IJAREEIE

e-ISSN: 2278-8875
p-ISSN: 2320-3765

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

Volume 13, Issue 5, May 2024

INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.317

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering (IJAREEIE)

| e-ISSN: 2278 – 8875, p-ISSN: 2320 – 3765| www.ijareeie.com | Impact Factor: 8.317|| A Monthly Peer Reviewed & Referred Journal |

||Volume 13, Issue 5, May 2024||

|DOI:10.15662/IJAREEIE.2024.1305024|

Malware Detection in Downloaded Files Using Various Machine Learning Approaches

R. Maheswari [1], T. Prathiba [2], J. Kavitha [3], M. Eswari [4]

Assistant Professor, Department of Electronics and Communication Engineering, Bharath Institute of Higher Education and Research, Selaiyur, Tamil Nadu, India [1],[2],[3] ,[4]

ABSTRACT: The fast propagation of computer networks has changed the viewpoint of network security. An easy accessibility conditions cause computer network as susceptible against several threats from hackers. Threats to networks are numerous and potentially devastating. Up to the moment, researchers have developed Intrusion Detection Systems (IDS) capable of detecting attacks in several available environments. Many of the technologies proposed are complementary to each other, since for different kind of environments some approaches perform better than others. This project presents a new intrusion detection system that is then used to survey and classify them. In our project we have used algorithms like Random Forest (RF) as existing system and Support Vector Machine (SVM) as proposed system is used. All are measured in terms of accuracy. From the results its proved that proposed Support Vector Machine (SVM) works better than existing Random Forest (RF).

KEYWORDS: Malware detection, Support Vector Machine (SVM), Random Forests, Anomaly Detection.

I. INTRODUCTION

The rapid evolution of digital technology has brought numerous conveniences to our lives, but it has also given rise to new challenges in cybersecurity. One of the persistent threats in the digital landscape is malware, malicious software designed to infiltrate systems, steal data, or cause damage. With the increasing volume and complexity of malware, traditional signature-based detection methods have become less effective, necessitating the adoption of advanced techniques such as machine learing. Machine learning (ML) has emerged as a powerful tool in the fight against malware, offering the ability to detect previously unseen threats and adapt to evolving attack strategies. In this context, the focus of this research is on detecting malware in downloaded files, which is a common vector for malware distribution. The primary objective of this study is to explore various machine learning approaches and their effectiveness in identifying and classifying malware within downloaded files. By leveraging the capabilities of machine learning models, we aim to enhance the accuracy and efficiency of malware detection, ultimately bolstering the cybersecurity posture of organizations and individuals. The taxonomy consists of the detection principle, and second of certain operational aspects of the intrusion detection system.

Many of the technologies proposed are complementary to each other, since for different kind of environments some approaches perform better than others.

II. OBJECTIVES

The primary objective of this research is to assess and compare the efficacy of different machine learning techniques in detecting malware within downloaded files. Specifically, the study aims to achieve the following objectives. Evaluate the performance of machine learning algorithms, such as support vector machines (SVM), deep learning neural networks, random forests, and others, in accurately identifying and classifying malware signatures and behaviors within downloaded files.

III. LITERATURE SURVEY

The team looked for and reviewed various patents, research papers, documents, newspapers, and magazine articles from diverse scenes for this project's literature review.

NL-IDS: Trust Based Intrusion Detection System for Network layer in Wireless Sensor Networks Umashankar Ghugar, Jayaram Pradhan IEEE 2022. we have proposed a trust based intrusion detection system (NL-IDS) for network layer in WSN to detect the Black hole attackers in the network. The sensor node trust is calculated as per the deviation of key factor at the network layer based on the Black hole attack.

Analyzing the Vulnerability of Wireless Sensor Networks to a Malicious Matched Protocol Attack George D. O'Mahon, Philip J. Harris, Colin C. Murphy IEEE 2022. In this paper, a specific vulnerability of WSNs is explored, termed here the matched protocol attack. This malicious attack uses protocol-specific structures to compromise a network using that protocol. Furthermore, a ZigBee cluster head network, which co-exists with ISM band services, consisting of XBee COTS devices is utilized, along with a real time spectrum analyzer, to experimentally evaluate the effect of matched protocol interference on a realistic network model.

Secure Knowledge and Cluster-Based Intrusion Detection Mechanism for Smart Wireless Sensor Networks Houbing Song IEEE 2022. This paper proposes a knowledge-based context-aware approach for handling the intrusions generated by malicious nodes. The system operates on a knowledge base, located at the base station, which is used to store the events generated by the nodes inside the network. The events are categorized and the cluster heads (CHs) are acknowledged to block maliciously repeated activities generated. The CHs can also get informational records about the maliciousness of intruder nodes by using their inference engines. The mechanism of events logging and analysis by the base station greatly affects the performance of nodes in the network by reducing the extra security-related load on them.

A Light-Weight Countermeasure to Forwarding Misbehavior in Wireless Sensor Networks: Design, Analysis, and Evaluation Cong Pu, Sunho Lim IEEE 2022. Network layer attacks are more severe since if the routing information is disregarded, disturbances may bring about routing loops, changing of routes etc. Selective forwarding attack is a type of active attack affecting network layers that selectively drops or refuses to forward the data packets. This paper discusses about an energy efficient detection-removal algorithm for effective detection of selective forwarding attack in a clustered WSN scenario.

Energy Efficient Detection-Removal Algorithm for Selective Forwarding Attack In Wireless Sensor Networks T.R Sreelakshmi, G.S Binu IEEE 2022. Network layer attacks are more severe since if the routing information is disregarded, disturbances may bring about routing loops, changing of routes etc. Selective forwarding attack is a type of active attack affecting network layers that selectively drops or refuses to forward the data packets. This paper discusses about an energy efficient detection-removal algorithm for effective detection of selective forwarding attack in a clustered WSN scenario. The impact of the malicious node in network parameters like packet delivery ratio, throughput, residual energy of network and end to end delay are analyzed.

IV. METHODOLOGY

- Data Collection and Preprocessing
- Feature Selection and Engineering
- Model Training and Evaluation
- Hyperparameter Tuning and Optimization
- Performance Metrics and Analysis
- Comparison of Approaches
- Validation and Deployment
- Documentation and Reporting

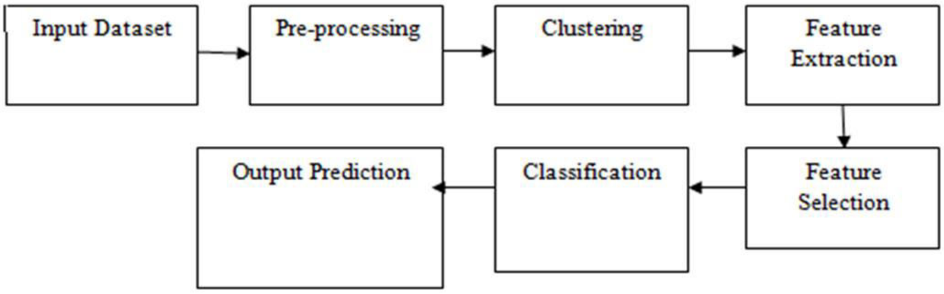


Fig 1: Block diagram Methodology

V. SYSTEM IMPLEMENTATION

There are two modes are there one is the input mode and another one is the final output mode the working of both nodes is explained below.

Fig2: Input mode

Input mode:

In the input mode we have to upload or choose the file which is needed for testing the malware. After the uploadation of the file the file under goes on some steps like the data collection and feature extraction from the uploaded file. Once the uploaded file meets the criteria which is customized by ourself then, the final output will be given.

VI. HARDWARE REQUIREMENTS

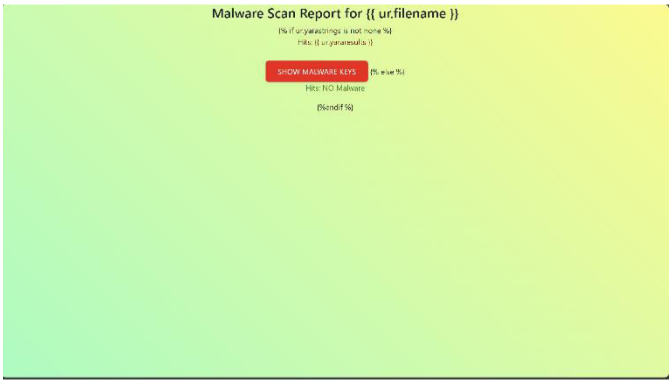
- Processor: Intel Core I5 Processor.GPS
- Ram: 8 GB RAMVibration sensor
- Hard Disk: 1 T.B Hard Disk
- 14 inch monitor

SOFTWARE REQUIREMENTS

- Technology : Python
- IDE : Python IDE

- Web Server : Jupyter/Anaconda/Panda

VII. RESULTS AND DISCUSSION



Malware Scan Report for {{ ur.filename }}

{% if ur.yarastrings is not none %}

Hits: {{ ur.yararesults }}

{% else %}

Hits: NO Malware

{%endif %}

Fig 3 : Final output

By using this model can that a file contains a malware or not.As explained in the system implementation in the previous section the input mode or the interface is created by the help of python language. The results that have been observed in the case of supervised machine learning models are much better as compared to those obtained in the case of unsupervised machine learning models. Among supervised machine learning models, the Random Forest Model is the most accurate. On the other hand, in the case of unsupervised machine learning models, Principal Component Analysis has been observed to achieve the highest accuracy.

VIII. CONCLUSION

In conclusion, Intrusion detection is currently attracting interest from both the research community and commercial companies. We have given background of the current-state-of-the-arts of IDS, based on a proposed taxonomy illustrated with examples of past and current projects. This taxonomy also highlights the recent work and covers the past and current developments adequately. Each of its technique has its own advantages and disadvantages. We believe that no single criterion can be used to completely defend against computer network intrusion.

FUTURE WORK

Investigate techniques for enhancing the interpretability and explainability of machine learning models used for malware detection. XAI methods can help cybersecurity analysts understand the reasoning behind model predictions, identify vulnerabilities, and improve trust in automated detection systems.Zero-Day Malware Detection: Focus on detecting zero-day malware, which are previously unknown threats with no existing signatures or patterns. Explore unsupervised learning, anomaly detection, and behavior-based approaches to identify suspicious activities and zero-day attacks in downloaded files.

REFERENCES

[1] W. Han, J. Xue, Y. Wang, L. Huang, Z. Kong, L. Mao "Maldae: Detecting and explaining malware based on correlation and fusion of static and dynamic characteristics."

[2] P. Burnap, R. French, F. Turner, K. Jones"Malware classification using self organizing feature maps and machine activity data".

[3] Sayadi H, Patel N, SMPD, Sasan A, Rafatirad S, Homayoun H. Ensemble learning for effective run-time hardware-based malware detection "A comprehensive analysis and classification. In 2018 55th ACM/ESDA/IEEE Design Automation Conference (DAC); 2018. pp. 1–6. oi:10.1109/DAC.2018.8465828.

[4] J. Acharya, A. Chuadhary, A. Chhabria, S. Detecting malware, malicious urls and virus using machine learning and signature matching.

[5] F. Akyildiz et al., “Wireless Sensor Networks: A Survey, “Elsevier Comp. Networks, vol. 3, no. 2, 2019, pp. 393–422.

[6] G.Li, J.He, Y. Fu. “Group-based intrusion detection system in wireless sensor networks” Computer Communications, Volume 31, Issue 18 (December 2019).

[7] Michael Brownfield, “Wireless Sensor Network Denial of Sleep Attack”, Proceedings of the 2019 IEEE Workshop on Information Assurance and Security United States Military Academy, West Point, NY.

[8] FarooqAnjum, DhanantSubhadrabandhu, SaswatiSarkar *, Rahul Shetty, “On Optimal Placement of Intrusion Detection Modules in Sensor Networks”, Proceedings of the First International Conference on Broadband Networks (BROADNETS19).

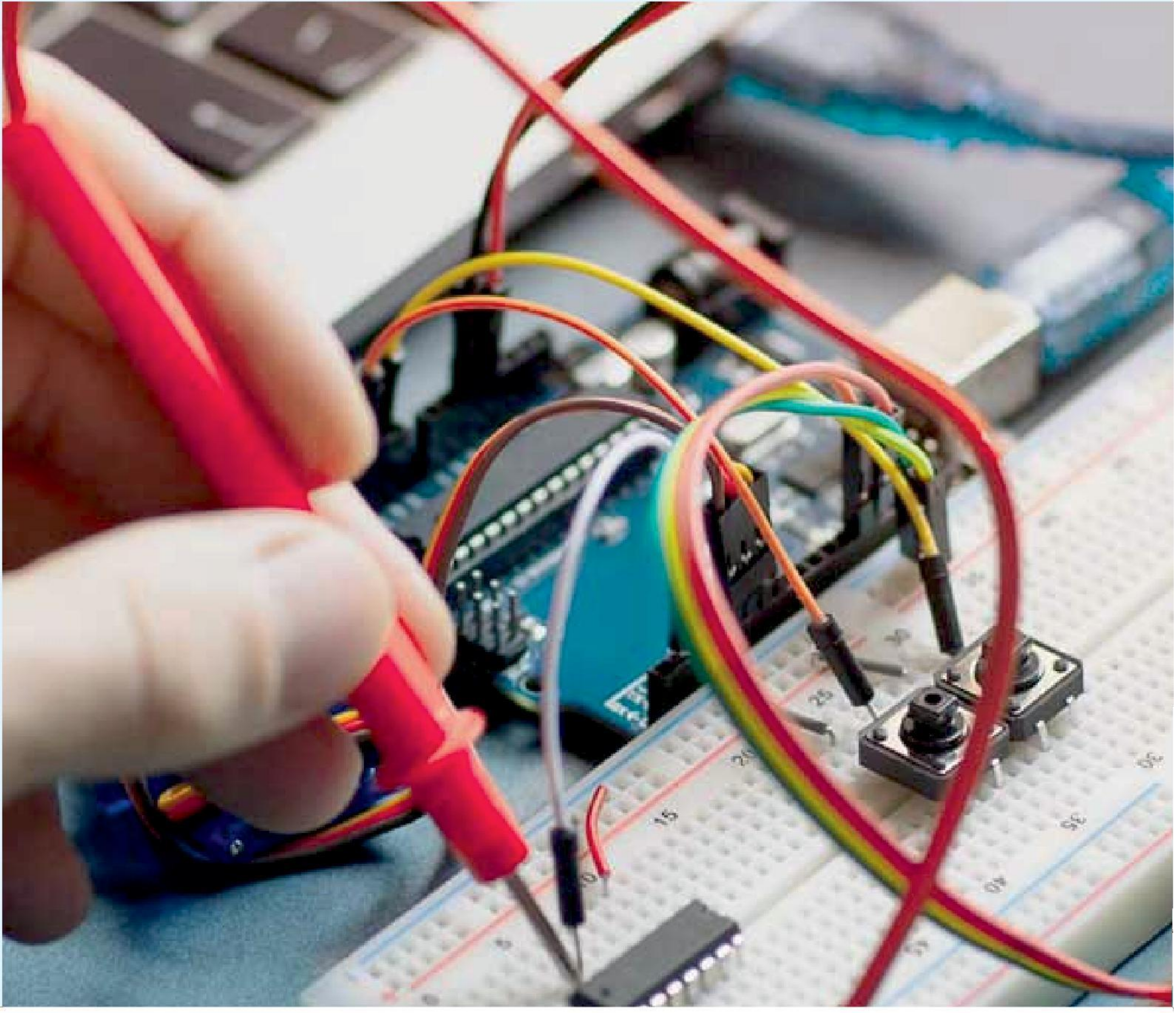
[9] : N. Udayakumar, V.J. Saglani, A.V. Cupta, T. Subbulakshmi“Malware classification using machine learning algorithms”

[10] : A. Kumar, K. Abhishek, K. Shah, D. Patel, Y. Jain, H. Chheda, P. Nerurkar“Malware detection using machine learning”

[11] : M. Naseer, J.F. Rusdi, N.M. Shanono, S. Salam, Z.B. Muslim, N.A. Abu, I. Abadi“Malware detection: Issues and challenges”

[12] : S. Sharma, C. Rama Krishna, S.K. Sahay“Detection of advanced malware by machine learning techniques”

[13] : M. Ahmadi, D. Ulyanov, S. Semenov, M. Trofimov, G. Giacinto“Novel feature extraction, selection and fusion for effective malware family classification”



INNO SPACE
SJIF Scientific Journal Impact Factor

doi crossref

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

International Journal of Advanced Research

in Electrical, Electronics and Instrumentation Engineering