ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Vol 7, Issue 8, August 2020

Analysis of Efficient Multiclass Cyber-Attack Classification

[1] Prof. Samleti Sandeep Dwarkanath, [2] Dr. P.Balamurugan

Abstract: Inline of vast growth of Internet based uses in recent years, necessity for the security of computer based applications have increased manifolds. As a major source of defense against all the attacks coming its way that needs to adopt to the ever changing threats coming its way. Techniques like machine learning and deep learning can be employed to recognize the reliable detection of anomalies. Anomalies can affect the performance of wireless sensor networks

Index Terms— WSN, IDS, Wireless Sensor Networks, Machine Learning, Deep Learning, Deep Neural Networks

INTRODUCTION

Different machine learning techniques were developed over a period of time for the detection of anomalies in the Wireless sensor networks, protecting them from adverse attacks in information technology. Wireless sensor network is a type of network which can be deployed with the help of tiny or small sensors those are densely located in an adverse areas. WSNs have no predefined infrastructure. The resources constraints play the vital role in the way that WSNs are deployed and maintained.

At present Wireless Sensor Networks are primarily used in aerospace, defense military, weather monitoring, object tracking, bio-medical applications, accumulating the weather-related parameters such as temperature, pressure and disaster management.

Many information and computer based system today are dealing with the sensitive and vital data protecting them from adversaries attacks either internally or externally impacting the performance of overall system. To avoid such adversaries, a well-designed and robust system must be deployed for smooth working of the aforementioned system.

The attacks could be human generated or machine specific generated, diverse and gradually resulting in undetected data theft or breaching at network-level and host-level. Such cyber-attacks could result in loss of money or any other important data. Cyber-attacks are evolving with sophisticated algorithms and techniques used by attackers with ever evolving hardware, software and various network tools and learning approaches. Malicious attacks could impose serious threat not only to financial institutions but also to individuals who are involved in convention of Information and Communication Technology (ICT) tools.

Founded on insensitive behavior, intrusion detection system can be divided as network-based intrusion detection system (NIDS) and host-based intrusion detection system (HIDS) The network activities of any network can be gathered using network components such as switches, routers and analyzed in order to identify attacks and possible threats within the scope of network traffic.

In this work following assumptions are considered :-

- An attacker's try to pretends to be normal user to remain hidden from the IDS.
- The usage pattern of network resources will be captured however the existing methods may end up in high false positive rate and less negative rate.
- The patterns of network intrusions exist in normal traffic with a very low profile over long time interval.

2 RELATED WORK

Some of the research work of the dominant researchers are discussed and relatively studied here in this part of the paper.

Elmasry et.al., in [9] have implemented double Particle Swarm Optimization (PSO) based algorithm and methods to select both feature subset and hyper parameters in one process. The mentioned techniques was exploited in the pre-training phase for selecting the optimized features and models hyper parameters spontaneously. In order to investigate the performance differences, they have utilized three deep learning models, namely Deep Neural networks (DNN), Long Short Term Memory Recurrent Neural Networks (LSTM-RNN) and Deep Belief

Networks (DBN). Also, they have used two common IDS datasets in their experimentation to validate their methodology and proved the effectiveness of developed models. Moreover, many evaluation metrics are used for both binary and multiclass classification to assess the models performance in each of the datasets. Finally, intensive quantitative, Friedman test and ranking method analyses of their results are provided at the end of their publication. Experimental Results proved a significant progress in area like network intrusion detection when adopting their approach by improving Detection Rate (DR) by 4% to 7% and reducing the False Alarm Rate (FAR) by 1 % to 5% from corresponding models without pre-training on the same dataset.

Poornima et.al in [11] has formulated an Online Locally Weighted Projection Regression (OLWPR) for anomaly detection in WSN. Linear Weighted Projection Regression methods are non-parametric and the current predictions were performed by local functions that are used only the subset of data. Hence, the mathematical complexity was turned out be Low which is one the requirement of Wireless Sensor Network. The dimensionality reduction in LWPR was preformed online by Principal Component Analysis(PCA) to handle the unnecessary and duplicate data in the given input data. After the predication process, the dynamic threshold value was found by the dynamic thresholding method to find the deviations of predicted value from the actual sensed value. OLWPR achieves the detection rate of 85% and very low error rate of 16%

Naseer (2018) et.al., in [10] expected to investigate the suitability of deep learning techniques for anomaly-based intrusion detection system. For this research, they designed anomaly detection models based on different deep neural network structures, including Convolutional Neural Networks , Autoencoders and recurrent neural networks . These deep neural network models were trained on NSLKDD training data set and examined on both test dta sets provided by NSLKDD, namely NSLKDDTest+ and NSLKDDTEst21. All the experimentation carried by authors on advance GPU based test bed for testing purpose. Conventional machine learning based intrusion detection models were implemented using well-known classification techniques including extreme learning machine, nearest neighbor, decision tree, random forest, support vector machine, naïve-bays and quadratic discriminant analysis. Both the deep and conventional machine learning models were evaluated and tested using well-known classification metrics and techniques, including receiver operating characteristics, area under curve, precision-recall curve, mean average precision and accuracy of classification. Experimental outcome of deep IDS models shown promising results for real world application in anomaly detection systems.

Wang et.al.,(2017)[12] has proposed a novel IDS called the hierarchical spatial temporal feature based IDS (HAST-IDS) in which it first learns the low level spatial features of network traffic using deep convolutional neural networks (CNNs) and the learns high-level temporal features using long short-term memory network automatically, no feature engineering techniques are mandatory. The automatically learned traffic features effectively reduce the FAR. The standard DARPA1998 and ISCX2012 data sets were used to evaluate the performance of the proposed system practically. The experiment results proved that the HAST-IDS outperforms other published approaches in terms of accuracy, detection rate and FAR, which successfully demonstrates its effectiveness in both feature learning and FAR reductions. HAST-IDS uses deep neural networks which can automatically learn hierarchical spatial temporal directly from raw network traffic data without any intervention making the model more robust for external attacks.

Further two more problems needs to be studied in future work. The first requires improvement in detection of performance on imbalance datasets.

Second problem involves the combining traditional traffic features. Many more published research work proved that in certain cases few manually designed traffic features will be very helpful.

Kurt (2018) et.al in [14] had proposed a robust online detection algorithmic program for (possible combined) false data injection and jamming attacks which also provides online estimates of the unknown and time-varying parameters and recovered state estimates. Also, taking into consideration of smarter attackers who are capable of designing stealthy attacks to prevent the detection or to increase the detection delay of the proposed algorithm , they proposed additional countermeasures . Numerical studies illustrates the quick and reliable response of the proposed detection mechanism against hybrid and stealthy cyber-attacks in smart grids. The drawback of proposed hybrid attack model did not cover network topology attacks as a special case. In further work the generalized state estimation mechanism can be considered in which both the system state and network topology are estimated.

Aditham (2017) et.al., in [13] have put forward a new system architecture in which insider attacks will be detected by utilizing the replication of data on various nodes in the system. Attack models mainly concentrates on misuse of program information by system admins of various big data platforms. The proposed system uses the two step attack detection algorithm and a more secure communication protocol to analyze processes executing in the system. The first step involves the construction of control instruction sequences for each process in the system. The next (second) step carries the matching of these instruction sequences among the replica nodes. The proposed system in this paper used cryptographic system. Initial experiments on real-world hadoop and spark test proved that the proposed system needs to consider only 20% of the code to analyse a program and suffers 3.3% time overhead. The proposed security system can be implemented and built for any big data system due to its extrinsic workflow models.

Kim (2018) et.al in [15] have proposed a novel framework for android malware detection. In their work, this framework uses different types of features to reflect the properties of android applications from various aspects and the features were refined using existence based and similarity based feature extraction method for effective feature representation on malware detection. Besides a multimodal deep learning method was proposed and build to be used in android as a malware detection. This work was first study of the multimodal deep learning to be used in the android malware detection area of research. With the detection model, it was possible to enhance maximum advantages of encompassing multiple types of features. To examine the performance, they have carried out different experiments with total of 41,260 samples. They also compared the accuracy of their developed model with that of other deep neural network models. Also further, they evaluated their model in various aspects including the efficiency in model updates, the usefulness of diverse features and their feature representation method. In addition to above, they compared the performance of framework with those of other existing models and methods including deep learning methods.

3. PROPOSED METHODOLOGY

3.1 Motivation Behind Research Work

An IDS is a proactive intrusion detection tool can be used to detect well in advance and classify intrusion, attacks or violations of the securities policies set up automatically at network-level and host-level infrastructure in a timely manner. Misuse of detection uses predefined signatures and filters to detect and avoid the attacks. It totally depends upon human interface and inputs to constantly update the signature database and depositories. This technique can be more accurate in finding and detecting the known attacks but is completely ineffective in the case of unknown sources of attacks.

Anomaly detection uses heuristic approaches and mechanisms to find the unknown malicious activities and impacts on the overall infrastructure impacting the performance of the system.

In most of the cases and the scenarios, anomaly detection results into a high positive rate [16]. To combat this problem most of the organizations uses the combination of both the misuse and anomaly detection in their commercial solution to the systems. Stateful protocol analysis is most powerful and robust in comparison to the above mentioned detection mechanisms due to the fact that Stateful protocol analysis acts on the network layer, application layer and transport layer. This may use predefined vendor specifications settings to deviations of appropriate protocols and applications.

Hence we can say that deep learning approaches and techniques are most recently can be used to enhance the intelligence and powerfulness of such kind of intrusion detection techniques, in machine learning we have shortage of availability dataset in public domains. So machine learning techniques proves to be inappropriate in our case.

The most common issues in the existing solutions based on machine learning methods are :Firstly the models produces high false positive rate with wider range and different class of attacks; secondly the models are not generalizable as existing studies have mainly used only a single dataset to report the performance of the machine learning model; third approach the model discussed so far have completely unseen in today’s huge network traffic and finally the solutions are required to preserve today’s rapidly increasing high-speed network size, speed and dynamics[17]. The above challenges are major source of motivation for the rest of work.

3.2 Our Contribution :-

i) In this work we tried to propose a Multiclass Cyber classification strategy, an Krill Herd optimized Deep Learning Neural Network (KH-DNN)[18,19] model by combining the network based Intrusion Detection System(NIDS) and host-based intrusion detection system (HIDS) to develop a flexible and effective IDS to detect and classify unforeseen and unpredictable cyber-attacks.

ii) The continuous dynamically change in network and rapid evolution of attacks makes it mandatory to evaluate datasets which can be generated over the period of time through various approaches.

iii) In this work we want to make use of a text representation methods [20] to classify and divide the process behaviors using system call trace instead of feature extractions methods used by classical classifiers.

iv) To prove the efficacy of our outcomes, multi class model in classification of attacks, the analysis and examination of various data set will be carried out under the supervision due to the fact that each dataset incurs from various issues such data corruption or loss, data duplication traffic variety, various in consistencies and contemporary attacks resulting in performance degradation.

4 CONCLUSION

The study of literature proved that a well-designed and well defined multi-stage deep learning mechanism can be worked upon to get the working model to avoid the Intrusion detection System (IDS) that will be robust enough to various kinds of attacks.

REFERENCES

[1] Gungor, V.C.; Lu, B.; Hancke, G.P. Opportunities and challenges of wireless sensor networks in smart grid. IEEE Trans. Ind. Electron. 2010, 57, 3557–3564.

[2] Rassam, M.A.; Maarof, M.A.; Zainal, A. A survey of intrusion detection schemes in wireless sensor networks. Am. J. Appl. Sci. 2012, 9, 1636–1652.

[3] J. Zheng and B. X. Zhang, Wireless Sensor Network Technology, China Machine Press, 2012.

[4] I. Butun, S. D. 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.

[5] A. Ghosal and S. Halder, “A survey on energy efficient intrusion detection in wireless sensor networks,” Journal of Ambient Intelligence and Smart Environments, vol. 9, no. 2, pp. 239–261, 2017.

[6] B. B. Zarpelão, R. S. Miani, C. T. Kawakani, and S. C. de Alvarenga, “A survey of intrusion detection in Internet of Things,” Journal of Network and Computer Applications, vol. 84, pp. 25–37, 2017.

[7] Zhang, F.; Kodituwakku, H.A.D.E.; Hines, W.; Coble, J.B. Multi-Layer Data-Driven Cyber-Attack Detection System for Industrial Control Systems Based on Network, System and Process Data. IEEE Trans. Ind. Inform. 2019, 15, 4362–4369.

[8] A. Abduvaliyev, A. S. K. Pathan, Jianying Zhou, R. Roman, and Wai-Choong Wong, “On the vital areas of intrusion detection systems in wireless sensor networks,” IEEE Communications Surveys & Tutorials, vol. 15, no. 3, pp. 1223–1237, 2013.

[9] Elmasry, Wisam, Akhan Akbulut, and Abdul Halim Zaim. "Evolving deep learning architectures for network intrusion detection using a double PSO metaheuristic." Computer Networks 168 (2020): 107042.

[10] Naseer, Sheraz, Yasir Saleem, Shehzad Khalid, Muhammad Khawar Bashir, Jihun Han, Muhammad Munwar Iqbal, and Kijun Han. "Enhanced network anomaly detection based on deep neural networks." IEEE Access 6 (2018): 48231-48246.

[11] Poornima, I. Gethzi Ahila, and B. Paramasivan. "Anomaly detection in wireless sensor network using machine learning algorithm." Computer Communications (2020).

[12] Wang, Wei, Yiqiang Sheng, Jinlin Wang, Xuewen Zeng, Xiaozhou Ye, Yongzhong Huang, and Ming Zhu. "HAST-IDS: Learning hierarchical spatial-temporal features using deep neural networks to improve intrusion detection." IEEE Access 6 (2017): 1792-1806.

[13] Aditham, Santosh, and Nagarajan Ranganathan. "A system architecture for the detection of insider attacks in big data systems." IEEE Transactions on Dependable and Secure Computing 15, no. 6 (2017): 974-987.

[14] Kurt, Mehmet Necip, Yasin Yılmaz, and Xiaodong Wang. "Real-time detection of hybrid and stealthy cyber-

attacks in smart grid." IEEE Transactions on Information Forensics and Security 14, no. 2 (2018): 498-513.

[15] Kim, TaeGuen, BooJoong Kang, Mina Rho, Sakir Sezer, and Eul Gyu Im. "A multimodal deep learning method for android malware detection using various features." IEEE Transactions on Information Forensics and Security 14, no. 3 (2018): 773-788.

[16] Mishra, P., Varadharajan, V., Tupakula, U., & Pilli, E. S. (2018). A detailed investigation and analysis of using machine learning techniques for intrusion detection. IEEE Communications Surveys & Tutorials.

[17] Azab, A., Alazab, M. & Aiash, M. (2016) "Machine Learning Based Botnet Identification Traffic" The 15th IEEE International Conference on Trust, Security and Privacy in Computing and Communications (Trustcom 2016), Tianjin, China, 23-26 August, pp. 1788-1794.

[18] Gandomi, Amir Hossein, and Amir Hossein Alavi. "Krill herd: a new bio-inspired optimization algorithm." Communications in nonlinear science and numerical simulation 17, no. 12 (2012): 4831-4845.

[19] Kim, Jin, Nara Shin, Seung Yeon Jo, and Sang Hyun Kim. "Method of intrusion detection using deep neural network." In 2017 IEEE International Conference on Big Data and Smart Computing (BigComp), pp. 313-316. IEEE, 2017.

[20] Vinayakumar, R., Mamoun Alazab, K. P. Soman, Prabaharan Poornachandran, Ameer Al-Nemrat, and Sitalakshmi Venkatraman. "Deep learning approach for intelligent intrusion detection system." IEEE Access 7 (2019): 41525-41550.