IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Evaluating Threat Detection Techniques In Software Defined Networking Environments

[1]Chetan Gupta, [2]Namrata Patel, [3]Sonam Dubey
[1]Assistant Professor, Dept. of CSE, SIRT, Bhopal, India
[2]M. Tech Research Scholar, Dept. of CSE, SIRTS, Bhopal, India
[3]Assistant Professor, Dept. of CSE, SIRT, Bhopal, India

ABSTRACT: A network architecture known as software-defined networking (SDN) moves the control panel to an application known as Controller, separating it from the data panel. The newest method for configuring, managing, and running networks is called software defined networking, or SDN. The network or application infrastructure's centralized store of policy and control instructions is called an SDN controller. Data across the network is vulnerable to assaults if it is not properly secured. We can set up SDN controllers to identify assaults like denial-of-service (DDOS) that can cause network congestion by using SDN policies. When the attack is identified, we can use the policies to determine whether to forward or drop the packet.

KEYWORDS: ***SDN, DDOS, Network Architecture, Network Attack, IDS.***

I. Introduction

The introduction of software-defined networks (SDN) has made monitoring network operations crucial. The controller in SDN is in charge of managing the network's complexity, which gives the network administrator abstraction. We rely on the controller's automated application of network regulations and rules, so we don't have to setup every device individually. SDN transfers the control functionality to a logical centralized software-based network controller, separating it from the data functionality.

A standard protocol called Open Flow determines how an SDN controller communicates with the network devices. An Open Flow-based switch exposes the controller to its flow table, provides an abstraction of it, and grants the controller the ability to modify values by adding, altering, or deleting rules from the table. Applications operating on the network controller can effortlessly manage the flow of incoming and outgoing packets in switches by utilizing Open Flow.

Network operators can use frameworks based on the Open Flow protocol to establish security rules and policies according to their own needs. Let's say a campus administrator has to monitor incoming email traffic to a spyware detection device and web traffic to an intrusion detection system (IDS). Instead of manually configuring every device, we want to use SDN to enable the operator to construct a high-level policy to accomplish this. Additionally, let's say that the sender needs to be prevented from accessing the network when the IDS identifies malicious activity. We want the edge router to automatically block the sender rather than requiring the operator to manually deny access to the source. Instead of individually setting each device in the network, operators may focus on defining simple rules and policies thanks to the network's abstraction. The framework includes a software layer that operates on top of the network controller and a number of external devices that report to the controller and carry out security-based functions including firewalls and intrusion detection systems (IDS). Giving network operators the ability to create security rules and policies for desired flows is the framework's true goal. A description of the

flow, a list of security services that apply to the flow, and the rules' response in the event that malicious material is detected are all included in the rules. This response could be a simple warning or a way to limit traffic or prevent all packets from a particular source from flowing.

II. Literature Review

The increasing reliance on data-driven techniques in financial decision-making has fueled substantial research into machine learning algorithms for credit risk assessment and loan prediction. Early applications of neural networks demonstrated their capability in handling non-linear relationships in credit [25], while more recent advancements in deep learning further improved prediction accuracy [26]. Traditional models such as Naïve Bayes and logistic regression, though simple, provided a baseline for performance comparison [5]. The emergence of ensemble methods such as Random Forest [29], Gradient Boosting [27], and the more powerful XGBoost [28] has significantly advanced the predictive capacity of credit models. Moreover, Support Vector Machines (SVMs) have been explored extensively for their robustness in high-dimensional spaces. Techniques like Elastic Net regularization have been employed for variable selection, improving model interpretability and reducing over fitting. Data preprocessing methods, such as the SMOTE algorithm for addressing class imbalance and various missing data imputation strategies, play crucial roles in preparing reliable training datasets. Several studies have also utilized Principal Component Analysis and cross-validation technique to enhance model performance and generalizability. Furthermore, real-world implementations in peer-to-peer lending platforms illustrate the practical relevance of these models. Challenges such as rising default rates highlight the need for robust and adaptive models capable of mitigating financial risk. Ultimately, the integration of machine learning in banking and finance continues to evolve, promising more efficient, accurate, and scalable solutions to credit risk management [30][31][32][33].

Neural networks are among the earliest models used for credit risk assessment, offering promising results in capturing complex, non-linear relationships within financial data [1]. As outlined in the report by the Department of Computing, Imperial College London, the application of back propagation and multilayer perceptron's helped automate decision-making processes in financial domains. These models laid the foundation for more advanced deep learning architectures.

The Lending Club platform [2], exemplifies the practical application of machine learning in peer-to-peer lending, enabling investors to make informed decisions based on borrower profiles and risk factors (Lending Club, n.d.). By leveraging alternative data and algorithmic models, such platforms facilitate more efficient credit evaluations and democratize investment opportunities in the financial sector.

Elastic Net regularization [3], as discussed on R-bloggers, is a robust technique for variable selection that combines the strengths of both LASSO and Ridge regression. It is particularly useful in high-dimensional datasets typical of credit data, allowing for more interpretable models while avoiding over fitting through regularization of coefficients.

Addo, Guegan, and Hassani [4], explored the application of machine learning and deep learning models for credit risk analysis demonstrating superior performance compared to traditional statistical methods. Their study in Risks highlights the growing reliance on advanced computational techniques in assessing borrower default probabilities.

Antonakis and Sfakianakis [5], evaluated the Naïve Bayes classifier as a screening tool for credit applicants, finding that while the model is computationally simple, it performs competitively in scenarios where independence assumptions are reasonably satisfied. Their research in the Journal of Applied Statistics contributes to validating classical probabilistic models in financial decision-making.

Attigeri, Pai, and Pai [6], conducted a comparative study of machine learning algorithms such as Decision Trees and Support Vector Machines for credit risk assessment. Published in Advanced Science Letters, their work underscores the adaptability and accuracy of these models in classifying borrowers based on risk.

Batista and Monard [7], analyzed various methods for handling missing data in supervised learning tasks. Their work in Applied Artificial Intelligence is particularly relevant for financial datasets, which often suffer from incomplete records, and emphasized the impact of imputation techniques on model performance.

Bekhet and Eletter [8], developed a neural scoring model for credit risk assessment in Jordanian commercial banks. Their findings, published in the Review of Development Finance, demonstrated the effectiveness of neural networks in predicting creditworthiness, offering a viable alternative to conventional scoring systems.

Chawla et al. [9], introduced SMOTE (Synthetic Minority Over-sampling Technique), a method designed to address class imbalance by generating synthetic examples of minority classes. This technique is especially relevant in credit risk modeling, where default cases are often underrepresented.

Chen et al. [10], presented XGBoost, an efficient and scalable gradient boosting framework. Widely adopted in the financial industry, XGBoost offers high performance in classification tasks such as credit scoring due to its regularization capabilities and handling of missing data.

Davis, Edelman, and Gammerman [11], assessed various machine learning algorithms for credit card applications, highlighting the potential of AI in automating financial decisions. Their study in the IMA Journal of Management Mathematics marks an early contribution to the field of credit analytics.

Duan and Keerthi [12], compared several multiclass SVM methods, identifying those best suited for practical applications. Their empirical study, presented at the International Workshop on Multiple Classifier Systems, provides insights into model selection for complex classification problems in finance.

Faggella [13], emphasized the increasing integration of neural networks and deep learning in everyday applications, including financial services. Through a conversation with Yoshua Bengio, the article discusses how such technologies are reshaping predictive analytics and personalized services.

In a follow-up article, Faggella [14], provided a comprehensive overview of machine learning, defining its principles and applications in various domains, including credit risk assessment. The piece offers an accessible yet informed perspective on how data-driven approaches are transforming traditional industries.

Friedman [15], introduced the concept of Gradient Boosting Machines (GBMs), offering a powerful framework for predictive modeling. His work, published in the Annals of Statistics, has become foundational in machine learning, including its applications in credit scoring.

In a subsequent study, Friedman [16], proposed Stochastic Gradient Boosting, which incorporates randomness into the boosting process to enhance model robustness. This technique has been instrumental in improving the generalization of credit risk models.

Hamid and Ahmed [17], developed a predictive model for loan risk using data mining techniques, demonstrating that machine learning models outperform traditional methods in accuracy and scalability. Their study appears in the Machine Learning and Applications journal.

Hsu, Chang, and Lin [18], provided a practical guide for implementing Support Vector Classification. This guide, published by National Taiwan University, remains a key resource for researchers applying SVMs in fields like credit risk prediction.

Islam [19] reported on the deteriorating state of the banking sector due to increasing bad loans. This article underscores the importance of effective risk modeling to safeguard financial stability and prevent systemic crises.

Khandani, Kim, and Lo [20], proposed consumer credit-risk models using machine learning algorithms, achieving higher predictive accuracy than traditional methods. Their work in the Journal of Banking & Finance demonstrates the transformative potential of data-driven approaches in consumer finance.

Khashman [21], compared various neural network architectures and learning schemes for credit risk evaluation. His findings in Expert Systems with Applications highlight the advantages of neural models in capturing nonlinear patterns in borrower data.

Kohavi [22], examined the effectiveness of cross-validation and bootstrap methods for model selection and accuracy estimation. His study remains influential in ensuring the reliability and generalizability of machine learning models in financial applications.

Liaw and Wiener [23], introduced the Random Forest algorithm for classification and regression tasks. As described in R News, Random Forests are highly effective in handling large, noisy datasets common in credit modeling.

Lopes et al. [24], applied machine learning to predict the recovery of credit operations in a Brazilian bank. Their research, presented at the IEEE ICMLA conference, illustrates real-world applications of AI in enhancing financial decision-making.

S. No.	Method Used	Limitation
1	Back Propagation Neural Networks	High computational cost and risk of over fitting without proper regularization.
2	Peer-to-Peer Lending with ML-based Credit Scoring	Lack of transparency in proprietary algorithms and borrower risk profiles.
3	Elastic Net Regularization	Requires careful tuning of hyper parameters; interpretability may still be limited.
4	Deep Learning (ANN, CNN) & Machine Learning (Random Forest, SVM)	Needs large datasets; black-box nature makes results hard to interpret.
5	Naive Bayes Classifier	Assumes feature independence, which is often unrealistic in financial datasets.
6	Decision Tree, SVM	Susceptible to over fitting; SVMs can be computationally expensive.
7	Missing Data Treatment Techniques (Mean, Deletion, Imputation)	Imputation can introduce bias; deletion reduces data volume.
8	Neural Scoring Model	Sensitive to input data scaling; lacks transparency in credit decisions.
9	SMOTE (Synthetic Minority Over-sampling Technique)	Risk of over fitting; may generate noisy synthetic examples.
10	XG_Boost (Gradient Boosting Algorithm)	Can overfit on noisy data; complex model with many parameters to tune.

III. Problem Domain

I. Centralized Controller Vulnerability – The SDN controller is a single point of failure, making it a prime target for attackers.

II. Lack of Standardized Detection Frameworks – Existing solutions are fragmented and lack consistency across different SDN environments.

III. Scalability and Real-Time Detection Challenges – Many detection systems struggle to perform effectively in large, dynamic, or high-speed networks.

IV. Evasion of Detection Techniques – Sophisticated attacks can bypass traditional and even some machine learning-based detection methods.

V. Emergence of New Attack Surfaces – The dynamic and programmable nature of SDN introduces novel vulnerabilities that are not yet well understood or addressed.

IV. Propose Work

There are various types of network simulating tools such as NS2, Mininet, W3, and FatTire. NS2 is used for this experiment. A normal network is created by deploying number of nodes which will act as hosts in which normal traffic is generated between the nodes and one of the intermediate nodes will be implemented with policies and rules of controller. All the messages should pass through this node so that malicious data can be detected. This node will be called the controller. An attack will be generated through an attack node which will try to exhaust the resources of the controller by flooding the network with malicious data.

In this project, we find the weak point of the SDN controller by which it can be overwhelmed when a DDoS attack happens and, propose a solution that is, specifically, tailored for SDN. Entropy is the method used in this research to detect DDoS attacks in SDN. Few parameters to DDoS detection using entropy includes; window size and a threshold value. Window size is either based on a time period or number of packets. Entropy is calculated within this window to measure uncertainty in the coming packets. A threshold value is needed to detect the attack. If the calculated entropy passes a threshold value or is below it, depending on the scheme, an attack is detected and as per specified rules the node acting as the controller will react to the attack either to forward the packet or to discard it. Here in this project we will test following scenarios and compare the outcome based on results and will prove that the security policies and rules applied on the controller works against DDOS attacks.

In the below flowchart figure 1 when the network is created a sender node will initiate communication and the message will be forwarded to the switch, the switch will look into its flow table and if present the message will be forwarded to the destination and if not present the message will be forwarded to the controller. The controller is equipped with security policies where the sender node will be scanned for malicious content. If the sender founds normal then a new entry will be made in the flow table of the controller and will be forwarded. If malicious data is found, the controller will react to the node according to the policies implemented.

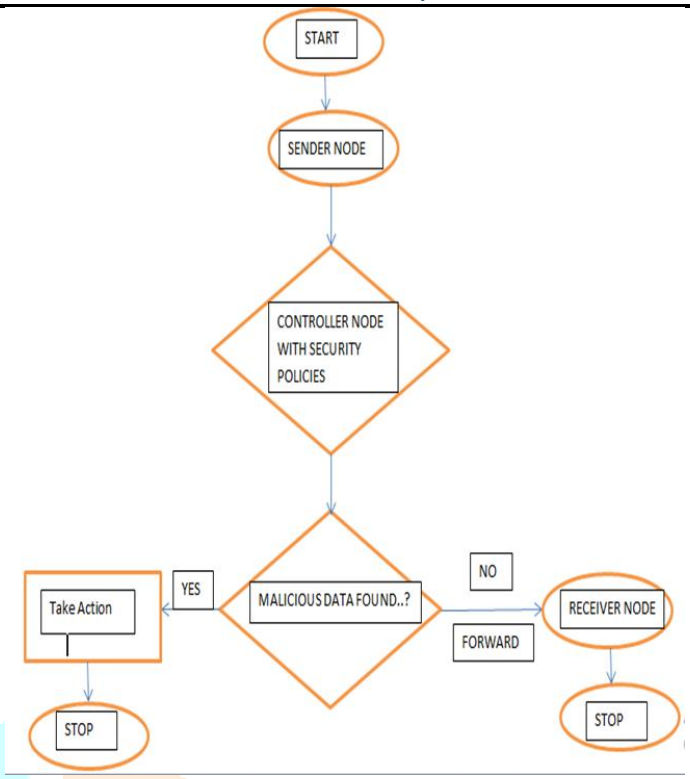

Fig 1. Architecture of system

V. Conclusion

The key component of SDN, the controller, is comparable to the operating system and needs constant security. We looked for any weaknesses that would allow the controller to get overloaded and cause the network to shut down. This study suggests a method for identifying denial-of-service (DDOS) assaults on the software-defined network controller. There are several techniques for identifying attacks. We are trying to use entropy to detect attacks. We can identify assaults on a single host or a subnet of hosts in a network by using entropy as a detection technique.

REFERENCES

[1]Neural network. https://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol4/cs11/report.
[2] Peer to peer lending and alternative investing. https://www.lendingclub.com/.
[3] Variable selection with elastic net. https://www.r-bloggers.com/ variable- selection-with-elastic-net/.
[4] Addo, P. M., Guegan, D., and Hassani, B. (2018). Credit risk analysis using machine and deep learning models. Risks, 6(2):38.
[5] Antonakis, A. and Sfakianakis, M. (2009). Assessing naive bayes as a method for screening credit applicants. Journal of applied Statistics, 36(5):537–545.
[6] Attigeri, G. V., Pai, M., and Pai, R. M. (2017). Credit risk assessment using machine learning algorithms. Advanced Science Letters, 23(4):3649–3653.
[7] Batista, G. E. and Monard, M. C. (2003). An analysis of four missing data treatment methods for supervised learning. Applied artificial intelligence, 17(5-6):519–533.
[8] Bekhet, H. A. and Eletter, S. F. K. (2014). Credit risk assessment model for jordanian commercial banks: neural scoring approach. Review of Development Finance, Universiti Tenaga Nasional (UNITEN), 43000 Kajang, Selangor, Malaysia, 4(1):20–28.
[9] Chawla, N. V., Bowyer, K. W., Hall, L. O., and Kegelmeyer, W. P. (2002). Smote: synthetic minority over-sampling technique. Journal of artificial intelligence research, 16:321–357.
[10] Chen, T., He, T., Benesty, M., et al. (2015). Xgboost: extreme gradient boosting. R package version 0.4-2, pages 1–4.
[11] DAVIS, R. H., Edelman, D., and Gammerman, A. (1992). Machine-learning algorithms for credit-card applications. IMA Journal of Management Mathematics, 4(1):43–51.

[12] Duan, K.-B. and Keerthi, S. S. (2005). Which is the best multiclass svm method? An empirical study. In International workshop on multiple classifier systems, Nanyang Technological University, Nanyang Avenue, Singapore, pages 278–285. Springer.
[13] Faggella, D. (2017). The rise of neural networks and deep learning in our everyday lives - a conversation with yoshua bengio -.
[14] Faggella, D. (2018). what is machine learning? - An informed definition. https://www.techemergence.com/what-is-machine-learning/.
[15] Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. Annals of statistics, Stanford University, USA, pages 1189–1232.
[16] Friedman, J. H. (2002). Stochastic gradient boosting. Computational Statistics & Data Analysis, Stanford University, Stanford, CA 94305, USA, 38(4):367–378.
[17]Hamid, A. J. and Ahmed, T. M. (2016). Developing prediction model of loan risk in banks using data mining. Machine Learning and Applications: An International Journal, University Khartoum, Sudan, 3(1):1–9.
[18] Hsu, C.-W., Chang, C.-C., Lin, C.-J., et al. (2003). A practical guide to support vector classification. National Taiwan University, Taipei 106, Taiwan.
[19] Islam, S. (2017). Bad loans cripple the banking sector.
[20] Khandani, A. E., Kim, A. J., and Lo, A. W. (2010). Consumer credit-risk models via machine-learning algorithms. Journal of Banking & Finance, 34(11):2767–2787.
[21] Khashman, A. (2010). Neural networks for credit risk evaluation: Investigation of different neural models and learning schemes. Expert Systems with Applications, Lefkosa, Mersin 10, Turkey, 37(9):6233–6239.
[22] Kohavi, R. et al. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. In Ijcai, Stanford University Stanford, CA., volume 14, pages 1137–1145. Montreal, Canada.
[23] Liaw, A., Wiener, M., et al. (2002). Classification and regression by random forest. R news, 2(3):18–22.
[24] Lopes, R. G., Carvalho, R. N., Ladeira, M., and Carvalho, R. S. (2016). Predicting recovery of credit operations on a brazilian bank. In Machine Learning and Applications (ICMLA), 2016 15th IEEE International Conference on, pages 780–784. IEEE.
[25] Dubey, S., Gupta, C. (2024). An Effective Model for Binary and Multi-classification Based on RFE and XGBoost Methods in Intrusion Detection System. Cyber Security and Digital Forensics. Lecture Notes in Networks and Systems, vol. 896. Springer, Singapore. https://doi.org/10.1007/978-981-99-9811-1_3.
[26] Gupta, C., Kumar, A. & Jain, N.K. (2024). An Enhanced Hybrid Intrusion Detection Based on Crow Search Analysis Optimizations and Artificial Neural Network. Wireless Personal Communication 134, 43–68. https://doi.org/10.1007/s11277-024-10880-3.
[27] Gupta, C., Kumar, A., Jain, N.K. (2023). A Detailed Analysis on Intrusion Detection Systems, Datasets, and Challenges. Advances in Data Science and Computing Technologies. ADSC 2022. Lecture Notes in Electrical Engineering, vol 1056. Springer, Singapore. https://doi.org/10.1007/978-981-99-3656-4_26.
[28] Gupta, C., Kumar, A. & Jain, N.K. Intrusion defense: Leveraging ant colony optimization for enhanced multi-optimization in network security. Peer-to-Peer Netw. Appl. 18, 98 (2025). https://doi.org/10.1007/s12083-025-01911-2.
[29] Solanki, S., Gupta, C., & Rai, K. (2020). A survey on machine learning based Intrusion Detection System on NSL-KDD dataset. Int. J. Comput. Appl, 176, 36-39.
[30] Gupta, C., Sinhal, A., Kamble, R. (2015). An "Enhanced Associative Ant Colony Optimization Technique-based Intrusion Detection System". Artificial Intelligence and Evolutionary Algorithms in Engineering Systems. Advances in Intelligent Systems and Computing, vol 325. Springer, New Delhi. https://doi.org/10.1007/978-81-322-2135-7_58.
[31] C Gupta, A Sinhal, R Kamble, "Intrusion detection based on k-means clustering and ant colony optimization: A survey", International Journal of Computer Applications, 20 Volume 79 – No 6, October 2013.
[32] Jain, T., Gupta, C. (2022). Multi-Agent Intrusion Detection System Using Sparse PSO K-Mean Clustering and Deep Learning. In: Mathur, G., Bundele, M., Lalwani, M., Paprzycki, M. (eds) Proceedings of 2nd International Conference on Artificial Intelligence: Advances and Applications. Algorithms for Intelligent Systems. Springer, Singapore. https://doi.org/10.1007/978-981-16-6332-1_10.
[33] Namdev, P., Gupta, C., Dubey, S. (2023). An Improved Intrusion Detection System Using Data Clustering and Support Vector Machine. In: Buyya, R., Misra, S., Leung, YW., Mondal, A. (eds)

Proceedings of International Conference on Advanced Communications and Machine Intelligence. MICA 2022. Studies in Autonomic, Data-driven and Industrial Computing. Springer, Singapore. https://doi.org/10.1007/978-981-99-2768-5_37.