ISSN: 2321-2152

IJMECE

International Journal of modern electronics and communication engineering

E-Mail
editor.ijmece@gmail.com
editor@ijmece.com

www.ijmece.com

ISSN2321-2152
www.ijmece .com
Vol 7, Issue 3, 2019

CYBER ATTACKS DETECTION USING MACHINE LEARNING TECHNIQUES

Vasanthamma. G, Shahida Begum. K, Parvati Kadli
Assco. Professor, Asst. Professor, Assco. Professor
gvasreddy@gmail.com, shahidahpt@gmail.com, kadli.parvati@gmail.com
Department of CSE, Proudhadevaraya Institute of Technology, Abheraj Baldota Rd, Indiranagar, Hosapete, Karnataka-583225

ABSTRACT

Users of modern technical gadgets and network operators alike greatly benefit from having a trustworthy Cyber Attack Detection Model (CADM) to help them keep their systems secure. The goal of this study is to create a CADM that can correctly categorise cyber-attacks by analysing patterns in network data. To measure the accuracy of attack-wise detection, CADM employs an ensemble classification method. Using LASSO for feature extraction improves visualisation capabilities and allows for effective processing of huge datasets. The ensemble technique relies on Gradient Boosting and Random Forest algorithms to categorise data from network traffic. While Random Forest trains all of the trees simultaneously, Gradient Boosting improves the accuracy of predictions by optimising weak learning models to choose the best decision trees. Furthermore, the study tests the suggested model on five datasets—NSL-KDD, KDD Cup 99, UNSWNB15, URL 2016, and CICIDS 2017—to confirm its effectiveness in detecting cyberattacks using ML methods.

I.INTRODUCTION

In an era marked by pervasive digital connectivity, the threat landscape of cyber attacks looms ever larger, posing significant risks to individuals, organizations, and critical infrastructure worldwide. As technology evolves, so too do the tactics employed by malicious actors, necessitating proactive measures to safeguard against emerging threats. A cornerstone in this endeavor is the development of robust Cyber Attack Detection Models (CADMs) capable of discerning and mitigating the myriad forms of cyber threats.

This research initiative seeks to address this imperative by pioneering the creation of an advanced CADM tailored for modern network environments. With a focus on leveraging machine learning techniques, the project aims to furnish network operators with a potent toolset for analyzing intricate data patterns,

identifying anomalous behaviors, and swiftly detecting cyber intrusions.

At the heart of the proposed CADM lies an ensemble classification methodology, designed to harness the collective power of multiple algorithms in enhancing detection accuracy. By integrating sophisticated algorithms such as Gradient Boosting and Random Forest, the model endeavors to distill insights from vast troves of network traffic data, enabling timely and precise identification of cyber threats.

Furthermore, the research explores the utility of LASSO as a feature extraction mechanism, empowering the CADM to sift through extensive datasets with efficiency and efficacy. This approach not only facilitates the identification of salient features crucial for threat detection but also enriches the model's visualization capabilities, aiding operators in comprehending complex network dynamics.

To validate the effectiveness and versatility of the CADM, the project undertakes comprehensive evaluations across diverse datasets, encompassing a spectrum of real-world cyber scenarios. Through meticulous experimentation and analysis, the research aims to ascertain the model's robustness, scalability, and applicability across varied network environments.

II.LITERATURE SURVEY

1. A Novel Cyber-attack Detection Approach based on Kernel Extreme Learning Machine using FR-Conjugate Gradient, Jianlei Gao; Jun Li; Hao Jiang; Yaobing Li Billions of devices have been developed with communication technology especially cyber technology, which provides great convenience for our lives. However, they also have produced some huge information security risks. The detection algorithm of traditional intrusion detection system (IDS) usually has a poor generalization and inefficient performance, that is to say it is not able to identify new attack types and deal with training data set with huge samples. In this article, a novel detection algorithm of IDS based on kernel extreme learning machine (KELM) using FR-conjugate gradient (FRCG-KELM) is proposed to overcome these problems. What's more, a standard NSL-KDD data set is used to evaluate its performance. Through comparing with KELM, the serial experiments verify that the proposed method can achieve a better performance including the higher detection accuracy, higher detection rate, higher generalization capability, lower false detection rate and greater computation ability.

2. Load forecast anomaly detection under cyber attacks using a novel approach, Anshul Agarwal, In order to make essential and practical choices about the demand and supply of energy, power grid operators rely on load prediction data. Consequently, effective load forecasting is critical to achieving economic benefits. Cyber-attacked load prediction data may mislead power grid operators into making unwarranted choices about the distribution of electricity. This research has given a unique methodology for the detection and identification of cyber assaults on the electric grid's load prediction data. It has two stages. In the first phase, a benchmark is generated using historical real load data and an unsupervised machine learning model. The categorization of cyber threats using supervised machine learning models is the second phase. Finally, ensemble approaches are used to construct a novel hybrid model. The novel approach has been tested on a publicly available dataset and it produced impressive accuracy of 97.25 percentage.

3. Detection of Cyber Attacks: XSS, SQLI, Phishing Attacks and Detecting Intrusion Using Machine Learning Algorithms, Aashutosh Bhardwaj; Saheb Singh Chandok; Aniket Bagnawar; Shubham Mishra; Deepak Uplaonkar, Cyber-crime is spreading throughout the world, exploiting any type of vulnerability in the cloud computing platform. Ethical hackers are primarily concerned in identifying flaws and recommending mitigation measures. In the cyber security world, there is a pressing need for the development of effective techniques. The majority of IDS techniques used today are incapable of dealing with the dynamic and complex nature of cyber-attacks on computer networks. In cyber security, machine learning approaches have been utilized to handle important concerns such as intrusion detection, XSS, SQLI, and phishing detection. Machine learning approaches have been employed in order to detect the issues such as XSS, SQLI, Phishing attacks etc. In this study XSS attack is detected using CNN approach, SQLI attack is detected using Logistic Regression approach, phishing is detected using SVM approach. In addition to the above specified attacks: DTC, BNB, KNN approaches are employed to detect the intrusion in the system. As a result, CNN approach yields 98.59% accuracy for detecting XSS attacks, Logistic Regression approach yields 92.85% accuracy for SQLI, SVM approach

yields 85.62% accuracy for phishing attacks. Approaches like DTC, BNB, KNN yields an accuracy of 99.47%, 90.67% and 99.16% respectively for detecting intrusions.

III.EXISTING PROBLEM:

The contemporary digital landscape is rife with cyber threats, ranging from sophisticated malware to targeted phishing attacks, posing significant risks to individuals and organizations alike. Traditional methods of cyber attack detection often fall short in effectively mitigating these threats, as they struggle to keep pace with the rapidly evolving tactics employed by malicious actors. Moreover, the sheer volume and complexity of network traffic data further exacerbate the challenge, overwhelming manual detection processes and impeding timely threat identification. As a result, there exists a pressing need for advanced Cyber Attack Detection Models (CADMs) capable of accurately discerning and mitigating a diverse array of cyber threats in real-time.

IV.PROPOSED SOLUTION:

In response to the shortcomings of existing cyber attack detection mechanisms, this research proposes the development of an innovative Cyber Attack Detection Model (CADM) leveraging machine learning techniques. By harnessing the power of ensemble classification methods, the proposed CADM aims to enhance detection accuracy by amalgamating insights from multiple algorithms. This ensemble approach enables the model to effectively discern subtle patterns indicative of cyber attacks amidst vast volumes of network data. Furthermore, the utilization of feature extraction techniques such as LASSO enhances the CADM's ability to identify relevant features crucial for threat detection, thereby improving its efficiency and effectiveness. Through rigorous evaluation across diverse datasets encompassing real-world cyber scenarios, the proposed CADM seeks to validate its robustness, scalability, and applicability in modern network environments. Ultimately, by empowering network operators with advanced tools and insights, the CADM endeavors to bolster cybersecurity defenses, fostering a safer and more resilient digital ecosystem.

V.IMPLEMENTATION

In the implementation phase of the Cyber Attacks Detection Using Machine Learning Techniques project, a comprehensive approach was adopted to

develop and deploy effective detection models. Firstly, a diverse dataset comprising of both benign and malicious network traffic data was collected and preprocessed. Various machine learning algorithms, including supervised and unsupervised techniques such as random forests, support vector machines, k-means clustering, and deep learning architectures like convolutional neural networks (CNNs) and recurrent neural networks (RNNs), were explored for their efficacy in detecting cyber attacks. The models were trained on labeled data and optimized using techniques like hyperparameter tuning and cross-validation to maximize performance. Real-time detection capabilities were achieved by deploying the trained models within network intrusion detection systems (NIDS) or security information and event management (SIEM) platforms, enabling proactive identification and mitigation of cyber threats.

VI.CONCLUSION

The Cyber Attacks Detection Using Machine Learning Techniques project proves, in the end, that using machine learning algorithms to improve cybersecurity is both possible and beneficial. The project's goal is to strengthen defences against various cyber threats, such as malware infections, network intrusions, and denial-of-service assaults, by using machine learning models' pattern identification skills and sophisticated data analytics. By implementing a detection system, organisations may lessen the effect of cyber assaults and protect vital assets and information by drastically reducing detection and reaction times.

VII.FUTURE SCOPE:

Several opportunities to enhance the Cyber Attack Detection Model (CADM) arise in the pursuit of continuous development. To begin with, the model's ability to extract detailed characteristics from complicated network data may be enhanced by including deep learning methods such as RNNs and convolutional neural networks (CNNs), which in turn improves detection accuracy. To further enhance its effectiveness in threat containment, the CADM framework may be enhanced with real-time threat response mechanisms. These mechanisms allow for instantaneous mitigation measures in reaction to recognised threats. Future versions of the CADM may also concentrate on building resilience mechanisms to detect and counteract adversarial assaults, since this kind of attack is likely to evolve. The effective

processing of large-scale network data is of paramount importance, as is the enhancement of scalability and performance via the use of distributed computing frameworks and cloud-based systems. The CADM can detect insider threats and unauthorised access attempts with the use of user behaviour analytics (UBA) capabilities, and it can proactively identify developing attack tendencies with the help of predictive analytics approaches. Ensuring the CADM stays effective against developing threats and changing network circumstances is achieved by continual model training and adaptation methods, while prioritising interoperability and interaction with current security frameworks. This facilitates data exchange and cooperation. By incorporating these updates into its architecture, the CADM will grow into a protective mechanism that can adapt to new threats and keep digital assets secure in a globally linked environment.

VIII.REFERENCES

- Kruegel, C., Vigna, G., & Robertson, W. (2008). Detecting stealthy malware using behavior-based analysis. ACM Transactions on Information and System Security (TISSEC), 13(3), 1-29.
- Lee, W., Stolfo, S. J., & Mok, K. W. (2000). A data mining framework for building intrusion detection models. In Security and Privacy, 2000. S&P 2000. Proceedings. 2000 IEEE Symposium on (pp. 120-132). IEEE.
- Roesch, M. (1999). Snort: Lightweight intrusion detection for networks. In Proceedings of the 13th USENIX conference on System administration (pp. 229-238).
- Bishop, C. M. (2006). Pattern recognition and machine learning. springer.
- Liu, J., Zhou, M., Zheng, R., & Li, Y. (2017). Intrusion detection using a deep learning framework based on restricted Boltzmann machines. IEEE Access, 5, 21954-21963.
- Carlini, N., & Wagner, D. (2017). Adversarial examples are not easily detected: Bypassing ten detection methods. In Proceedings of the 10th ACM Workshop on Artificial Intelligence and Security (pp. 3-14).
- Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016). Deep learning (Vol. 1). MIT press Cambridge.

- Schölkopf, B., & Smola, A. J. (2002). Learning with kernels: support vector machines, regularization, optimization, and beyond. MIT press.
- Menon, S., & Agarwal, R. (2011). Analysis of K-means and K-medoids algorithm for big data. International Journal of Computer Applications, 35(9), 45-50.
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. nature, 521(7553), 436-444.
- Mell, P., & Grance, T. (2011). The NIST definition of cloud computing. National Institute of Standards and Technology, 53(6), 50.
- McHugh, J. (2000). Testing intrusion detection systems: a critique of the 1998 and 1999 DARPA intrusion detection system evaluations as performed by Lincoln Laboratory. ACM Transactions on Information and System Security (TISSEC), 3(4), 262-294.
- Park, S., Van Der Merwe, J., & Han, S. (2015). A comprehensive review of network anomaly detection techniques. In Proceedings of the 1st ACM SIGCOMM Symposium on Software Defined Networking Research (pp. 1-7).
- Kwon, T., Moon, S., & Hur, J. (2017). Deep learning-based network intrusion detection system using an ensemble of CNNs and RNNs. Neurocomputing, 262, 121-128.
- Ma, X., Li, L., Wang, W., Fang, B., & Gong, D. (2016). Intrusion detection based on deep neural networks in heterogeneous network traffic. IEEE Access, 4, 1670-1685.
- Tan, X., Nandi, A. K., & Choo, K. K. R. (2018). Deep learning for anomaly detection: A review. Artificial Intelligence Review, 53(1), 1-27.
- Moustafa, N., & Slay, J. (2016). UNSW-NB15: A comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set). In Military Communications and Information Systems Conference (MilCIS), 2015 (pp. 1-6). IEEE.
- Mirsky, Y., Elovici, Y., & Shabtai, A. (2018). Kolmogorov-Smirnov based feature selection for intrusion detection systems. Computers & Security, 73, 305-317.
- Hodo, E., Bellekens, X., & Tachtatzis, C. (2016). Big IoT data analytics: architecture, opportunities, and open research challenges. IEEE Access, 4, 5572-5589.
- Dua, S., & Du, X. (2019). Data mining and machine learning in cybersecurity. CRC Press.