International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048
Volume 7, Issue 10, pp 249-257, October-2023
https://doi.org/10.47001/IRJIET/2023.710032

Realtime Network Based Anomaly Detection and Malware Analysis for SMEs and Smart Homes

[1]K.N.H De Silva, [2]M.A.S.B Manchanayaka, [3]D.L.S.I Punyasiri, [4]H.A.D.N Perera, [5]Anjalie Gamage, [6]Narmada Gamage

[1,2,3,4,5,6]Department of Computer Systems Engineering, Sri Lanka Institute of Information Technology, New Kandy Rd, Malabe 10115, Sri Lanka
Authors E-mail: [1]it20232504@my.sliit.lk, [2]it20194130@my.sliit.lk, [3]it20147228@my.sliit.lk, [4]it17179218@my.sliit.lk, [5]anjalie.g@slitt.lk, [6]narmada.g@sliit.lk

***Abstract* - The increasing risk landscape of cyberattacks requires the implementation of strong solutions for detecting anomalies in real-time within network systems and analyzing malware. These solutions should be specifically designed to cater to the needs of Small and Medium-sized Enterprises (SMEs) and smart homes. This study presents a comprehensive framework to effectively address the pressing security concern at hand. To begin with, a cost-effective and advanced firewall architecture, based on Raspberry Pi, is implemented in order to protect connected devices from external malicious entities. Simultaneously, this architecture captures network traffic for further analysis. By utilizing supervised machine learning models, such as Random Forests, a system has been developed to detect anomalies in Internet of Things (IoT) environments. This system enables timely notifications and facilitates informative discussions regarding the identified irregularities. In order to enhance the ability to detect network anomalies, a unique methodology is implemented, which involves the utilization of Natural Language Processing (NLP). This approach is complemented by the strategic deployment of honeypots to confuse potential attackers and is further supported by a collaborative infrastructure for sharing information on threats. Moreover, this study conducts an analysis of malware using both signature-based and behavior-based approaches. These methodologies are facilitated by supervised machine learning models, which are seamlessly incorporated with an alerting mechanism to ensure prompt notifications. This research study makes a significant contribution to the field by addressing existing gaps in knowledge, such as comparative assessments, zero-day vulnerabilities, user-centric design, and cost-effectiveness considerations. Additionally, it provides a practical guide for enhancing cyber resilience in small and medium-sized enterprises (SMEs) and smart homes.**

Keywords: Internet of Things (IoT), Natural Language Processing (NLP), Small and Medium-sized Enterprises (SMEs), Raspberry-Pi, Firewall, Machine Learning Module.

I. INTRODUCTION

In today's interconnected and digitally transformed world, securing computer networks is of utmost importance. This issue encompasses SMEs and smart homes across different domains. The increasing number of connected devices and growing cyber threats highlight the need for proactive measures to protect sensitive data and maintain uninterrupted operations [1]. Real-time network-based anomaly detection and malware analysis solutions are promising defenses against evolving cyber risks. This research aims to contribute significantly to this critical field by introducing a comprehensive framework specifically designed for SMEs and smart homes. This framework fills the gap in efficient and user-friendly solutions in these environments [2].

The main goal of this study is to create, apply, and assess a new method for promptly identifying and analyzing network anomalies and malicious software infiltrations. This methodology aims to benefit SMEs and smart homes. This research aims to address challenges faced by resource-constrained entities in combating modern cyber threats by incorporating advanced technologies and methods. The objective is to create a robust, affordable, and flexible solution that overcomes the drawbacks of conventional security measures.

The proposed methodology effectively tackles major obstacles in network security. This paper presents a new and cost-effective firewall architecture that utilizes the features of Raspberry Pi. The first step of developing a comprehensive security framework entails the implementation of a sophisticated firewall system. The firewall has the dual purpose of enhancing protection against external threats and serving as a centralized hub for the effective administration and analysis of network traffic. By implementing this measure, small and medium-sized enterprises (SMEs) and smart households can enhance their security stance while also maximizing the efficient use of existing resources.

Anomaly identification is crucial in the context of the Internet of Things (IoT) [3]. The second step in the

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 10, pp 249-257, October-2023

https://doi.org/10.47001/IRJIET/2023.710032

implementation is to deploy a specialized anomaly detection system for IoT environments. This system uses supervised machine learning models, such as Random Forests, to effectively analyze network data and detect patterns that indicate real-time malicious activities [4]. Additionally, it enables smooth notification transmission for identified irregularities, promoting a collaborative approach to maintain network integrity.

The third step entails the incorporation of natural language processing (NLP) tools to enhance the scope of anomaly detection. The proposed methodology in this study aims to improve the accuracy of anomaly detection by implementing strategic honeypot deployments, which are designed to confuse and deter possible malicious actors [5]. Moreover, the establishment of a threat-sharing system facilitates the cultivation of joint endeavors aimed at addressing dynamic threats, hence enhancing the robustness of network cybersecurity.

The fourth step focuses on malware analysis, an essential aspect of network security. This study uses both signature-based and behavior-based methods [6]. Using supervised machine learning models, the system can detect known malicious software patterns and accurately identify abnormal behaviors that suggest zero-day attacks. The addition of a malware alert system improves these capabilities by providing prompt notifications and appropriate countermeasures.

This study aims to improve real-time network-based anomaly detection and malware analysis for small and medium-sized enterprises (SMEs) and smart homes. The research aims to develop a comprehensive defense strategy by combining cost-effective and advanced technological solutions, including machine learning, natural language processing (NLP), and collaborative frameworks. The following sections of this research paper provide a detailed analysis of the different components, implementation details, assessments, and evaluation of the proposed approach.

II. BACKGROUND AND LITERATURE REVIEW

In the ever-changing world of interconnected devices and digital transformation, protecting computer networks is crucial for various sectors, including SMEs and smart homes [7]. The increasing use of interconnected devices and the rise in cyber threats highlight the urgent need for proactive measures to protect sensitive data and maintain smooth operations. Real-time network-based anomaly detection and malware analysis solutions are effective in addressing the constantly evolving cyber risks [8]. This research aims to make a significant contribution to this important field by presenting a comprehensive framework specifically designed for the unique contexts of SMEs and smart homes. This framework aims to address the need for efficient and user-friendly solutions designed for these unique environments.

The increasing connectivity of devices in the digital age has changed how we understand and address cybersecurity. As organizations and residences increasingly rely on networked systems for critical functions, the risk of cyber threats becomes a significant and urgent concern. Small and medium-sized enterprises (SMEs) and smart homes face unique challenges in securing their networks [9]. These challenges arise due to resource and technical expertise limitations. This study explores real-time network-based anomaly detection and malware analysis, specifically targeting SMEs and smart homes. The goal is to fill the current gap in proactive cybersecurity measures for these important sectors.

In order to fully grasp the importance of real-time network-based anomaly detection and malware analysis, it is imperative to acknowledge the ever-changing nature of the threat landscape [10]. Cybercriminals consistently develop advanced techniques to exploit vulnerabilities, rendering conventional reactive security measures insufficient [11]. The proactive nature of real-time detection facilitates the timely identification of malicious activities as they occur, thereby enabling prompt implementation of countermeasures. According to a study conducted by Yuchong Li (2021), the implementation of real-time detection systems has been shown to effectively decrease the time it takes to detect and respond to incidents, thereby minimizing the potential for harm [12].

Machine learning (ML) has emerged as a highly effective tool in the field of network security, specifically in the areas of anomaly detection and malware analysis [13]. Machine learning algorithms, such as Random Forests, provide the capability to analyze vast datasets and detect complex patterns that are indicative of malicious activities [14]. The study conducted by Yu-kyung Kim (2022) showcased the effectiveness of machine learning (ML) in identifying previously unidentified security risks by utilizing pattern recognition techniques [15]. This research successfully tackled the issue of zero-day attacks. Nevertheless, despite the considerable potential of machine learning (ML), the crucial aspect that needs to be addressed is its effective integration for small and medium-sized enterprises (SMEs) and individuals without technical expertise.

To expand anomaly detection, the third component incorporates advancements from natural language processing (NLP) [16]. This methodology enhances anomaly detection accuracy and strategically deploys honeypots to mislead potential attackers. A threat sharing framework is integrated to encourage a community-driven response and enhance the network's cyber resilience.

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 10, pp 249-257, October-2023

https://doi.org/10.47001/IRJIET/2023.710032

The growing use of the Internet of Things (IoT) requires more attention on anomaly detection, the second aspect of this implementation [4]. By using a specialized system for IoT environments and supervised machine learning models like Random Forests, real-time analysis of network data can be achieved. This system efficiently detects patterns indicating malicious activities, allowing for prompt implementation of countermeasures. Furthermore, the efficient dissemination of notifications regarding identified irregularities fosters collaborative efforts to protect the network's integrity.

The emergence of the Internet of Things (IoT) has introduced a heightened level of intricacy to the realm of network security. Internet of Things (IoT) devices present numerous vulnerabilities that can be exploited by malicious actors, underscoring the critical importance of implementing effective anomaly detection mechanisms [11]. The significance of customized anomaly detection systems in Internet of Things (IoT) settings. The incorporation of natural language processing (NLP) techniques presents a new aspect to network security. By leveraging the capabilities of Natural Language Processing (NLP), systems have the ability to not only improve the accuracy of anomaly detection but also deceive potential attackers by strategically deploying honeypots [17].

The fourth aspect involves malware analysis, which is crucial for network security. This study utilizes a comprehensive methodology that integrates signature-based and behavior-based approaches [18]. This research uses supervised machine learning models to achieve its objectives. It identifies known malware signatures and detects unusual behaviors seen in zero-day attacks. The addition of a malware alert system improves this capability by providing timely notifications and appropriate countermeasures.

There is a significant research gap in comparing detection methods. Signature-based detection, typically used to identify known malware, differs in adaptability compared to machine learning and behavioral analysis, particularly in detecting new threats [19]. This analysis is crucial for assessing the effectiveness of different techniques in different contexts. It is important to also consider cost-effective solutions for small and medium-sized enterprises (SMEs) and smart homes. Regrettably, there is a lack of thorough studies assessing the cost-effectiveness of real-time detection systems tailored for these industries.

This research aims to address gaps in real-time network-based anomaly detection and malware analysis. It focuses on meeting the needs of SMEs and smart homes. The goal is to create a robust defense system by combining cost-effective and advanced technologies such as machine learning, natural language processing, and collaborative frameworks. The next sections of this research project will focus on examining the elements, implementation details, evaluations, and effectiveness of the proposed approach.

III. METHODOLOGY

Addressing research gaps requires a deliberate and comprehensive strategy. Design a Raspberry Pi-based firewall for traffic analysis. Supervised machine learning, particularly Random Forests, detects anomalies in Internet of Things (IoT) networks and alerts relevant parties about security issues. Natural language processing and honeypots help the system discover and stop intruders.

Signature-based and behavior-based malware analysis using supervised machine learning algorithms is also included. This framework improves zero-day attack detection and speed. Integrating a collaborative threat-sharing system allows community-driven responses to emerging dangers. An easy-to-use interface prioritizes usability and user experience. An economic analysis determines the solution's cost-effectiveness for SMEs and smart homes. An innovative and practical system for real-time anomaly detection and malware analysis in network environments, especially for SMEs and smart homes, accomplishes study objectives.

A) System Overview

A Raspberry Pi bridges the router and the user's internal network. This strategic stance controls and secures data movement between internal and external domains. A Raspberry Pi software-based firewall is methodically developed for security and traffic management. This firewall system carefully inspects incoming and outgoing data packets to protect the system.

The architectural framework enhances its functionalities by incorporating various specialized subsystems, each playing a crucial role in bolstering the overall security stance:

- Internet of Things (IoT) Anomaly Detection using Machine Learning
- Network Anomaly Detection utilizing Machine Learning
- Detection of External Malicious URLs through Machine Learning
- Signature and Behavior-Based Malware Detection using Machine Learning

The machine learning models are stored using the joblib library for easy deployment in real-time situations. And importing essential libraries such as matplotlib, pandas, seaborn, joblib, and numpy.

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 10, pp 249-257, October-2023

https://doi.org/10.47001/IRJIET/2023.710032

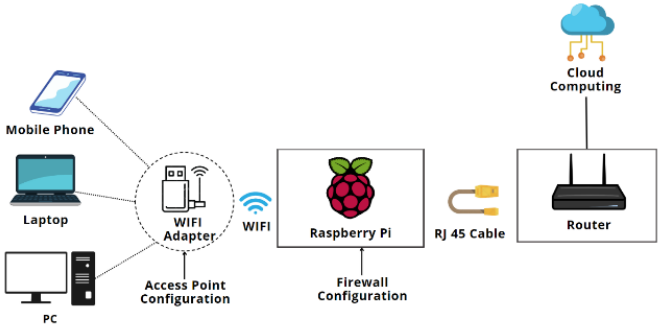

Figure 1: System Overview

These specialized modules for machine learning are designed to function on cloud-based servers, taking advantage of the scalability and computational capabilities provided by the cloud infrastructure. The database synchronization enables smooth communication between the Raspberry Pi device and the cloud servers. This communication protocol facilitates the efficient transfer of data, thereby enabling the processing and analysis of complex datasets.

B) Functionalities

1) Firewall and External Malicious URLs Detection

This methodology centers around using a repurposed Raspberry Pi device as the central hub for network security. The device has a software firewall called iptables firewall, which provides strong security features and there is GUI for manage this firewall name called easywall. Integrating iptables firewall on the Raspberry Pi is a significant advancement, providing the system with improved security and management capabilities.

The Raspberry Pi device is configured as an access point to optimize network traffic and enhance user satisfaction. This maneuver improves the connection between the user network and the router, enhancing data flow and maintaining strong security measures. The configuration of this access point improves the efficiency of the network anomaly detection and malware analysis system [19].

This implementation stands out for its use of cloud-hosted machine learning modules. These modules, hosted on remote servers, enable the system to identify potentially harmful URLs. The firewall analyzes each URL as it moves through the network and sends it to the machine learning module for analysis. Detecting a malicious URL requires immediate action. A Python script is used to connect with the firewall and block the malicious URL.

The combination of these elements creates a complex and carefully designed security system. The network's integrity is improved by using a Raspberry Pi-powered firewall system fortified with iptables firewall. Simultaneously, cloud-hosted machine learning modules enhance the system's real-time threat detection and prevention capabilities, particularly in identifying malicious URLs. This methodology combines advanced technologies including Raspberry Pi, iptables firewall software (easywall), cloud servers, SQLlite SB, and custom machine learning modules.

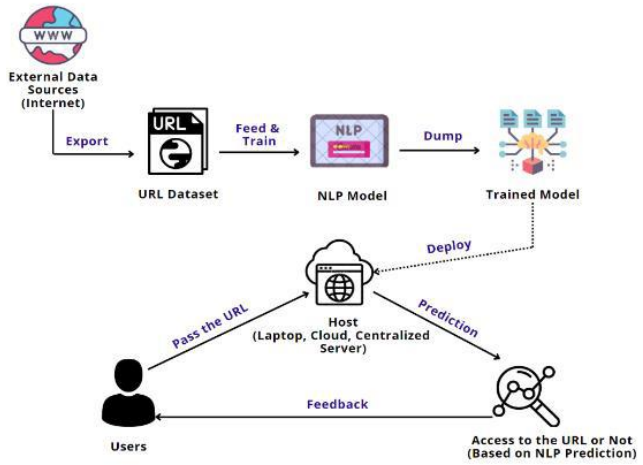

Figure 2: External URL Detection using ML

The diagram shows how advanced Natural Language Processing (NLP) techniques are used to identify external malicious URLs. The objective is to predict the risk level of these URLs by analyzing their malicious attributes. This research project collects datasets from Kaggle, including legitimate URLs, phishing URLs, and malicious URLs. Before being inputted into a Random Forest Regression algorithm for training, the data is subjected to feature engineering and preprocessing using NLP methodologies. The algorithm is highly proficient in handling complex interactions with text data and severity labels.

The Bag of Words model is used to convert preprocessed text into numerical format. The CountVectorizer tool from scikit-learn library is used to apply different machine learning algorithms to the transformed data.

2) Netowrk Anomaly Detection & Honeypot Setup

To achieve this goal, we utilize a comprehensive approach that includes cutting-edge technologies, advanced methodologies, and robust libraries. The methodology uses natural language processing (NLP) techniques to enhance anomaly detection precision and effectiveness. To conduct a comprehensive analysis, we capture network packets in real-time using the widely recognized 'tcpdump' tool. This method ensures obtaining a complete and up-to-date dataset. The data capture process is automated through a carefully designed cronjob, ensuring a seamless workflow.

International Research Journal of Innovations in Engineering and Technology (IRJIET)
ISSN (online): 2581-3048
Volume 7, Issue 10, pp 249-257, October-2023
https://doi.org/10.47001/IRJIET/2023.710032

The captured network packets are sent to a cloud-based server using a database synchronization. This step allows the use of a hosted machine learning model trained specifically for anomaly detection. By utilizing machine learning, the model efficiently analyzes incoming data to identify deviations from established patterns. This enables quick identification of possible anomalies. A Python script is used to detect anomalies and trigger a response, allowing the firewall to quickly implement protective measures. The anomaly triggers the use of a strategically designed honeypot to divert potential attackers from critical systems.

The use of honeypots is an essential part of the methodology. This honeypot is deployed on a virtual box environment using virtual router name called MikroTik Ubuntu instance. It creates a simulated environment that imitates real services, attracting potential attackers to engage with it instead of actual systems. Open-source honeypot solutions like "Cowrie" and "Dionaea" are commonly used to simulate various services such as SSH, TELNET, and web services. This methodology establishes a regulated framework for observing attacker behavior and gaining valuable insights into their strategies.

To improve analysis, it is important to integrate mechanisms for sharing threat intelligence into the methodology. This involves using the attacker IP addresses to access the VirusTotal API. This service extracts important contextual information about the attacker's IP address, improving analysis and understanding of potential threats.

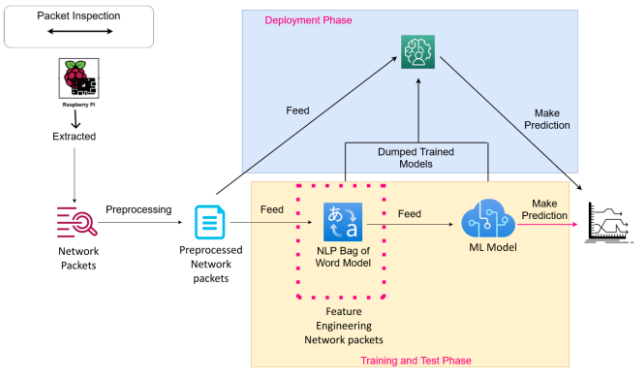

Figure 3: Network Anomaly Detection using ML

The network anomaly detection approach combines Natural Language Processing (NLP) and a naive Bayes machine learning (ML) model, as shown in the diagram. The dataset includes data from both malicious and benign network packets. It has a severity column that indicates the packet's acquisition method and distinguishes between malicious and non-malicious origins. Packets with harmful code are assigned a severity value of 1, while those without malicious content receive a score of 0. NLP techniques are effectively applied to handle this dataset and extract a comprehensive set of features.

In here used NLP techniques to create a bag-of-words model and feed it into a naive Bayes ML model to predict network anomalies. Prior research has emphasized the high predictive accuracy of the naive Bayes model for this task, which is attributed to its remarkable precision.

3) IoT related Netowrk Anomaly Detection

The process begins by obtaining network traffic data using the tcpdump tool. The implementation automates packet capture by creating a cron job, ensuring a smooth and uninterrupted data collection process. The collected network packets are sent to a cloud server for efficient data storage and management. A database synchronization is used to transfer captured packets to the cloud-hosted environment, ensuring data integrity and real-time accessibility.

The use of supervised machine learning models, specifically Random Forests, is an important part of the methodology. This entails creating a robust machine learning module trained on a dataset of network packet information. The model can identify unique anomalies in IoT devices. This is a major concern in the current technological landscape. When anomalies are detected, a Python script is activated to initiate firewall rules, enabling an agile and dynamic responsive mechanism.

The next step in the process involves intentionally redirecting potential attackers to a internal VM hosted honeypot. This strategic maneuver has two benefits: it hinders adversaries and provides valuable insights into potential attackers' operational methods. The honeypot is intentionally designed to create a controlled environment that promotes interaction, allowing for a thorough examination of attacker behavior.

After detecting and controlling network abnormalities, the approach involves increasing complexity by utilizing external services. The integration of the "fraudguard.io" public API improves the research ecosystem by enabling instant alert generation. The alerts contain crucial information about the IP addresses of the attackers. This integration seamlessly integrates into the workflow, providing stakeholders with timely information for effective decision-making and proactive defense strategies.

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 10, pp 249-257, October-2023

https://doi.org/10.47001/IRJIET/2023.710032

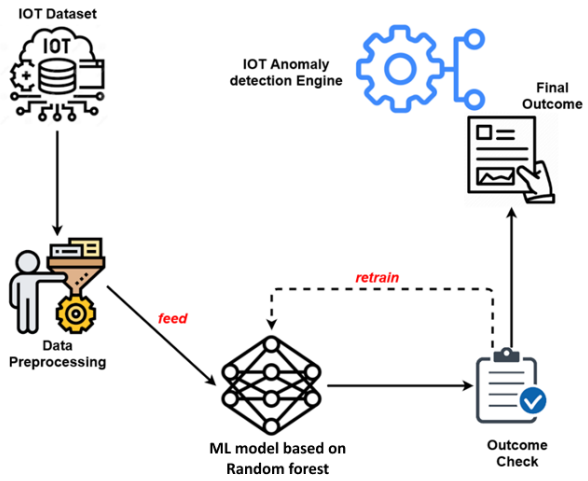

Figure 3: IoT related Network Anomaly Detection using ML

The diagram shows the use of a forest algorithm to detect anomalies in an IoT network. The methodology involves using a pre-processed dataset from Kaggle that includes engineered features. Python is used to preprocess data before inputting it into a random forest regression algorithm. This algorithm is chosen for its ability to accurately detect and categorize various types of anomalies.

To improve anomaly severity prediction, a Random Forest Regression model is trained with processed data. The algorithm was selected for its strong capability to efficiently handle complex interactions between text data and severity labels using ensemble learning techniques.

4) Signature & Behavior Based Mlaware Detection

"Real-time Network-based Anomaly Detection and Malware Analysis for SMEs and Smart Homes" uses supervised machine learning models to detect signature-based and behavior-based malware. This plan aims to provide SMEs with a dependable real-time solution for network anomalies and virus detection. This boosts cybersecurity.

Signature-based malware detection begins this strategy. This approach finds known malware patterns. Python and the dpkt package derive unique signatures from intercepted network traffic. The SQLite DBs send extracted signatures to cloud servers. The cloud server employs Snort, an open-source IDS/IPS, or Emerging Threats. Snort's rule-based architecture finds packets with extracted signatures. Snort promptly alerts of potential malware threats.

Snort and Barnyard2 enhance signature detection. Barnyard2 reads Snort's binary logs. Databases hold processed logs for efficient analysis. Comparing with open-source malware signature databases like ClamAV improves detection. This signature-based security prevents known malware penetration.

Behavior-based malware detection, a dynamic approach that evaluates system deviations, confirms the method's efficacy. This component detects harmful software patterns using trained supervised machine learning models. Network packets provide the data for analysis. The cloud-based machine learning module evaluates packets in real time.

Signature-based and behavior-based malware detection builds a complete ecosystem. Hosting Snort and machine learning modules on cloud servers boosts computational power and scalability. This capability analyzes network traffic in real time, helping the system detect and eliminate malware threats.

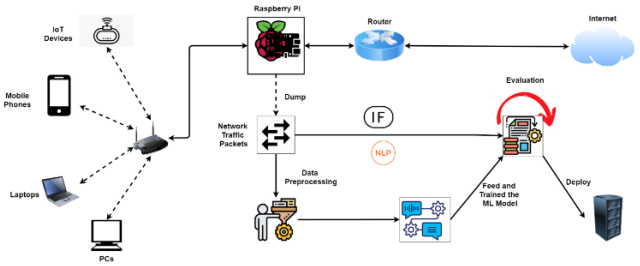

Figure 3: Signature and Behavior base Malware Detection using ML

The diagram offers insights into the integration of behavior-based malware detection and the Naive Bayes machine learning model. The process starts by obtaining data from a pre-engineered Kaggle dataset. The data then goes through extensive preprocessing and filtering using a Python script. The processed data is inputted into a well-trained Naive Bayes machine learning model, which uses its probabilistic framework to analyze the dataset's subtle behavioral patterns.

IV. RESULTS AND DISCUSSION

The system functions smoothly and demonstrates outstanding precision by leveraging its integrated machine learning modules. It has exceptional aptitude in carrying out diverse tasks.

The methodology given in this study demonstrates its effectiveness in real-time anomaly detection and malware analysis in network settings that serve small and medium-sized organizations (SMEs) and smart homes. The integration of machine learning components, such as Internet of Things (IoT) anomaly detection, network anomaly detection, and malware analysis, significantly enhances the effectiveness of cybersecurity measures.

In addition, the system has exceptional proficiency in providing notifications in a punctual manner. The system

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 10, pp 249-257, October-2023

https://doi.org/10.47001/IRJIET/2023.710032

utilizes a threat-sharing framework that guarantees a prompt reaction to developing threats, accomplished by employing a collaborative approach driven by the community. The utilization of natural language processing (NLP) enhances the precision of anomaly detection, resulting in the production of clear notifications that improve user understanding and expedite efficient responses. This innovation improves usability by addressing the disparity in understanding between individuals with technical expertise and those without technical knowledge.

The current research presents a secure system specifically designed for smart homes and small and medium enterprises (SMEs). The fundamental elements of this solution have been carefully crafted to achieve specified objectives, and the tactics implemented to accomplish and progress each aim have been well deliberated. The table below concisely presents the summary results and impacts of the applied approach, with a specific emphasis on real-time anomaly detection and malware analysis in the context of small and medium-sized enterprises (SMEs) and smart homes.

Table 1: Key Feature on Proposed System

Aspect	Description
ML-based Anomaly Detection	The solution detects anomalies in IoT networks using supervised ML like Random Forests to protect SMEs and smart homes.
Network Anomaly Detection and Timely Alerts	Network anomaly detection quickly discovers anomalous traffic patterns, alerting stakeholders with relevant insights to efficiently address security threats.
Raspberry Pi-based Firewall	By inspecting data packets, the Raspberry Pi-based firewall secures internal networks from external attacks.
Cost-Effectiveness Analysis	The economic research shows SMEs and smart homes a cost-effective cybersecurity solution.
Practicality and Usability	Its accurate anomaly detection, clear alarms, and internal network security demonstrate its use. This comprehensive defense approach meets SMEs and smart home cybersecurity needs.

The following section provides an overview of the outcomes obtained from the assessment of machine learning models that have been implemented within the primary system. The evaluation of these machine learning models encompassed the utilization of diverse metrics, such as Accuracy, Mean Squared Error (MSE) Value, Mean Absolute Error (MAE) Value, R2 Score, Precision, Recall, and F1 Score.

- The External URL Severity ML model attained an accuracy of 0.850, accompanied by mean squared error (MSE) and mean absolute error (MAE) values of 0.150. The R2 score for the model was computed to be 0.716.
- The Internet of Things (IoT) Anomaly Detection machine learning (ML) model was developed utilizing the Random Forest Regressor technique. Nevertheless, the utilization of evaluation metrics for this model was not specified.
- The Malware Detection machine learning model achieved an Accuracy of 0.873, accompanied by Precision, Recall, and F1 Score values of 0.977, 0.863, and 0.882, respectively.
- The Machine Learning model for Network Anomaly Detection achieved an Accuracy score of 0.992. Furthermore, the Precision, Recall, and F1 Score metrics for this model were calculated to be 0.998, 0.993, and 0.996, correspondingly.

The system's cost-effectiveness was also assessed. The economic study shows that the recommended cybersecurity strategy is cheap for SMEs and smart homes. Smaller businesses and homeowners need affordable, trustworthy cybersecurity solutions. The system's cost-effectiveness and extensive protection make it practical and appealing to its intended customers.

V. CONCLUSION AND FUTURE WORK

This study proposes a framework for real-time network-based anomaly detection and malware analysis specifically developed for small and medium-sized enterprises (SMEs) and smart homes. This approach enhances security while maintaining user-friendliness by utilizing advanced technologies like Raspberry Pi firewalls, supervised machine learning, NLP, and collaborative threat-sharing systems. The inclusion of signature-based and behavior-based malware analysis, such as zero-day attack detection, improves network resilience.

Future research should prioritize comparative studies of signature-based detection and machine learning, as well as behavioral analysis, to determine their relative effectiveness. Enhancing zero-day attack detection through evolving technologies is crucial. Enhancing usability and creating user-friendly interfaces can boost adoption among SMEs and homeowners. A cost-effectiveness study will enhance cyber security solutions for organizations with limited resources. The framework improves protection in the ever-changing digital landscape of real-time network security for SMEs and smart homes.

ISSN (online): 2581-3048
Volume 7, Issue 10, pp 249-257, October-2023
https://doi.org/10.47001/IRJIET/2023.710032

REFERENCES

[1] "How to stay protected against Cyber Threats? - Advancing Digital Life." https://www.axiatadigitallabs.com/how-to-stay-protected-against-cyber-threats/ (accessed Aug. 14, 2023).

[2] A.Almogren, H. Almajed, A. Alzahrani, and T. H. H Aldhyani, "Design of Efficient Based Artificial Intelligence Approaches for Sustainable of Cyber Security in Smart Industrial Control System," Sustain. 2023, Vol. 15, Page 8076, vol. 15, no. 10, p. 8076, May 2023, doi: 10.3390/SU15108076.

[3] B. A. Bhuvaneswari and S. S., "Anomaly detection framework for Internet of things traffic using vector convolutional deep learning approach in fog environment," Futur. Gener. Comput. Syst., vol. 113, pp. 255–265, Dec. 2020, doi: 10.1016/J.FUTURE.2020.07.020.

[4] A.Chatterjee and B. S. Ahmed, "IoT anomaly detection methods and applications: A survey," Internet of Things, vol. 19, p. 100568, Aug. 2022, doi: 10.1016/J.IOT.2022.100568.

[5] M. R. Amal and P. Venkadesh, "H-DOCTOR: Honeypot based firewall tuning for attack prevention," Meas. Sensors, vol. 25, p. 100664, Feb. 2023, doi: 10.1016/J.MEASEN.2022.100664.

[6] S. Sibi Chakkaravarthy, D. Sangeetha, and V. Vaidehi, "A Survey on malware analysis and mitigation techniques," Comput. Sci. Rev., vol. 32, pp. 1–23, May 2019, doi: 10.1016/J.COSREV.2019.01.002.

[7] "EU's Cybersecurity Strategy for the Digital Decade." https://www.cyber-diplomacy-toolbox.com/EU_Cybersecurity_Strategy_for_the_Digital_Decade.html (accessed Aug. 14, 2023).

[8] A.Alqudhaibi, M. Albarrak, A. Aloseel, S. Jagtap, and K. Salonitis, "Predicting Cybersecurity Threats in Critical Infrastructure for Industry 4.0: A Proactive Approach Based on Attacker Motivations," Sensors 2023, Vol. 23, Page 4539, vol. 23, no. 9, p. 4539, May 2023, doi: 10.3390/S23094539.

[9] K. L. Kermanidis, M. Maragoudakis, N. Rawindaran, A. Jayal, and E. Prakash, "Exploration of the Impact of Cybersecurity Awareness on Small and Medium Enterprises (SMEs) in Wales Using Intelligent Software to Combat Cybercrime," Comput. 2022, Vol. 11, Page 174, vol. 11, no. 12, p. 174, Dec. 2022, doi: 10.3390/COMPUTERS11120174.

[10] N. Jeffrey and Q. Tan, "A Review of Anomaly Detection Strategies to Detect Threats to Cyber-Physical Systems," pp. 1–34, 2023.

[11] U. Tariq, I. Ahmed, A. K. Bashir, and K. Shaukat, "A Critical Cybersecurity Analysis and Future Research Directions for the Internet of Things: A Comprehensive Review," Sensors 2023, Vol. 23, Page 4117, vol. 23, no. 8, p. 4117, Apr. 2023, doi: 10.3390/S23084117.

[12] Y. Li and Q. Liu, "A comprehensive review study of cyber-attacks and cyber security; Emerging trends and recent developments," Energy Reports, vol. 7, pp. 8176–8186, Nov. 2021, doi: 10.1016/J.EGYR.2021.08.126.

[13] G. Apruzzese et al., "The Role of Machine Learning in Cybersecurity," Digit. Threat. Res. Pract., vol. 4, no. 1, Mar. 2023, doi: 10.1145/3545574.

[14] I.H. Sarker, "Machine Learning: Algorithms, Real-World Applications and Research Directions," SN Comput. Sci., vol. 2, no. 3, pp. 1–21, May 2021, doi: 10.1007/S42979-021-00592-X/FIGURES/11.

[15] Y. K. Kim, J. J. Lee, M. H. Go, H. Y. Kang, and K. Lee, "A Systematic Overview of the Machine Learning Methods for Mobile Malware Detection," Secur. Commun. Networks, vol. 2022, 2022, doi: 10.1155/2022/8621083.

[16] P. Ryciak, K. Wasielewska, and A. Janicki, "Anomaly Detection in Log Files Using Selected Natural Language Processing Methods," Appl. Sci. 2022, Vol. 12, Page 5089, vol. 12, no. 10, p. 5089, May 2022, doi: 10.3390/APP12105089.

[17] Y. Kang, Z. Cai, C. W. Tan, Q. Huang, and H. Liu, "Natural language processing (NLP) in management research: A literature review," J. Manag. Anal., vol. 7, no. 2, pp. 139–172, Apr. 2020, doi: 10.1080/23270012.2020.1756939.

[18] Y. Ding, X. Xia, S. Chen, and Y. Li, "A malware detection method based on family behavior graph," Comput. Secur., vol. 73, pp. 73–86, Mar. 2018, doi: 10.1016/J.COSE.2017.10.007.

[19] M. S. Akhtar and T. Feng, "Malware Analysis and Detection Using Machine Learning Algorithms," Symmetry 2022, Vol. 14, Page 2304, vol. 14, no. 11, p. 2304, Nov. 2022, doi: 10.3390/SYM14112304.

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 7, Issue 10, pp 249-257, October-2023

https://doi.org/10.47001/IRJIET/2023.710032

Citation of this Article:

K.N.H De Silva, M.A.S.B Manchanayaka, D.L.S.I Punyasiri, H.A.D.N Perera, Anjalie Gamage, Narmada Gamage, "Realtime Network Based Anomaly Detection and Malware Analysis for SMEs and Smart Homes" Published in *International Research Journal of Innovations in Engineering and Technology - IRJIET*, Volume 7, Issue 10, pp 249-257, October 2023. Article DOI https://doi.org/10.47001/IRJIET/2023.710032

****

