WI-FI DATA ANALYSIS BASED ON MACHINE LEARNING

Nanumura Gedara Umal Anuraga Nanumura

University of South Wales, United Kingdom

ABSTRACT

This study proposes using machine learning to improve Wi-Fi network security. As Wi-Fi networks spread from industrial to residential areas, the necessity for strong security has risen. The rise of smart networking, especially in the IoT, has created data security and vulnerability issues. A unique method that uses machine learning to detect abnormalities and probable security breaches in Wi-Fi networks addresses these difficulties. We gather, preprocess, and analyse network data to create a complete dataset. This dataset trains machine learning algorithms to identify and classify network anomalies. Using agile methods, data mining, and machine learning algorithms, we created a Wi-Fi network intrusion detection system (WNIDS) that can detect diverse network threats. The proposed WNIDS contains two linked stages with specific machine learning models. These algorithms accurately classify network data as normal or attack specific. Our technology protects against malicious attacks and provides a robust Wi-Fi network for users across domains by incorporating machine learning. Modern network security dangers were fully understood by surveys and data analysis. The WNIDS was implemented and deployed through a structured system development life cycle. This tool eliminates network weaknesses and advances distant enterprises, offering safe and smooth access for consumers globally.

KEYWORDS

Machine learning, Wi-Fi network security, data security, vulnerability, dataset.

1. INTRODUCTION

The way in which people engage with technology has been fundamentally altered as a result of the widespread availability of mobile computing devices, in particular smartphones, and the rising processing capacity of these devices [1]. These gadgets have become indispensable to day-to-day life since they make possible a diverse array of activities such as communication, the exchange of documents, video streaming, and many more. The proliferation of cloud-based services has added to the already enormous volume of Internet traffic, which has led to projections of a significant rise in the amount of data that is transmitted over the Internet [2]. Wi-Fi networks, which account for a sizeable percentage of this expansion, have developed to meet the need for connectivity posed by portable devices [3].

However, the proliferation of Wi-Fi networks has also resulted in an increase in worries over their level of security. Because of the broad use of these networks, they have grown susceptible to unwanted access and infiltration as a result. Instances of network breaches, illegal control of public Wi-Fi networks, and flaws in encryption techniques such as WPA2 have brought to light the essential requirement for highly effective security procedures [4]. It is of the utmost importance, not only for individual users but also for sectors that are dependent on network connectivity, to protect the privacy, authenticity, and accessibility of the data that is transmitted through Wi-Fi networks [5].

The deployment of methods that utilize machine learning has become more popular as a means of addressing these security concerns. To identify unusual occurrences and potential security risks, machine learning analyzes massive volumes of data for recurring patterns. These methods, when applied to Wi-Fi networks, have the ability to recognize odd network behavior, differentiate between genuine and malicious access points, and increase overall network security [6].

The purpose of this study is to investigate how the power of machine learning may be utilized to strengthen the security of Wi-Fi networks. This study makes a contribution to the protection of both human and corporate data by detecting illegal access, identifying hostile network activity, and proactively addressing possible threats. These are all ways in which the data might be compromised. The technique that has been suggested comprises the gathering and examination of data from Wi-Fi networks, the training of machine learning models, and the implementation of an intelligent Intrusion Detection System (IDS) in order to detect and prevent possible security breaches [7].

In the following study, we will dig into the complexities of the landscape of Wi-Fi security, address the issues faced by unauthorized access and harmful actions, and suggest a solution to strengthen network security that is powered by machine learning. Collecting data, training models, putting them into practice, and conducting evaluations are all part of the study that is being done to prove that the suggested method is beneficial. In the end, the purpose of this study is to make a contribution to the development of Wi-Fi network security and to ensure that the use of networks is both secure and dependable for individuals as well as for industries [8].

2. FORMAT GUIDE

2.1. Deep Learning-based Intrusion Detection for Wi-Fi Networks

This study focuses on using deep learning methods, more especially Convolutional Neural Networks (CNNs), to the problem of identifying malicious activity in wireless local area networks (Wi-Fi). The purpose of this project is to investigate the use of CNNs to evaluate network traffic patterns and categorize them as normal or abnormal, with the end goal of improving the safety of Wi-Fi networks [9].

2.2. Machine Learning Techniques for Wi-Fi Intrusion Detection

This study provides a detailed assessment of the many different machine learning approaches that are deployed for the purpose of detecting Wi-Fi intrusions. The purpose of this study is to explore the efficacy of several algorithms, including Random Forest, Support Vector Machines (SVM), and Neural Networks, when it comes to detecting abnormalities in networks and improving network security [10].

The template is used to format your paper and style the text. All margins, column widths, line spaces, and text fonts are prescribed; please do not alter them. You may note peculiarities. For example, the head margin in this template measures proportionately more than is customary. This measurement and others are deliberate, using specifications that anticipate your paper as one part of the entire proceedings, and not as an independent document. Please do not revise any of the current designations [12].

2.3. Wi-Fi Intrusion Detection Using Machine Learning Techniques

The authors of this study investigate whether or not the use of machine learning techniques is beneficial in identifying malicious activity on Wi-Fi networks. They experiment with classifiers such as Naive Bayes, Decision Tree, and k-Nearest Neighbours in order to differentiate between normal and invasive patterns in the network data [11].

2.4. A Comparative Study of Machine Learning Techniques for Wi-Fi Intrusion Detection

The template is used to format your paper and style the text. All margins, column widths, line spaces, and text fonts are prescribed; please do not alter them. You may note peculiarities. For example, the head margin in this template measures proportionately more than is customary. This measurement and others are deliberate, using specifications that anticipate your paper as one part of the entire proceedings, and not as an independent document. Please do not revise any of the current designations [12].

2.5. Wi-Fi Intrusion Detection System using Machine Learning Algorithms

The purpose of this study is to offer an intrusion detection system (IDS) architecture that makes use of machine learning methods to identify illegal access and intrusions into Wi-Fi networks. In this study, various intrusion detection algorithms, such as Random Forest, SVM, and k-Nearest neighbours, are put through their paces to see how well they function [13].

2.6. Anomaly Detection in Wi-Fi Networks Using Machine Learning

In this research, we investigate the use of machine learning approaches designed for the goal of detecting unusual behaviour in Wi-Fi networks. The study makes use of autoencoders, a type of neural network, to learn the usual behaviour of the network and notice any deviations from this pattern that may be indicators of possible intrusions or attacks. The goal of the study is to improve the network's ability to detect and prevent attacks and intrusions [14].

2.7. Machine Learning Techniques for Wireless Network Security: A Comprehensive Survey

This study presents an overview of machine learning approaches that may be used to the security of wireless networks. Although the survey does not primarily focus on Wi-Fi networks. It addresses detection of anomalies and intrusions, as well as other areas of network security that are pertinent to Wi-Fi networks [15].

2.8. Network Intrusion Detection System using Machine Learning for Wi-Fi Networks

Making use of several machine learning approaches, the primary aim of this investigation is the development of an Intrusion Detection System (IDS) for Wi-Fi networks. The research investigates the efficacy of several intrusion detection algorithms, including Decision Trees, SVM, and Naive Bayes, among others [16].

2.9. Deep Learning Approach for Wi-Fi Intrusion Detection

A technique based on deep learning is proposed by the authors as a means of detecting Wi-Fi intrusions. They use Long Short-Term Memory networks to capture sequential patterns in network traffic and identify aberrant behaviour, hence improving network security. This allows them to spot potential threats to the network [17].

2.10. Deep Learning Approach for Wi-Fi Intrusion Detection

A technique based on deep learning is proposed by the authors as a means of detecting Wi-Fi intrusions. They use Long Short-Term Memory networks to capture sequential patterns in network traffic and identify aberrant behavior, hence improving network security. This allows them to spot potential threats to the network [17].

2.11. Wi-Fi Intrusion Detection Using Recurrent Neural Networks

This study investigates the use of Recurrent Neural Networks (RNNs), a type of artificial neural network, to identify Wi-Fi intrusions. The purpose of this study is to evaluate the capacity of RNNs to perform time-series network data analysis and recognize abnormalities that point to the presence of possible intrusions [18].

2.12. A Hybrid Machine Learning Approach for Wi-Fi Intrusion Detection

In this research, a hybrid machine learning strategy to detecting Wi-Fi infiltration is presented. This approach includes several distinct algorithms, including SVM and Genetic Algorithms, among others. The research reveals that the use of many security measures together may significantly improve network safety [19].

2.13. Enhancing Wi-Fi Network Security through Machine Learning-based Anomaly Detection

To improve the safety of Wi-Fi networks, Choudhary and Gupta have proposed an anomaly detection system that makes use of machine learning. The purpose of the study is to find unexpected patterns in the behaviour of networks so that possible vulnerabilities in network security may be discovered earlier [20].

2.14. Federated Learning for Privacy-Preserving Wi-Fi Intrusion Detection

This research proposes a method for the detection of Wi-Fi intrusions that takes use of federated learning, protects the privacy of Wi-Fi users, and does not affect the security of the network. This project aims to examine how machine learning models may be taught in a collaborative manner across several devices, without compromising the privacy of individual users. Specifically, the research will investigate how this can be done [21].

2.15. Figures and Tables

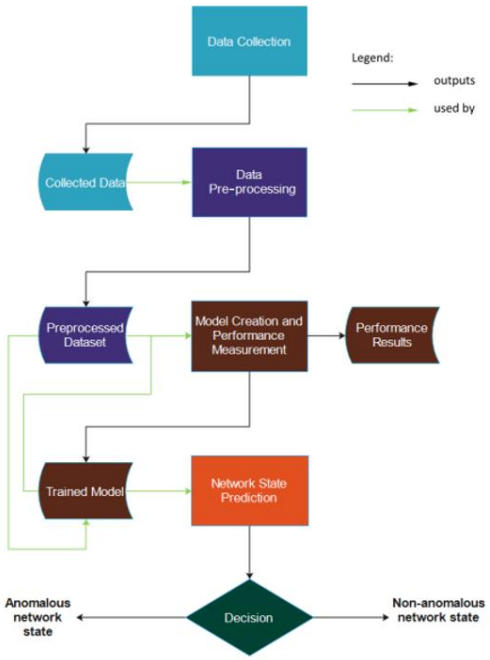

Figure 1. Desire model of SNMP process

3. REQUIREMENTS AND METHODOLOGY

3.1. Requirements

3.1.1. Essential Requirements for Software

- The program known as Jupyter Notebook.
- The infrastructure for wireless networks.
- Access points (AP) in addition to antennas.
- The gadget that will be the focus of the network's surveillance.

3.1.2. Minimum Requirements for the Hardware

- Requirements for the barest minimum of a system.
- Hardware for networks, including routers.

3.1.3. Essential Functional Requirements

- The collection of data of a high quality for use in training models
- Analysis of networks and classification of different kinds of networks
- The identification of suspicious data traffics as well as possible efforts at hacking
- The detection of fake networks that are connected to real networks.
- A Taxonomy of the Different Kinds of Network Attacks

3.1.4. Non-functional Requirements

- An intuitive and easy-to-use interface within Jupyter Notebook to facilitate its use
- Capable management of both guided and unsupervised procedures in a productive manner
- A high degree of accuracy obtained through the use of cross-validation
- Resilience was achieved by the use of a cross-validation data set function.

3.2. Methodology

1) Data Collection: NMP is utilized during the first step of the process in order to collect data from the wireless network. In order to provide information about devices and networks at regular intervals of 15 seconds, the SNMP protocol is utilized. The obtained data values come from the activities that take place on wireless networks. These activities include a variety of characteristics, such as Clear to Send (CTS) and Request to Send (RTS) codes, as well as CPU utilization and Media Access Control (MAC) addresses [22]. In order to diagnose problems with the transmission of data packets, CTS and RTS are utilized, while other characteristics offer insightful information on the network.

2) The Preprocessing of the Data: In order to get the gathered SNMP data ready for further analysis, a preprocessing step is performed on them. The information is then converted into an asymmetric matrix format that can be read by machine learning techniques. Addressing connection issues, maintaining counter signatures, and handling incomplete SNMP signatures are some of the tasks included in preprocessing. The end product is a matrix consisting of real-world values that are arranged in a manner that is consistent with the timestamp and SNMP signature type.

3) Model Construction and the Evaluation of Its Performance: During the process of model construction, machine learning algorithms are incorporated for the purpose of anomaly detection. Deep neural networks, support vector machines (SVMs), and one-class support vector machines are the primary methods that are investigated. In order to attain the best possible results from the model, hyperparameter tweaking is performed via fivefold cross-validation. To evaluate and evaluate and compare the performance of the different models, metrics such as accuracy, recall, F1-score, and Area Under the Curve (AUC) are utilized. The chosen models will be kept for possible usage in the future.

4) Identification of the Current State of the Network: A real-time network state identification system is included as the last component, which also incorporates the learned models. SNMP data is standardized by using Jupyter Notebook, and the mean and standard deviation of the training set are used as the basis for the standardization. The selected model computes the current state of the network based on each SNMP data input and outputs the results not only through the command-line interface but also via a text file. Repeating this procedure until certain requirements are satisfied ensures that irregularities in the network will be discovered in a timely manner.

This research makes use of a methodology that incorporates a number of essential stages to improve network security by utilizing SNMP data analysis and other machine learning strategies [23]. Utilizing Simple Network Management Protocol (SNMP) to collect information about devices and networks at 15-second intervals, including characteristics such as CTS, RTS, MAC addresses, and CPU utilization, is required for data gathering [24]. After that, the data that was collected undergoes preprocessing, which results in the creation of an asymmetric matrix format that may be used by machine learning algorithms. The building of a model requires the use of

deep neural networks, support vector machines, and one-class support vector machines (one-class SVMs), followed by the tuning of hyperparameters and the monitoring of performance using metrics like as accuracy and recall. The completed system incorporates these models in order to identify the current state of the network in real time, while also standardizing SNMP data and using the model that was selected in order to categorize the various network states. In order to give a thorough foundation for the method that has been presented, the requirements for software, hardware, system functions, and non-functional components have been outlined.

4. CONCLUSIONS

In the course of this research endeavour, we set out to conduct an exhaustive investigation of the process of detecting anomalies in wireless networks by utilizing machine learning strategies. Our investigation addressed the crucial need for the rapid discovery and mitigation of network anomalies, which are pivotal in guaranteeing uninterrupted and secure wireless communication. This requirement was brought to light by the fact that our investigation was conducted. We came at a number of significant conclusions that shed light on the potential, limits, and future prospects of utilizing machine learning in this area after conducting an in-depth study, conducting experiments, and analysing the results of those experiments.

The assessment of several machine learning algorithms was the central focus of our investigation. In particular, we focused on deep neural networks (DNN), supervised single-class support vector machines (SVM), and supervised multi-class SVM. According to the findings of our comparison study, DNN, with its ingrained capacity to recognize complicated patterns, is a viable route for network anomaly detection. On the other hand, it was clear that the high level of skill came at the expense of a lengthy and labour-intensive training procedure. On the other hand, SVM models demonstrated impressive durability and consistency over a wide variety of network settings, which exemplified their usefulness in practical applications due to their ability to accurately predict network behaviour.

The conception and execution of a real-time anomaly detection system is the fruition of all of our labour and represents the zenith of our efforts. This approach, which was encased in a Jupyter Notebook dashboard that was easy to use, was able to successfully catch network discrepancies in the early phases of their development. We anticipate a significant impact on network quality of service and security as a result of providing network managers with insights that can be put into action. Despite the fact that the findings are rather convincing, we are aware that the system is dependent on the currently available dataset and that it may be susceptible to new types of network paradigms.

While we are aware of the constraints that our study places on us, we are also aware of the boundless possibilities that lie ahead for further investigation. The incorporation of more powerful machine learning algorithms, the enlargement of the dataset to include a wider variety of network settings, and the investigation of various approaches to risk reduction are all potential directions for future research. Our study lays a foundation, paving the way for following researchers to look further into the complex dynamic that exists between machine learning and wireless network management. By presenting actual proof of the efficacy of machine learning strategies, our study makes a significant contribution to the field of wireless network anomaly detection. The findings of our analysis have the potential to bolster wireless networks against threats that were not anticipated, as well as to strengthen the pillars of security and connection that support our contemporary world. These developments are the result of the ongoing transformation of the digital landscape.

ACKNOWLEDGEMENTS

Thank you for everyone who help me to success this research.

REFERENCES

[1] S. F. Ng, N. S. I. Che-Hassan, N. H. Mohammad-Nor, and N. A. Abdul-Malek, "The Relationship between Smartphone Use, Symptoms of Depression, Symptoms of Anxiety, and Academic Performance in College Students," Malaysian Online J. Educ. Technol., vol. 5, no. 4, p. 72, 2017, [Online]. Available: http://www.mojet.net/frontend//articles/pdf/v5i4/v05i04-05pdf.pdf.

[2] S. Kumar, P. Tiwari, and M. Zymbler, "Internet of Things is a revolutionary approach for future technology enhancement: a review," J. Big Data, vol. 6, no. 1, pp. 1–21, Dec. 2019, doi: 10.1186/S40537-019-0268-2/FIGURES/9.

[3] T. Cisco and A. Internet, "Cisco: 2020 CISO Benchmark Report," Comput. Fraud Secur., vol. 2020, no. 3, pp. 4–4, 2020, doi: 10.1016/s1361-3723(20)30026-9.

[4] J. Jang-Jaccard and S. Nepal, "A survey of emerging threats in cybersecurity," J. Comput. Syst. Sci., vol. 80, no. 5, pp. 973–993, Aug. 2014, doi: 10.1016/J.JCSS.2014.02.005.

[5] B. L. Filkins et al., "Privacy and security in the era of digital health: What should translational researchers know and do about it?," Am. J. Transl. Res., vol. 8, no. 3, pp. 1560–1580, 2016.

[6] I. H. Sarker, "Machine Learning for Intelligent Data Analysis and Automation in Cybersecurity: Current and Future Prospects," Ann. Data Sci., pp. 1–26, Sep. 2022, doi: 10.1007/S40745-022-00444-2/FIGURES/6.

[7] A. F. J. Jasim and S. Kurnaz, "New automatic (IDS) in IoTs with artificial intelligence technique," Optik (Stuttg)., vol. 273, p. 170417, 2023, doi: https://doi.org/10.1016/j.ijleo.2022.170417.

[8] R. Nazir, A. A. laghari, K. Kumar, S. David, and M. Ali, "Survey on Wireless Network Security," Arch. Comput. Methods Eng., vol. 29, no. 3, pp. 1591–1610, May 2022, doi: 10.1007/S11831-021-09631-5.

[9] M. Natkaniec and M. Bednarz, "Wireless Local Area Networks Threat Detection Using 1D-CNN," Sensors, vol. 23, no. 12, p. 5507, 2023, doi: 10.3390/s23125507.

[10] T. Sowmya and E. A. Mary Anita, "A comprehensive review of AI based intrusion detection system," Meas. Sensors, vol. 28, p. 100827, Aug. 2023, doi: 10.1016/J.MEASEN.2023.100827.

[11] M. M. Rashid, J. Kamruzzaman, M. M. Hassan, T. Imam, and S. Gordon, "Cyberattacks detection in iot-based smart city applications using machine learning techniques," Int. J. Environ. Res. Public Health, vol. 17, no. 24, pp. 1–21, 2020, doi: 10.3390/ijerph17249347.

[12] Y. Kayode Saheed, A. Idris Abiodun, S. Misra, M. Kristiansen Holone, and R. Colomo-Palacios, "A machine learning-based intrusion detection for detecting internet of things network attacks," Alexandria Eng. J., vol. 61, no. 12, pp. 9395–9409, Dec. 2022, doi: 10.1016/J.AEJ.2022.02.063.

[13] P. Vanin et al., "A Study of Network Intrusion Detection Systems Using Artificial Intelligence/Machine Learning," Appl. Sci., vol. 12, no. 22, 2022, doi: 10.3390/app122211752.

[14] Y. Song, S. Hyun, and Y. G. Cheong, "Analysis of autoencoders for network intrusion detection†," Sensors, vol. 21, no. 13, pp. 1–23, 2021, doi: 10.3390/s21134294.

[15] T. Sowmya and E. A. Mary Anita, "A comprehensive review of AI based intrusion detection system," Meas. Sensors, vol. 28, p. 100827, 2023, doi: https://doi.org/10.1016/j.measen.2023.100827.

[16] H. Premakumar, A. Kunarathinam, and V. Ravi, "Detection of Intrusions in networks using Machine Learning Approach," pp. 1–7, 2022.

[17] E. Ul, H. Qazi, M. H. Faheem, and T. Zia, "HDLNIDS: Hybrid Deep-Learning-Based Network Intrusion Detection System," Appl. Sci. 2023, Vol. 13, Page 4921, vol. 13, no. 8, p. 4921, Apr. 2023, doi: 10.3390/APP13084921.

[18] J. J. Montaño Moreno, A. Palmer Pol, and P. Muñoz Gracia, "Artificial neural networks applied to forecasting time series.," Psicothema, vol. 23, no. 2, pp. 322–329, Apr. 2011.

[19] İ. Yazici, I. Shayea, and J. Din, "A survey of applications of artificial intelligence and machine learning in future mobile networks-enabled systems," Eng. Sci. Technol. an Int. J., vol. 44, p. 101455, 2023, doi: https://doi.org/10.1016/j.jestch.2023.101455.

[20] R. Kaur, D. Gabrijelčič, and T. Klobučar, "Artificial intelligence for cybersecurity: Literature review and future research directions," Inf. Fusion, vol. 97, p. 101804, 2023, doi: https://doi.org/10.1016/j.inffus.2023.101804.

[21] V. Rey, P. M. Sánchez Sánchez, A. Huertas Celdrán, and G. Bovet, "Federated learning for malware detection in IoT devices," Comput. Networks, vol. 204, p. 108693, 2022, doi: https://doi.org/10.1016/j.comnet.2021.108693.

[22] K. Bicakci and B. Tavli, "Denial-of-Service attacks and countermeasures in IEEE 802.11 wireless networks," Comput. Stand. Interfaces, vol. 31, no. 5, pp. 931–941, Sep. 2009, doi: 10.1016/j.csi.2008.09.038.

[23] M. Wang, G. Song, Y. Yu, and B. Zhang, "The Current Research Status of AI-Based Network Security Situational Awareness," Electron., vol. 12, no. 10, 2023, doi: 10.3390/electronics12102309.

[24] Huawei, "display mac-address - Fat AP and Cloud AP V200R010C00 Command Reference - Huawei," 2023. https://support.huawei.com/enterprise/en/doc/EDOC1100064352/25e69c90/display-mac-address (accessed Aug. 13, 2023).

[25] A. T. Chala and R. Ray, "Assessing the Performance of Machine Learning Algorithms for Soil Classification Using Cone Penetration Test Data," Appl. Sci., vol. 13, no. 9, 2023, doi: 10.3390/app13095758.