IJETIR

ISSN (Online) 2583-0554

International Journal of Emerging Technologies and Innovative Research (IJETIR)

Volume 4, Issue 6, June 2024

Impact Factor: 5.731

Efficient Malware Detection in Android Devices to Improve Cyber Security

Faizur Rahaman.R[1] and Dr. S. Prasanna[2]

PG Student, Department of Computer Applications[1]

Professor, Department of Computer Applications[2]

vrfaizur1134@gmail.com & prasanna.scs@velsuniv.ac.in

Vels Institute of Science, Technology & Advanced Studies (VISTAS), Chennai, Tamil Nadu, India

Abstract: *The project introduces a novel framework for detecting Android malware based on permissions, utilizing multiple linear regression methods. Permissions play a crucial role in the security of the Android operating system, serving as fundamental indicators of an application's behavior. Through static analysis, the framework extracts application permissions and employs machine learning techniques to conduct security analyses.*

Specifically, the framework employs multiple linear regression techniques to develop two classifiers for permission-based Android malware detection. These classifiers leverage the relationships between various permission attributes to accurately identify potentially malicious applications. Notably, the framework achieves notable performance levels using classification algorithms without the need for overly complex models.

In the project, the existing system utilizes the Random Forest (RF) algorithm, while the proposed system adopts the Support Vector Machine (SVM) algorithm. Both algorithms are evaluated in terms of accuracy, with the results demonstrating that the proposed SVM approach outperforms the existing RF method. This highlights the effectiveness of SVM in accurately detecting Android malware based on permission analysis.

Keywords: Android Malware Detection, Cybersecurity, Machine Learning, Android Security, Threat Detection, Behavioral Analysis, Signature-based Detection

I. INTRODUCTION

In the evolution of mobile phones, their utility has expanded far beyond simple text messaging to encompass critical activities such as banking transactions, social media engagement, and personal data storage. However, this increased functionality has also made mobile devices prime targets for malware developers. Android, an open-source Linux-based mobile operating system, has gained widespread adoption among mobile device manufacturers due to its flexibility and cost-effectiveness. As a result, the majority of the market is now dominated by Android devices. According to Statista's data, Android's market share has surgedfrom 30% in the fourth quarter of 2010 to 88% in the second quarter of 2018.

The open-source nature of Android, combined with its flexibility for users to access applications from sources beyond official app stores, further enhances its appeal. While third-party application repositories offer users a wider selection of apps, they also present increased security risks. Despite rigorous vetting processes in official app repositories like Google Play, instances of malware can still slip through. Wang et al.'s research evaluated over 6 million applications from 17 different app stores, revealing that malware is prevalent across various platforms. Although Google Play is generally considered more reliable, malware can still be found in almost all app stores.

To combat the threat of Android malware, researchers have developed detection systems that leverage application permissions as key attributes. When users install an application, they are prompted to grant various permissions, which govern the app's behavior. Malicious apps may misuse these permissions to carry out harmful activities, highlighting theimportance of users paying attention to requested permissions. In this study, machine learningmodels

IJETIR
ISSN (Online) 2583-0554
International Journal of Emerging Technologies and Innovative Research (IJETIR)
Volume 4, Issue 6, June 2024
Impact Factor: 5.731

are employed to analyze the permissions requested by applications and determine whether they exhibit malicious behavior. By correlating requested permissions with known malware characteristics, these models can effectively identify potentially harmful apps and safeguard users against security threats.

Objective

- **Developing Lightweight Detection Algorithms**: Create efficient algorithms that minimize resource consumption while maximizing detection accuracy.
- **Real-time Detection Capabilities**: Enable detection mechanisms that can identify malware in real-time to mitigate immediate threats.
- **Enhancing Detection Sensitivity and Specificity**: Improve the ability to distinguish between malware and legitimate applications through advanced detection techniques.

II. LITERATURE SURVEY

Analyzing the Vulnerability of Wireless Sensor Networks to a Malicious Matched Protocol Attack George D. O'Mahon, Philip J. Harris, Colin C. Murphy IEEE 2023.

Safety critical, Internet of Things (IoT) and space-based applications have recently begun to adopt wireless networks based on commercial off the shelf (COTS) devices and standardized protocols, which inherently establishes the security challenge of malicious Malwares. Malicious Malwares can cause severe consequences if undetected, including, complete denial of services. Particularly, any safety critical application requires all services to operate correctly,as any loss can be detrimental to safety and/or privacy. Therefore, in order for these safety critical services to remain operational and available, any and all Malwares need to be detected and mitigated. Whilst Malware detection is not a new research area, new vulnerabilities in wireless networks, especially wireless sensor networks (WSNs), can be identified. 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. Through attack exploration, this paper provides evidence that traditional spectral techniques are not sufficient to detect an Malware using this style of attack. Furthermore, a ZigBee clusterhead 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. Results of this evaluation are provided in terms of device errors and spectrum use. This malicious challenge is also examinedthrough Monte-Carlo simulations. A potential detection technique, based on coarse inter-nodedistance measurements, which can theoretically be used to detect matched protocol interferenceand localize the origin of the source, is also suggested as a future progression of this work. Insights into how this attack style preys on some of the main security risks of any WSN (interoperability, device

Abnormal-Node Detection Based on Spatio-Temporal and Multivariate-Attribute Correlation in Wireless Sensor Networks Nesrine Berjab, Hieu Hanh Le, Chia-Mu Yu, Sy-Yen Kuo, Haruo Yokota IEEE 2023.

In wireless sensor networks (WSNs), data can be subject to malicious attacks and failures, leading to unreliability. This vulnerability poses a challenge to environmental monitoring applications by creating false alarms. To guarantee a trustworthy system, we therefore need todetect abnormal nodes. In this paper, we propose a new framework for detecting abnormal nodes in clustered heterogeneous WSNs. It makes use of observed spatiotemporal (ST) and multivariate-attribute (MVA) sensor correlations, while considering the background knowledge of the monitored environment. Based on the ST correlations, the collected data is analyzed by computing the crosscorrelation between sensor streams. A new method is proposedfor evaluating the intensity of the correlation between two sensor streams. The crosscorrelation value obtained is compared against two thresholds, the lag threshold and the correlation threshold. Based on available background knowledge and the observed MVA correlations, a number of rules are presented to detect abnormal nodes while identifying real events. Our experiments on real-world sensor data demonstrate that our approach captures the correlation and discovers abnormal nodes efficiently.

IJETIR
ISSN (Online) 2583-0554
International Journal of Emerging Technologies and Innovative Research (IJETIR)
Volume 4, Issue 6, June 2024
Impact Factor: 5.731

Secure Knowledge and Cluster-Based Malware Detection Mechanism for Smart Wireless Sensor Networks Amjad Mehmood, Akbar Khanan, Muhammad Muneer Umar, Salwani Abdulla, Khairul Akram Zainol Ariffin, Houbing Song IEEE 2023.

Wireless sensor networks, due to their nature, are more prone to security threats than other networks. Developments in WSNs have led to the introduction of many protocols specially developed for security purposes. Most of these protocols are not efficient in terms of putting an excessive computational and energy consumption burden on small nodes in WSNs. This paper proposes a knowledge-based context-aware approach for handling the Malwares 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 eventsare 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 analysisby the base station greatly affects the performance of nodes in the network by reducing the extra security-related load on them.

III. METHODOLOGY SECTION

Random Forest (RF) is a widely-used ensemble learning method in machine learning forclassification, regression, and other tasks. It operates by constructing numerous decisiontrees during training and then outputs the mode of the classes (classification) or the mean/average prediction (regression) of the individual trees. One of the main advantages of random forests is that they address the tendency of decision trees to overfit their training data. They achieve this by aggregating multiple decision trees trained on different subsets of the training data, thereby reducing variance while introducing a small increase in bias.

However, despite their popularity and effectiveness, random forests do have limitations and disadvantages. Firstly, their accuracy may be lower compared to more advanced techniques like gradient boosted trees. Additionally, random forests may not perform well in datasets with certain characteristics, as their performance can be affected by the nature of the data.

Moreover, random forests can be computationally expensive and time-consuming, especially when dealing with large datasets. This can lead to higher computational costs, making them less practical for real-time applications or scenarios where computational resources are limited.

Furthermore, random forests lack standards in terms of their implementation, whichcan make it challenging to apply them uniformly across different datasets or environments. Their performance may vary depending on the specific dataset and parameters chosen for training.

Another disadvantage is that random forests may not be suitable for all types of datasets or machine learning tasks. While they are effective for many applications, theremay be scenarios where other algorithms or techniques are more appropriate.

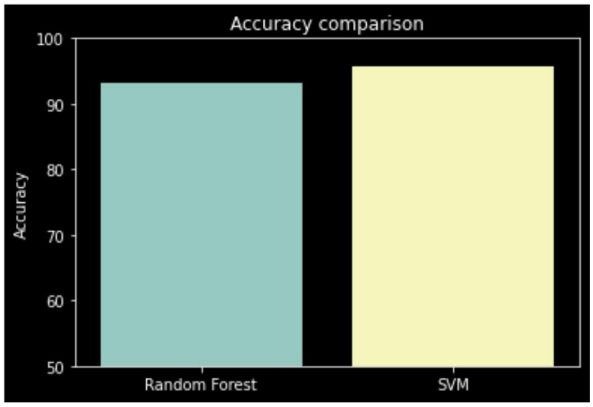
In summary, while random forests offer many benefits, including robustness to overfitting and flexibility in handling various types of data, they also have limitationssuch as potential lower accuracy, computational complexity, lack of standards, and limited applicability to certain datasets. These factors should be carefully consideredwhen choosing an appropriate machine learning algorithm for a given task or application.

IJETIR
ISSN (Online) 2583-0554
International Journal of Emerging Technologies and Innovative Research (IJETIR)
Volume 4, Issue 6, June 2024
Impact Factor: 5.731

IV. OUTPUT SCREENS

Out[12]:

	Agorithms	accuracy
0	Random Forest	93.143003
1	SVM	95.729000



V. CONCLUSION

In the realm of Android operating system security, application permissions play a pivotal role. These permissions, gleaned from applications, serve as crucial attributes in the identification of malicious software through machine learning algorithms in a recentstudy. Android malware detection is executed through the utilization of two rule-based classification models integrated with Hybrid models. Nevertheless, what sets these approaches apart is the simplicity and user-friendliness of the classifiers. This stands out as the most notable advantage of the proposed methodologies.

The importance of application permissions in ensuring the security of the Android operating system cannot be overstated. These permissions, extracted from various applications, serve ascritical indicators in the identification of potentially malicious software through the employment of machine learning algorithms in a recent study. The detection of Android malware is facilitated by employing two rule-based classification models integrated with Hybrid models. Notably, the classifiers employed in these approaches are characterized by their simplicity and user-friendliness, which represent the most significant advantage of the proposed methodologies.

REFERNECES

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

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

[3] 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.

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

[5] Parveen Sadotra et al, International Journal of Computer Science and Mobile Computing, Vol.5 Issue.9, September- 2019, pg. 23-28

[6] K. Akkayaand M. Younis, ―A Survey of Routing Protocols in Wireless Sensor Networks, in the Elsevier Ad Hoc Network Journal, Vol. 3/3 pp. 325-349, 2019.

[7] A. Abduvaliyev, S. Lee, Y.K Lee, “Energy Efficient Hybrid Malware Detection System for Wireless Sensor Networks”, IEEE International Conference on Electronics and Information Engineering, Vol.2, pp. 25-29, August 2019.
[8] Parveen Sadotra and Chandrakant Sharma. A Survey: Intelligent Malware Detection System in Computer Security. International Journal of Computer Applications 151(3):18-22, October 2019.
[9] A. Araujo, J. Blesa, E. Romero, D. Villanueva, “Security in cognitive wireless sensor networks. Challenges and open problems”, EURASIP Journal on Wireless Communications and Networking, February 2019.
[10] A. Becher, Z. Benenson, and M. Dorsey, \Tampering with motes: Real-world physical attacks on wireless sensor networks." in SPC (J. A. Clark, R. F. Paige, F. Polack, and P. J.Brooke, eds.), vol. 3934 of Lecture Notes in Computer Science, pp. 104{118, Springer, 2019.
[11] I. Krontiris and T. Dimitriou, \A practical authentication scheme for in-network programming in wireless sensor networks," in ACM Workshop on Real- World Wireless Sensor Networks, 2019.