AN INTRUSION DETECTION SYSTEM MODEL FOR DETECTING KNOWN AND INNOVATIVE CYBER ATTACKS USING SVM ALGORITHM

[1]Arul Selvan M
[1] Department of Computer Science & Engineering, K.L.N College of Engineering, Pottapalayam – 630612, Tamilnadu, India
[1]arul2591@gmail.com

Abstract: Nowadays, intrusions have become a major problem faced by users. To stop these cyber attacks from happening, the development of a reliable and effective Intrusion Detection System (IDS) for cyber security has become an urgent issue to be solved. The proposed IDS model is aimed at detecting network intrusions by classifying all the packet traffic in the network as benign or malicious classes. The Canadian Institute for Cyber security Intrusion Detection System (CICIDS2017) dataset has been used to train and validate the proposed model. The model has been evaluated in terms of the overall accuracy, attack detection rate, false alarm rate, and training overhead. DDOS attacks based on Canadian Institute for Cyber security Intrusion Detection System (KDD Cup 99) dataset has been used to train and validate. For validation, comparison for 2 dataset (CICIDS2017 and KDD Cup 99) is done. Then, to implement the Deep learning algorithms is proposed. Method Classification using SVM algorithm Model predict is done. Testing dataset for anomaly detection model classified as attack or normal. Finally, the experimental results shows that the performance metrics such as accuracy, precision, recall and confusion matrix.

Key words: CICIDS, Intrusion Detection System, Deep Learning, Cyber Security & malicious

Introduction:

Machine learning is a branch of artificial intelligence and computer science which focuses on the use of data and algorithms to imitate the way that humans learn, gradually improving its accuracy. Machine learning is an important component of the growing field of data science. Through the use of statistical Algorithms are trained to make classifications or predictions, uncovering key insights within data mining projects. These insights subsequently drive decision

Corresponding Author: *Arul Selvan M*
K.L.N. College of Engineering, Pottapalayam, Tamil Nadu, India
Mail: arul2591@gmail.com

making within applications and businesses, ideally impacting key growth metrics. As big data continues to expand and grow, the market demand for data scientists will increase, requiring them to assist in the identification of the most relevant business questions and subsequently the data to answer them.

Support Vector Machine (SVM) is a supervised machine learning algorithm used for both classification and regression. Though regression problems are well its best suited for classification. The objective of SVM algorithm is to find a hyperplane in an N-dimensional space that distinctly classifies the data points. SVMs are used in applications like handwriting recognition, intrusion detection, face detection, email classification, gene classification, and in web pages. This is one of the reasons for using SVMs in machine learning. It can handle both classification and regression on linear and non-linear data. The Naive Bayes classification algorithm is a probabilistic classifier. It is based on probability models that incorporate strong independence assumptions. The independence assumptions often do not have an impact on reality. Therefore they are considered as naive. Naive Bayes algorithm is a supervised learning algorithm, which is based on Bayes theorem and used for solving classification problems. It is mainly used in text classification that includes a high-dimensional training dataset.

Innovative Cyberattacks Using Convolutional Neural Network:

As a tremendous amount of service being streamed online to their users along with massive digital privacy information transmitted in recent years, the internet has become the backbone of most people's everyday workflow. The extending usage of the internet, however, also expands the attack surface for cyberattacks. If no effective protection mechanism is implemented, the internet will only be much vulnerable and this will raise the risk of data getting leaked or hacked. The focus of this paper is to propose an Intrusion Detection System (IDS) based on the Convolutional Neural Network (CNN) to reinforce the security of the internet. The proposed IDS model is aimed at detecting network intrusions by classifying all the packet traffic in the network as benign or malicious classes. The Canadian Institute for Cybersecurity Intrusion Detection System (CICIDS2017) dataset has been used to train and validate the proposed model. The model has been evaluated in terms of the overall accuracy, attack detection rate, false alarm rate, and training overhead. A comparative study of the proposed model's performance against nine other well-known classifiers has been presented.

Multi-step Attack Detection:

Since the beginning of the Internet, cyber-attacks have threatened users and organizations. They have become more complex concurrently with computer networks. Nowadays, attackers need to perform several intrusion steps to reach their final objective. The set of these steps is known as multi-step attack, multi-stage attack or attack scenario. Their multi-step nature

hinders intrusion detection, as the correlation of more than one action is needed to understand the attack strategy and identify the threat. Since the beginning of 2000s, the security research community has tried to propose solutions to detect this kind of threat and to predict further steps. This survey aims to gather all the publications proposing multi-step attack detection methods.Focusing on methods that go beyond the detection of a symptom and try to reveal the whole structure of the attack and the links between its steps. Following a systematic approach to bibliographic research in order to identify the relevant literature. Our effort results in a corpus of publications covering methods, which describe and classify. The analysis of the publications allows us to extract some conclusions about the state of research in multi-step attack detection.

Clustering-based real-time anomaly detection:

Off late, the ever increasing usage of a connected Internet-of-Things devices has consequently augmented the volume of real-time network data with high velocity. At the same time, threats on networks become inevitable; hence, identifying anomalies in real time network data has become crucial. The evaluation is done by performing critical comparative analysis using existing approaches, such as K-means, hierarchical density-based spatial clustering of applications with noise (HDBSCAN), isolation forest, spectral clustering and agglomerative clustering. The outcome of the evaluation has substantially proven the efficacy of the proposed framework with a much higher accuracy rate of 96.51% when compared to other algorithms. Besides, the proposed framework also outperformed the existing algorithms in terms of lesser memory consumption and execution time. Ultimately the proposed solution enable analysts to precisely track and detect anomalies in real time.

The spark iterative computation architectural enables large-scale machine learning algorithms to achieve high level efficiency in results, and spark.ml. API for pipeline offers developers with extensive range of new module to integrate with their architecture. Here algorithms have their own benefits in terms of anomaly detection, processing of data, providing accuracy, memory consumption, and execution time. Network traffic anomaly may indicate a possible intrusion in the network and therefore anomaly detection is important to detect and prevent the security attacks. The early research work in this area and commercially available Intrusion Detection Systems (IDS) are mostly signature-based. The problem of signature based method is that the database signature needs to be updated as new attack signatures become available and therefore it is not suitable for the real-time network anomaly detection. The recent trend in anomaly detection is based on machine learning classification techniques. Applying seven different machine learning techniques with information entropy calculation to Kyoto 2006+ data set and evaluate the performance of these techniques. Our findings show that, for this particular data set, most machine learning techniques provide higher than 90% precision, recall

and accuracy. However, using area under the Receiver Operating Curve (ROC) metric, find that Radial Basis Function (RBF) performs the best among the seven algorithms studied in this work.

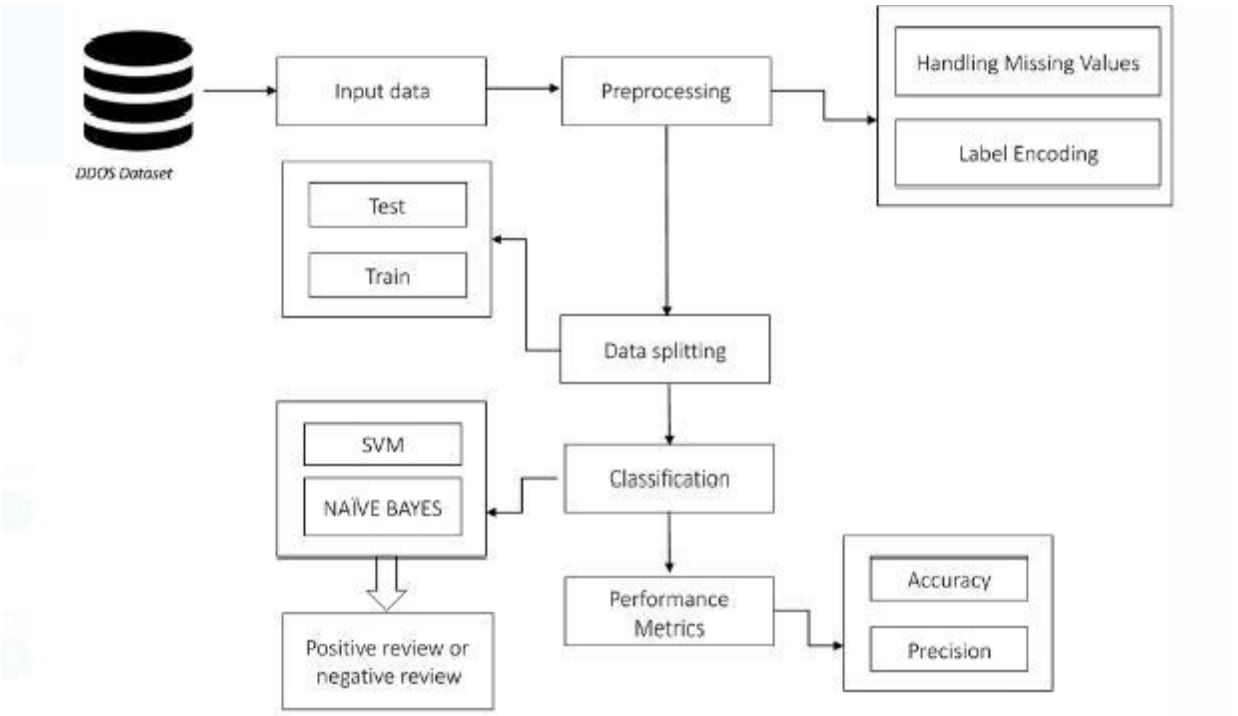

Fig.1. Architecture Diagram:

Machine Learning Techniques for Network Intrusion:

As an essential component of the critical infrastructure, the Industrial Control System (ICS) is facing increasing cyber threats. The emergence of the Shodan search engine also magnified this threat. Since it can identify and index Internet-connected industrial control devices, the Shodan search engine has become a favorite toolkit for attackers and penetration testers. In this paper, honeypot technology is used to conduct a comprehensive exploring on Shodan search engine. First deploy six distributed honeypot systems and collect three-month traffic data. For exploring Shodan, design a hierarchical DFA-SVM recognition model to identify Shodan scans based on the function code and traffic feature, which is adapted to find the Shodan and Shodan-like scanners superior to the predominant method of reverse resolving IPs. Finally, conducting an in-depth analysis for Shodan scans and evaluate the impact of Shodan on industrial control systems in terms of scanning time, scanning frequency, scanning port, region preferences, ICS protocol preferences and ICS protocol function code proportion. Accordingly, providing some defensive measures to mitigate Shodan threat. The main advantage of SVM is a machine learning model with the advantages of high detection rate of small samples and strong generalization ability, which is suitable for handling high-dimensional and non-linear Shodan

traffic from a small amount of Shodan scanners. Flow-based intrusion detection is an innovative way of detecting intrusions in high-speed networks. Flow-based intrusion detection only inspects the packet header and does not analyse the packet payload. This paper gives an introduction to a flow-based intrusion detection system and surveys state of the art in flow-based intrusion detection. It also describes the available flow-based datasets used for evaluation of flow-based intrusion detection systems. The paper proposes a taxonomy for flow-based intrusion detection systems on the basis of the technique used for detection of maliciousness in IP flow records.Reviewing the architecture and evaluation results of available flow-based intrusion detection systems and identify important research challenges for future research in the area of flow-based intrusion detection.

Industrial Control Systems:

Intrusion Detection has been heavily studied in both industry and academia, but cyber security analysts still desire much more alert accuracy and overall threat analysis in order to secure their systems within cyberspace. Improvements to Intrusion Detection could be achieved by embracing a more comprehensive approach in monitoring security events from many different heterogeneous sources. Correlating security events from heterogeneous sources can grant a more holistic view and greater situational awareness of cyber threats. One problem with this approach is that currently, even a single event source (e.g., network traffic) can experience Big Data challenges when considered alone. Attempts to use more heterogeneous data sources pose an even greater Big Data challenge. Big Data technologies for Intrusion Detection can help solve these Big Heterogeneous Data challenges. In this paper, reviewing the scope of works considering the problem of heterogeneous data and in particular Big Heterogeneous Data. Discussing the specific issues of Data Fusion, Heterogeneous Intrusion Detection Architectures, and Security Information and Event Management (SIEM) systems, as well as presenting areas where more research opportunities exist. Overall, both cyber threat analysis and cyber intelligence could be enhanced by correlating security events across many diverse heterogeneous sources.

The main advantage is Heterogeneity among the actual Sensors, IDSs, Analyzers, or even SIEMs can be beneficial for Intrusion Detection where detection accuracy can be improved. The system requirement is the first step in the requirements analysis process. It lists the requirements of a particular software system including functional, performance and security requirements. The requirements also provide usage scenarios from a user, an operational and an administrative perspective. The purpose of software requirements specification is to provide a detailed overview of the software project, its parameters and goals. This describes the project target audience and its user interface, hardware and software requirements.

Conclusions:

Intrusion Detection System (IDS) for cybersecurity based on a SVM (Support vector Machine) is proposed. The Support vector machine (SVM) allow the proposed IDS model to learn complicated patterns of features form network traffic, while maintaining reasonable storage and computation overhead. In this present study, proposed various predictive models trained on several ML algorithms for predicting Intrusion detection in social network is reliable. The Algorithms are implemented and predict the result based on accuracy, precision, recall and f1-measure. Thus any intrusions that may happen in future may be analyzed and predicted. Preventive measures can be taken if any attack is predicted as the result of the module. In future, will explore the application of more advanced deep learning methods and possible combinations of machine learning. A Denial-of-Service (DoS) attack is an attack meant to shut down a machine or network, making it inaccessible to its intended users. DoS attacks accomplish this by flooding the target with traffic, or sending it information that triggers a crash. This can further be enhanced my monitoring the system using hardware. Warnings can be given to the administrator or user if flooding or attack occurs. The attacks or purposeful flooding of the server may be analyzed and preventive measures or scaling of the server or cloud can be done in order for smooth work without unnecessary traffic.

References:

1. Dilip, R., & Bhagirathi, V. Remote monitoring and control systems in hazardous area.
2. Bhagirathi, V., Meghana, M., & Dilip, R. (2013). Image Processing Techniques for Coin Classification using Labview.
3. Dilip, R., & Bhagirathi, V. (2013). LAN Based Industrial Automation with GSM Connectivity. ICSEM-2013 Conference Proceedings.
4. Dilip, R., & Bhagirathi, V. (2013). Cell Phone Based Liquid Inventory Management Using Wireless System.
5. Dilip, R., & Bhagirathi, V. (2013). Image processing techniques for coin classification using LabVIEW. *OJAI 2013*, *1*(1), 13-17.
6. Dilip, R. (2019). DESIGN AND DEVELOPMENT OF INTELLIGENT SYSTEM FOR HUMAN BODY DESIGN AND DEVELOPMENT OF INTELLIGENT SYSTEM FOR HUMAN BODY. *no. July*, 0-3.
7. Dilip, R., Akash, G., Bhat, A. U., Agil, K., & Tej, R. K. (2020). Design of prototype battery management system.
8. Dilip, R., Sandeep, K., Chandrashekhar, L., & Swamy, S. R. (2020). Revitalized Screen Reader for Visually Challenged.
9. Dilip, R., Bharath, N., Kushal, D. K., Karunakar, L., & Varghese, S. (2020). Design and Development of E-Prescription System.
10. Dilip, R., Vidya, C., Ayushi, G., Kurup, K. R., & Sarveshwar, S. (2020). Design and Development of E-Rodeo A Hybrid Electric-Cycle.
11. Dilip, R., & Ramesh, K. B. (2020). Development of Graphical System for Patient Monitoring using Cloud Computing.
12. Rekha, C. M., Shivakumar, K. S., & Dilip, R. (2020, October). Comparison of spacefactor, capacitance value and impregnated temperature in mpp oil impregnated polypropylene film AC

capacitors. In *2020 International Conference on Smart Technologies in Computing, Electrical and Electronics (ICSTCEE)* (pp. 544-547). IEEE.

13. Krishna, K. M., Borole, Y. D., Rout, S., Negi, P., Deivakani, M., & Dilip, R. (2021, September). Inclusion of cloud, blockchain and iot based technologies in agriculture sector. In *2021 9th international conference on cyber and IT service management (CITSM)* (pp. 1-8). IEEE.
14. Dilip, R., Borole, Y. D., Sumalatha, S., & Nethravathi, H. M. (2021, September). Speech based biomedical devices monitoring using LabVIEW. In *2021 9th International Conference on Cyber and IT Service Management (CITSM)* (pp. 1-7). IEEE.
15. Dilip, R., Milan, R. K., Vajrangi, A., Chavadi, K. S., & Puneeth, A. S. (2021, November). Jumping robot: a pneumatic jumping locomotion across rough terrain. In *Journal of Physics: Conference Series* (Vol. 2115, No. 1, p. 012008). IOP Publishing.
16. Seshanna, S., & Seshanna, M. (2015). Learning Inclusiveness and Under-served Communities in India. *International Journal of Physical and Social Sciences*, *5*(10), 142-147.
17. Seshanna, S., & Seshanna, M. (2016). The impact personality traits, role conflict and work family conflicton customer orientation: a review of extant literature. *International Journal of Research in Social Sciences*, *6*(2), 466-480.
18. Seshanna, S., & Seshann, M. (2017). The applied experiential learningmethod in entrepreneurship education: A conceptual approach. *International Journal of Research in Social Sciences*, *7*(5), 481-488.
19. Seshanna, S., & Seshanna, M. (2018). Midas Ventures A case of a financial services aggregator. *International Journal of Research in Social Sciences*, *8*(4), 159-162.
20. Seshanna, M. INVESTORS BEHAVIOURAL STUDY ON ART AS AN ALTERNATIVE INVESTMENT ASSET CLASS.
21. Seshanna, M., Kumar, H., Seshanna, S., & Alur, N. (2021). THE INFLUENCE OF FINANCIAL LITERACY ON COLLECTIBLES AS AN ALTERNATIVE INVESTMENT AVENUE: EFFECTS OF FINANCIAL SKILL, FINANCIAL BEHAVIOUR AND PERCEIVED KNOWLEDGE ON INVESTORS'FINANCIAL WELLBEING. *Turkish Online Journal of Qualitative Inquiry*, *12*(4).
22. Seshanna, M., Periasamy, P., & Seshanna, S. (2021). ART AS AN ALTERNATIVE INVESTMENT ASSET CLASS IN EMERGING ECONOMIES: A STUDY LINKING PERSONALITY FACTORS TO INVESTOR BEHAVIOUR. *Turkish Online Journal of Qualitative Inquiry*, *12*(6).