An Execution of Intrusion Detection Using Boltzmann Machine and Its Applications

Shailendra Singh Gaur[1], C. M. Sharma[2], Varsha Sharma[3]

Bhagwan Parshuram Institute of Technology, GGSIPU, Delhi
shailendrasinghgaur@bpitindia.com, cmsharma@bpitindia.com, varshasharma@bpitindia.com

Abstract: In today's era wireless sensor networks are very challenging to reduce the attack and risk of getting affected. Several methods for image compression are used to solve the WSN problems like RBM, DRBM, CRBM and many more. Most of the image compression algorithms are random content, change in image and results in low quality of images after deducting. Multilayer random Boltzmann machine learning network is implemented in this paper to solve the intrusion detection in network traffic to improve the accuracy of the algorithm, which will help us to understand the different types of attacks.

Keywords: *WSN, Restricted Boltzmann Machine Learning Algorithm, Intrusions Detection*

1. INTRODUCTION

Wireless Sensor Networks monitor sound, motion, pressure and temperature that are physical or environmental conditions. Advancement in the technology has led to loss of important information and exposure of sensitive data. It also required to create a system which can reduce attack and detect risk on security. Wireless Sensor Networks (WSNs) are self- configured and infrastructure-less wireless network to monitor the physical or environmental conditions such as sound, temperature, pressure, motion, and vibration etc. With the advancement of technology in recent years, the various attacks over these data have increased rapidly and WSN has gained attention too. This often leads to loss of important information and exposure of sensitive data. Thus, a system is required to be placed which can automate the process of attack and their detection and thus reduce the risk on security teams.

2. OBJECTIVES

- The use of AI along with Data science and Machine learning prioritizes the security alerts and automate the responses significantly to reduce the stress placed on security teams.
- AI and machine learning is used widely in data science application and skill sets to identify the behavior pattern that can't be detect by the preset rules.
- ML algorithm is allowed to develop a defense response after analyzing and processing the previous cyber-attack data.
- Encryption techniques are needed to operate on many different types of data, both user and machine generated inputs with different analysis tools and big data storage formats.

3. MATERIALS AND METHODOLOGY

In the absence of a Standard Dataset for Attack detection in WSNs, the appropriate dataset found for the purpose was NSL- KDD dataset. This data set contains the records of the internet traffic seen by a simple intrusion detection network and is the ghosts of the traffic encountered by a real IDS and just the traces of its existence remained. The dataset consists of 125973 rows and 43 columns. No missing values were found in any of the columns. Three features were object datatype which were unsuitable for machine learning algorithms, hence would require pre-processing steps. All other features were either integer or float values which were directly used. Type was selected as the target ones. All other features present were used for determination of type of attack.

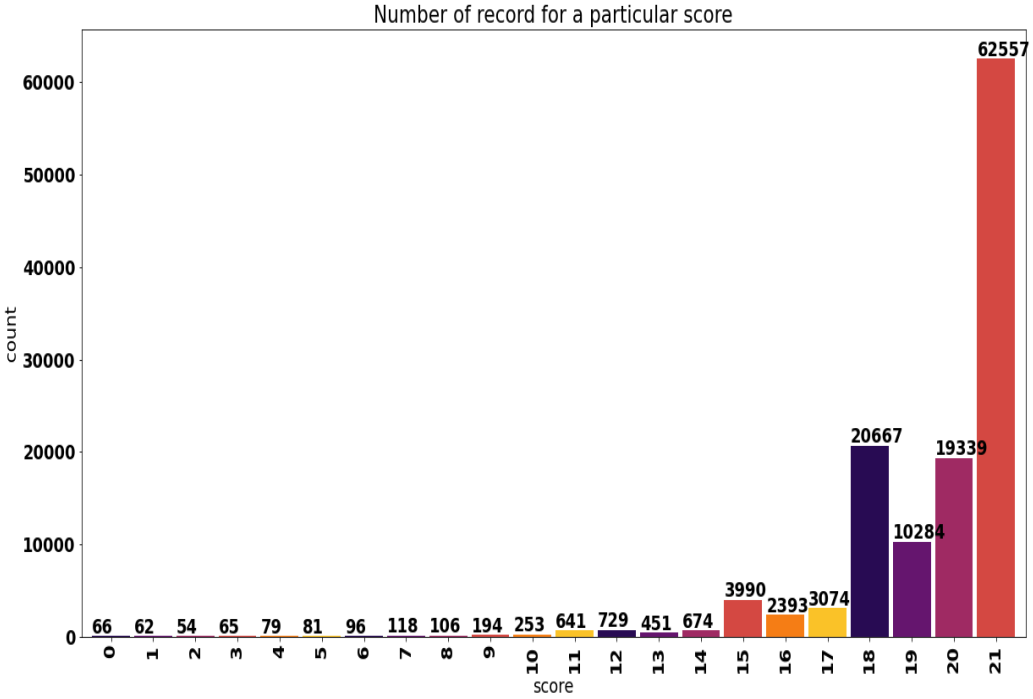

Fig.1. Score Distribution of Attacks for a particular record

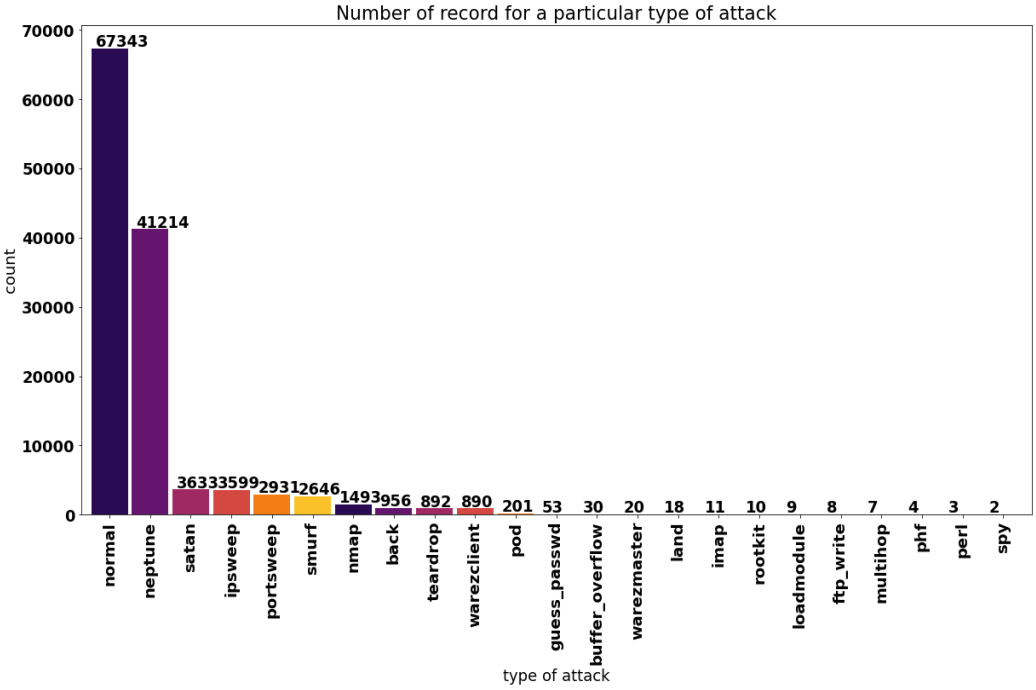

Fig. 2. Attack Frequency for a particular type of attacks

The dataset selected was first used for pre-processing. All the features in object format were one-hot encoded. Therefore, Restricted Boltzmann Machine Algorithm is applied to a processed dataset. The various WSN image analysis techniques are Auto encoder, RBM, NMF, LSNMF and PNMF based image compression techniques. Restricted Boltzmann machines (RBMs) are unsupervised machine learning algorithms. These algorithms are used to represent the internal representation of data and also sample the output of visible units and hidden layer input or vice versa. RBM is used for classification and generation.

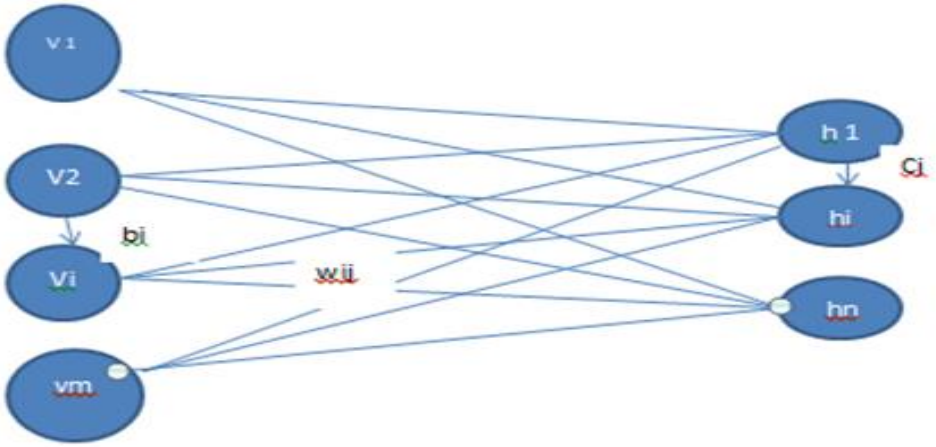

Fig.3. Restricted Boltzmann Machine Architecture

Here i is visible in nit and j is hidden unit and wij belongs to W where W belongs to Rm x n set of weights from hidden to visible units. The bias unit is represented by bi belongs to b. And the hidden unit is cj belongs to c.

This algorithm used hidden h and joint vector v distribution which is proportional to exponential of negative energy of configuration.

Before compression of image, we need to build a matrix factorization algorithm which converts the user matrix into 2 lower D matrices. Where H is the m users x k latent factor and W is latent factor into n user item matrix. Resultant is calculated by multiplying H x W that is R. K represents the model capacity and increases in K factor represents the improvement in prediction chances and if k is too high, the model is considered to be fit. In RBM all parameters are set to be 10 epochs with no predefined weight.

We pass the input data from each of the visible node to the hidden layer.

Steps for Restricted Boltzmann Machine Architecture Algorithm:

• We multiply the input data by the weight assigned to the hidden layer, add the bias term and applying an activation function like sigmoid or softmax activation function.
• Forward propagation gives us probability of output for a given weight w ,this gives P(a|x) for weights w.
• During back propagation we reconstruct the input. During
reconstruction RBM estimates the probability of input x given activation a, this gives us P(x|a) for weight w. Reconstruction is about the probability distribution of the original input.
•We compare the difference between input and reconstruction using
KL divergence.

A network intrusion detection system is designed to detect any breach in the system within the network and also monitor and analyze the data. 10:1 I'd the ratio of attack these days. High accuracy results in high true negative ratio having low number of true positive. In this paper we use RBM as a classifier where NetFlow network traffic info is used to analyze the network.

3.1 Data Analysis

Network traffic collection is the first and key step in intrusion detection. The location of the network collector plays a decisive role in the efficiency of intrusion detection. To provide the best protection for the target host and network, the proposed intrusion detection model is deployed on the nearest victim's router or switch to monitor the inbound network traffic. During the training phase, the collected data samples are categorized according to the transport layer and network layer protocol and are labeled based on the domain knowledge. However, the data collected during the test phase are classified according to the trained hybrid model.

4. REVIEW

Theis, Lucas [7] proposed a method to compress the images using an autoencoder. The technique they have introduced results in the lossy images. Compressive Autoencoder Architecture Younghoon [8](2018) proposed a method based on wireless sensor networks. They are capturing images from a camera sensor and transmit via a wireless sensor network. They are not using any compression algorithm before transferring the data. Shakev et al. has used the rMQTT protocol in WSN. Ramnik et al.2017 [9] proposed WSN image transfer model and used cross layer optimization to transfer the images with limited capability of sensor nodes. Hasin et al.2017 [10] implemented using MQTT, a real time data acquisition. Sahoo et al. 2017[11] used IOT to implement the WSN and used Zigbee to transmit images but zigbee transmit image in 2.4kb/s. Tramel et al. used RBM to solve the CS observation matching problem and found in comparison to Variational autoencoder and generative adversarial networks that RBM is the simplest model to use with minimum no of parameters, even RBM consumes very less energy.

5. ANALYSIS

In this study intrusion detection of network traffic where NSL-KDD dataset is used which contain record of network traffic RBM algorithm is implemented over this dataset. Achieved 99% accuracy. The proposed system showed an overall training accuracy of 99.125% and training loss of 0.029. Therefore the proposed system shows promising results towards intrusion detection in traffic over internet [12] – [20].

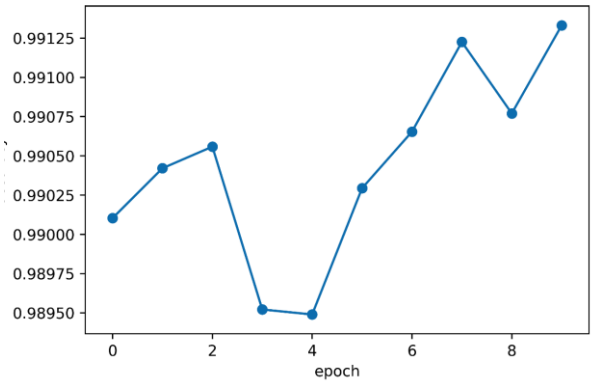

Fig .5. Accuracy of the model

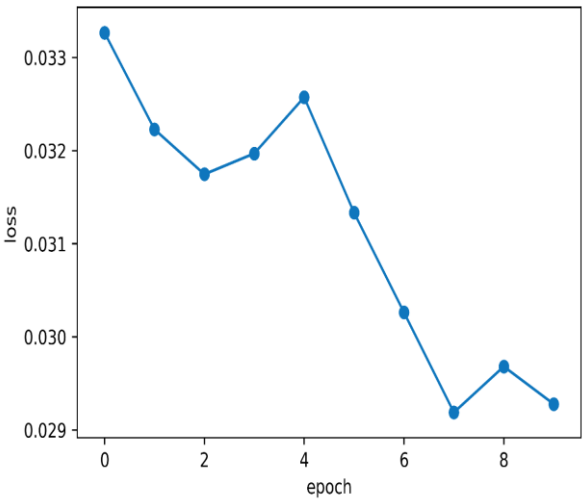

Fig .6. Loss of the model

6. CONCLUSION

The processing capacity and power of nodes in a Wireless Sensor Network (WSN) are restricted. The quality of the images is deficient, and the contents of the images may vary after decoding when we apply image compression algorithms in WSN. Wireless Sensor Nodes (WSNs) play a very significant role in our day-to-day applications. As the resources in each sensor node are limited, it is challenging situation to reduce the energy consumption and increase the lifetime of a sensor node. Currently, the Image Compression algorithms in WSNs are subject to the random changes in image contents. It is difficult to describe various images in the real world with only one kind of image compression. The neural network model is adopted in WSNs to compress the images. In this study we implemented RBM algorithm and achieved 99% accuracy.

7.FUTURE SCOPE

As a future work, we can focus on the following points:

- Reducing the systematic error of the model.
- VAE and GAN deep learning model is used for future work which can help to reduce the semantic error.
- COAP transfer protocol can also be implemented to enhance the security and reliability of data.

REFERENCES

[1]. S. Aruna Deepthi, E.Sreenivasa Rao, M.N.Giriprasad, Design of various Image Compression Methods in Wireless Sensor Networks, International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Volume-9 Issue-1, October 2019.

[2]. Cheng, Chunling, et al. "A multilayer improved RBM network based image compression method in wireless sensor networks." International Journal of Distributed Sensor Networks, vol. 2016, 2016.

[3]. I. Butun, S. D. Morgera, and R. Sankar. A survey of intrusion detection systems in wireless sensor networks. IEEE Communications Surveys Tutorials, 16(1):266–282, First 2014.

[4]. S. R. J. Ramson and D. J. Moni. Applications of wireless sensor networks a survey. In 2017 International Conference on Innovations in Electrical, Electronics, Instrumentation and Media Technology (ICEEIMT), pages 325–329, Feb 2017.

[5]. A. Gouveia and M. Correia. A Systematic Approach for the Application of Restricted Boltzmann Machines in Network Intrusion Detection, volume 10305. 05 2017.

[6]. Lin Bao,1,2 Xiaoyan Sun , 1 Yang Chen,1 Guangyi Man,1 and Hui Shao1, Restricted Boltzmann Machine-Assisted Estimation of Distribution Algorithm for Complex Problems, Hindawi Complexity Volume 2018, Article ID 2609014, 13 pages https://doi.org/10.1155/2018/2609014.

[7]. Jianlin Liu, Fenxiong Chen * and Dianhong Wang, Data Compression Based on Stacked RBM-AE Model for Wireless Sensor Networks, Sensors 2018, 18, 4273; doi:10.3390/s18124273 www.mdpi.com/journal/sensors

[8]. Norouzi, M.; Ranjbar, M.; Mori, G. Stacks of convolutional restricted boltzmann machines for shift invariant feature learning. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2009), Miami, FL, USA, 20–25 June 2009; pp. 2735–2742.

[9]. Hinton, G.E.; Salakhutdinov, R.R. A better way to pretrain deep boltzmann machines. In Proceedings of the Twenty-Sixth Conference on Neural Information Processing Systems, Lake Tahoe, NV, USA, 3–8 December 2012; pp. 2447–2455.

[10]. Tramel, E.W.; Manoel, A.; Caltagirone, F.; Gabrié, M.; Krzakala, F. Inferring sparsity: Compressed sensing using generalized restricted Boltzmann machines.

[11]. M. Tavallaee, E. Bagheri, W. Lu, and A. Ghorbani, “A Detailed Analysis of the KDD CUP 99 Data Set,” Submitted to Second IEEE Symposium on Computational Intelligence for Security and Defense Applications (CISDA), 2009.

[12]. Gupta, M., Kumar, R., & Dewari, S. (2021). Digital twin techniques in recognition of human action using the fusion of convolutional neural network. In *Digital Twin Technology* (pp. 165-186). CRC Press.

[13]. Kour, S., Kumar, R., & Gupta, M. (2021, September). Analysis of student performance using Machine learning Algorithms. In *2021 Third International Conference on Inventive Research in Computing Applications (ICIRCA)* (pp. 1395-1403). IEEE.

[14]. Gupta, M., Kumar, R., Chaudhary, R. K., & Kumari, J. (2021, December). IoT Based Voice Controlled Autonomous Robotic Vehicle Through Google Assistant. In *2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N)* (pp. 713-717). IEEE.

[15]. Gupta, M., Kumar, R., Walia, H., & Kaur, G. (2021, October). Airlines based twitter sentiment analysis using deep learning. In *2021 5th International Conference on Information Systems and Computer Networks (ISCON)* (pp. 1-6). IEEE.

[16]. Kumar, R., Gupta, M., Shukla, S., & Yadav, R. K. (2021, September). E-challan automation for RTO using OCR. In *2021 third international conference on inventive research in computing applications (ICIRCA)* (pp. 1-8). IEEE.

[17]. Bawa, H., Singh, P., & Kumar, R. (2012). An Efficient Novel Key management scheme using NchooseK algorithm for Wireless Sensor Networks. *International Journal of Computer Networks & Communications (IJCNC) Vol, 4*.

[18]. Gupta, M., Solanki, V. K., Singh, V. K., & García-Díaz, V. (2018). Data mining approach of accident occurrences identification with effective methodology and implementation. *International Journal of Electrical and Computer Engineering*, *8*(5), 4033.

[19]. Gupta, M., Solanki, V. K., & Singh, V. K. (2017). A novel framework to use association rule mining for classification of traffic accident severity. *Ingeniería solidaria*, *13*(21), 37-44.

[20]. Puneet, Kumar, R., & Gupta, M. (2022). Optical coherence tomography image based eye disease detection using deep convolutional neural network. *Health Information Science and Systems*, *10*(1), 13.