 2009, M. Tavallaee[11] conducted an in-depth study of the KDD CUP 99 dataset.

S.Revathi [12] suggested in-depth examination of the NSL-KDD dataset in 2013. Several forms of ML are used. This was done so that infiltration might be detected.

N. Paulauskas [13] analysed raw data in 2017. They took into account how preparing the data may affect intrusion detection. NSL-KDD was the dataset that was analysed in depth.

In 2017, P. S. Bhattacharjee [14] proposed Intrusion detection system. They make use of NSL-KDD data set. The mechanism used in research was vectorised fitness function. They did research in area of genetic algorithm.

In 2017, R. A. R. Ashfaq [15] proposed semi-supervised learning approach. Research has considered Fuzzy logic. The objective of research is to perform intrusion detection system.

In 2011, J. Martens [16] presented learning recurrent neural networks. They did these using hessian-free optimization techniques.

3. PROBLEM FORMULATION

However there are several researches in field of intrusion detection system. Existing research did research on learning system and trained the IDS data set in order to perform prediction. The limitation of existing research is overall accuracy. After training of data set the accuracy remains below 90% in previous researches. There is need to propose better solution by

introduction LSTM model with additional layers and drop out layers to increase the accuracy during testing.

4. ROLE OF LSTM IN IDS DETECTION

Neutral Network Based Ids

It's crucial since it deals with the NN architecture. Supervised learning is being carried out via neural network models. Using IDS data sets, it is constructing a body of knowledge. It is taking into account the predetermined right response. The networks then fine-tune their ability to correctly identify the right response. These are an effort to improve forecast precision. There are many problems that these NN might help with. Intrusion detection and other forms of network security may be related to these problems. It's possible that functions like sales forecasting and consumer research are intertwined with these problems. Validating data and controlling risks are two additional concerns.

LSTM and Its Training Mechanism

The system will keep the trained network, or "net," for future use. To create a trained network, two LSTM layers were used in the implementation. During the training process, the proposed model employs two LSTM layers and a drop out layer. Seventy percent of the dataset is used for training, while the remaining thirty percent is used for testing. The LSTM-dependent neural network is trained based on the features. Training duration is affected by a number of parameters, one of which is batch size. Accuracy has improved greatly thanks to the efforts of the hidden layers and dropout layers. In order to train an IDS, a dataset must first be obtained, and then characteristics must be chosen from it. Once the training/testing split is determined, a 12-layer LSTM1 layer and a 5-layer LSTM2 layer are implemented. When overfitting becomes a problem, a dropout layer is employed to fix it, followed by a fully connected layer and a softmax layer. Decisions on potential intrusions may be made with the use of a classification operation.

LSTM layer

The LSTM mechanism for DL and feedback connectivity is being investigated in the study as a means of training the network. The IDS attack categorization in the proposed model is handled by LSTM networks, which are equipped with hidden layers, dropout layers, fully connected layers, and classification layers. In order to analyse and predict data from the provided IDS dataset, LSTM method has been applied.

The abnormalities are categorised by the trained model. The IDS was modelled using a pair of LSTMs, which were then connected one after the other. There are a total of 12 and 5 hidden layers, respectively, in each. Although these concealed layers have improved accuracy, they have also introduced the problem of over fitting. When training a neural network model, over fitting occurs. Keeping training going means the model will eventually take on habits and routines that humans perform.

Long-term dependencies are something these layers must understand. It learns not just the contents of sequences, but also the many time steps that make up time series. The interactions it conducts are additive. The layers are used to improve gradient flow over extended periods of time when training.

Dropout layer

By monitoring the validation loss while plotting, overfitting may be detected and corrected. As the model receives more and more practice, it takes on its trainer's quirks. After a while, training stops being as effective on novel data. This information may represent various subsets of the whole population. When the model fits the training data and validation data too well, we say that it has been overfit. When training a neural network model, overfitting occurs. Overfitting may be dealt with in a number of ways. The dropout layer is one solution to the overfitting problem. Such an issue may be dealt with using the dropout layer.

When the training loss stays the same or goes down, the model is overfitting. Regularizes are a set of techniques developed to reduce the impact of overfitting. One of them is dropping out of school. Dropout is effective because it removes a layer's inputs. It might be that the output of the preceding layer is a set of input variables in the data sample. That is to say, the Dropout layer is a subsequent addition to the model. This rule is applicable to the layer above, which received the outputs from the previous layer. The modelling of a large network with different network architecture has an effect on it. Layer-specific dropout rates might be provided as probabilities of setting each input.

The dropout used in the proposed work works by removing the layer's inputs. The suggested model has been updated by adding a dropout layer on top of the existing ones.

Fully connected layer

The input is multiplied by the weight matrix in a fully linked layer. After that, a bias vector is included. The

proposed work introduces a completely Connected Layer (output size) function, which, as its name suggests, returns a fully connected layer while also defining its output size.

Softmax

The activation layer is often used as the last layer in a NN. In lieu of the ReLU, sigmoid, and tanh activation functions, it is used. The output of the previous layer must be converted to the neural network, which is why the Softmax layer is essential. Typically, NN layer provides assistance for softmax to be carried out shortly before the output layer. All of the output layer's nodes' counts should be recorded here.

Classification layer

Classification is often regarded as the most dynamic field of neural networks research and application. Separating large datasets into distinct categories in order to derive a rule relies heavily on classification's presence.

Batch size

The batch size has been considered as a hyperparameter. It defines the counting of samples to perform the task before modifying parameters inside the internal model. Batch is considered a loop. It is iterating multiple samples to perform a prediction operation. Predictions are compared to expected output variables, and error is found in the Batch. The use of an update mechanism helps rectify this error and is thus utilized to enhance the model. In this research, the considered batch size is 512. If the batch size is increased, then training time gets reduced, but if the batch size is reduced, then training time also increases. The batch size is also influencing accuracy because of observation. It is found that if the batch size is increased, then accuracy gets reduced and vice-versa.

Gradient threshold

It is a comma-separated pair that has been provided as an option. It consists of a positive scalar and a Gradient Threshold. According to the Gradient Threshold Algorithm, the gradient is clipped if it is larger than the threshold value. The suggested research makes use of a Gradient Threshold value of 2. This parameter is considered because there is an issue of exploding gradients when significant error gradients accumulate and output in huge modification to neural network model weights during training. It is not stable and not able to learn from previous training data.

Epoch

Several epochs are considered a hyperparameter. These are defining the time count. This time count belongs to the learning algorithm. It represents a time counter for a learning algorithm that works over a complete dataset used for training purpose. Epoch is representing every sample in the dataset used for training have changed to modify parameters in the internal model.

There are 30 epochs at the time of training of the model.

Learning Rate

This is an example of a hyperparameter that may be adjusted. Adjusting the magnitude of model changes in light of predicted error. The model's weights are updated periodically. In complicated operations, the selection of learning rates has been discovered, as the smallest value may lengthen the training process. This may be permanent training. However, if the value is too high, the output during training is accelerated while learning the ideal set of weights. When training neural networks, the learning rate is used if there are positive but relatively tiny values. These positive numbers range from zero to one. Model adaptation problems are being managed by the learning rate.

Testing Phase

The trained model is evaluated for its performance in the testing phase. To ensure the model is reliable, a sample dataset is used. Prediction is carried out using the network model trained with the prior dataset. The model's trustworthiness is shown on the testing side. In the evaluation stage, a supervised trained network is taken into account. The testing dataset is then extracted. Accuracy, f-score, and precision are calculated when applying a trained network to the test data that takes into account novel test values.

Let's do our analysis by first training the model on 70% of the data and then testing it on the remaining 30%.

5. SCOPE OF RESEARCH

The study's findings would be crucial in selecting an IDS prediction mechanism with high precision. When it comes to IDS detection, new studies promise a scalable and adaptable method that takes the training model into account. Since the proposed model employs a large dataset during its training phase, there should be less room for mistake in the final accuracy calculation. To improve IDS detection in the future, similar studies should be able to utilise the same model.

REFERENCES

[1] P. Li and Y. Zhang, "A Novel Intrusion Detection Method for Internet of Things," 2019 Chinese Control And Decision Conference (CCDC), Nanchang, China, 2019, pp. 4761-4765, doi: 10.1109/CCDC.2019.8832753.

[2] CHUANLONG YIN, YUEFEI ZHU, A Deep Learning Approach for IntrusionDetection Using Recurrent Neural Networks, IEEE Access, Received September 5, 2017, accepted October 5, 2017, date of publication October 12, 2017, date of current version November 7, 2017.

[3] B. Dong and X. Wang, "Comparison deep learning method to traditional methods using for network intrusion detection," in Proc. IEEE ICCSN,2016, pp. 581–585

[4] Ullah, Imtiaz, and Qusay H. Mahmoud. "A Scheme for Generating a Dataset for Anomalous Activity Detection in IoT Networks." Canadian Conference on Artificial Intelligence. Springer, Cham, 2020.

[5] Althubiti, Sara & Jones, Eric & Roy, Kaushik. (2018). LSTM for Anomaly-Based Network Intrusion Detection. 1-3. 10.1109/ATNAC.2018.8615300.

[6] W. Li, P. Yi, Y. Wu, L. Pan, and J. Li, ``A new intrusion detection system based on KNN classi_cation algorithm in wireless sensor network,"J. Elect. Computer. Eng., vol. 2014, Jun. 2014, Art. no. 240217.

[7] A. L. Buczak and E. Guven, ''A survey of data mining and machine learning methods for cyber security intrusion detection,''IEEE Commun.Surveys Tuts., vol. 18, no. 2, pp. 1153–1176, 2nd Quart., 2016.

[8] A. Javaid, Q. Niyaz, W. Sun, and M. Alam, ''A deep learning approach fornetwork intrusion detection system,'' presented at the 9th EAI Int. Conf.Bio-inspired Inf. Commun. Technol. (BIONETICS), New York, NY, USA,May 2016, pp. 21–26.

[9] T. A. Tang, L. Mhamdi, D. McLernon, S. A. R. Zaidi, and M. Ghogho,''Deep learning approach for network intrusion detection in soft-ware defined networking,'' inProc. Int. Conf. Wireless Netw. MobileCommun. (WINCOM), Oct. 2016, pp. 258–263.

[10] M. Sheikhan, Z. Jadidi, and A. Farrokhi, ''Intrusion detection using reduced-size RNN based on feature grouping,''Neural Comput. Appl.,vol. 21, no. 6, pp. 1185–1190, Sep. 2012.

[11] M. Tavallaee, E. Bagheri, W. Lu, and A. A. A. Ghorbani, ''A detailed analysis of the KDD CUP 99 data set,'' inProc. IEEE Symp. Comput. Intell.Secur. Defense Appl., Jul. 2009, pp. 1–6.

[12] S. Revathi and A. Malathi, ''A detailed analysis on NSL-KDD dataset using various machine learning techniques for intrusion detection, ''Int. J. Eng.Res. Technol., vol. 2, pp. 1848–1853, Dec. 2013.

[13] N. Paulauskas and J. Auskalnis, ''Analysis of data pre-processing influence on intrusion detection using NSL-KDD dataset,'' inProc. Open Conf.Elect., Electron. Inf. Sci. (eStream), Apr. 2017, pp. 1–5.

[14]P. S. Bhattacharjee, A. K. M. Fujail, and S. A. Begum, ''Intrusion detection system for NSL-KDD data set using vectorised fitness function in genetic algorithm,''Adv. Comput. Sci. Technol., vol. 10, no. 2, pp. 235–246, 2017.

[15] R. A. R. Ashfaq, X.-Z. Wang, J. Z. Huang, H. Abbas, and Y.-L. He,''Fuzziness based semi-supervised learning approach for intrusion detec-tion system,''Inf. Sci., vol. 378, pp. 484–497, Feb. 2017.

[16] J. Martens and I. Sutskever, ''Learning recurrent neural networks with hessian-free optimization,'' presented at the 28th Int. Conf. Int. Conf.Mach. Learn., Bellevue, WA, USA, Jul. 2011, pp. 1033–1040.

[17] https://datasilk.com/intrusion-detection-prevention