International Journal For Technological Research In Engineering
Volume 9, Issue 10, June-2022 *ISSN (Online): 2347 - 4718*

SURVEY PAPER ON ANDROID MALWARE DETECTION USING GENETIC ALGORITHM

Dheeraj S[1], Goutham GS[2], Jayasurya R[3], Sanjay G[4], Rakshitha R[5]
[1,2,3,4] Student, [5]Assistant Professor
Department of Computer Science and Engineering
Vidya Vikas Institute of Engineering and Technology, Mysuru

Abstract - Android platform due to open source characteristic and Google backing has the largest global market share. Being the world's most popular operating system, it has drawn the attention of cyber criminals operating particularly through wide distribution of malicious applications. This paper proposes an effectual machine-learning based approach for Android Malware Detection making use of evolutionary Genetic algorithm for discriminatory feature selection. Selected features from Genetic algorithm are used to train machine learning classifiers and their capability in identification of Malware before and after feature selection is compared. The experimentation results validate that Genetic algorithm gives most optimized feature subset helping in reduction of feature dimension to less than half of the original feature-set. Classification accuracy of more than 94% is maintained post feature selection for the machine learning based classifiers, while working on much reduced feature dimension, thereby, having a positive impact on computational complexity of learning classifiers.

Keywords - Android malware analysis, Genetic algorithm, Feature selection, Support vector classifier, artificial neural network.

1. INTRODUCTION

Android Apps are freely available on Google Playstore, the official Android app store as well as third-party app stores for users to download. Due to its open source nature and popularity, malware writers are increasingly focusing on developing malicious applications for Android operating system. In spite of various attempts by Google Play store to protect against malicious apps, they still find their way to mass market and cause harm to users by misusing personal information related to their phone book, mail accounts, Therefore, there is need to perform malware analysis or reverse-engineering of such malicious applications which pose serious threat to Android platforms. Broadly speaking, Android Malware analysis is of two types: Static Analysis and Dynamic Analysis. Static analysis basically involves analyzing the code structure without executing it while dynamic analysis is examination of the runtime behavior of Android Apps in constrained environment. Given in to the ever-increasing variants of Android Malware posing zero-day threats, an efficient mechanism for detection of Android malwares is required. In contrast to signature-based approach which requires regular update of signature database, machine-learning based approach in combination with static and dynamic analysis can be used to detect new variants of Android Malware posing zero-day threats. In [1], broad yet lightweight static analysis been performed achieving a decent detection accuracy of more than 94% using Support Vector Machine algorithm.

2. LITERATURE SURVEY

[1] Andrea Saracino, Daniele Sgandurra, Gianluca Dini and Fabio Martinelli
Android users are constantly threatened by an increasing number of malicious applications (apps), generically called malware. Malware constitutes a serious threat to user privacy, money, device and file integrity. In this paper we note that, by studying their actions, we can classify malware into a small number of behavioural classes, each of which performs a limited set of misbehaviours that characterize them. These misbehaviours can be defined by monitoring features belonging to different Android levels. In this paper we present MADAM, a novel host-based malware detection system for Android devices which simultaneously analyses and correlates features at four levels: kernel, application, user and package, to detect and stop malicious behaviours.

[2] Nikola Milosevic a , Ali Dehghantanha b , Kim-Kwang Raymond Chooc,∗
The majority of existing approaches need to perform analysis on the remote server or they require the Android device to be rooted. However, our permission-based approach can run on Android devices without root access and offers a relatively good accuracy in malware detection. The widespread adoption of Android devices and their capability to access significant private and confidential information have resulted in these devices being targeted by malware developers. Existing Android malware analysis techniques can be broadly categorized into static and dynamic analysis. In this paper, we present two machine learning aided approaches for static analysis of Android malware. The first approach is based on permissions and the other is based on source code analysis utilizing a bag-of-words representation model.

[3] Haisheng Yana), Lingling Peng been observed that mainly three approaches were considered which are as follows:

In the paper, an android malware detection method based on evolutionary super-network is proposed in order to improve the precision of android malware detection. Chi square statistics method is used for selecting characteristics on the basis of analyzing android authority. Boolean weighting is utilized for calculating characteristic weight. Processed characteristic vector is regarded as the system training set and test set; hyper edge alternative strategy is used for training super-network classification model, thereby classifying test set characteristic vectors, and it is compared with traditional classification algorithm. In this paper, the results show that the detection method proposed in the paper is close to or better than traditional classification algorithm. The experimental result shows that the detection method proposed in the paper has better performance in the collected dataset than traditional KNN, SVM and Bayes Net classifiers. Next, Android software API calling is combined. The performance of the super-network classifier is studied on the basis of greater dataset, thereby improving detection precision.

[4] G. Tejaswini1 , Ch. Monisha2 , B. Divya3 , D. Layasree4 , A. Bhavana5 , Ms. M. Pallavi6
Android was the most popular mobile operating system amongst smart phone users. Its high popularity, combined with the extended use of smart phones for everyday tasks as well as storing or accessing sensitive and personal data, has made Android applications the target of numerous malware attacks over the last few years and in the present. This paper proposes an effective machine-learning based approach for Android Malware Detection making use of evolutionary Genetic algorithm for discriminatory feature selection. Extract the features from APK files and these features can be applied to genetic algorithm. Selected features from Genetic algorithm are used to train machine learning classifiers and their capability in identification of Malware before and after feature selection is compared.

3. CONCLUSION

As the number of threats posed to Android platforms is increasing day to day, spreading mainly through malicious applications or malwares, therefore it is very important to design a framework which can detect such malwares with accurate results. Where signature-based approach fails to detect new variants of malware posing zero-day threats, machine learning based approaches are being used. The proposed methodology attempts to make use of evolutionary Genetic Algorithm to get most optimized feature subset which can be used to train machine learning algorithms in most efficient way. From experimentations, it can be seen that a decent classification accuracy of more than 94% is maintained using Support Vector Machine and Neural Network classifiers while working on lower dimension feature-set, thereby reducing the training complexity of the classifiers.

REFERENCE

[1] https://www.researchgate.net/publication/296624384
[2] www.elsevier.com/locate/compeleceng
[3] https://doi.org/10.1063/1.5033818
[4] www.ijaresm.com

