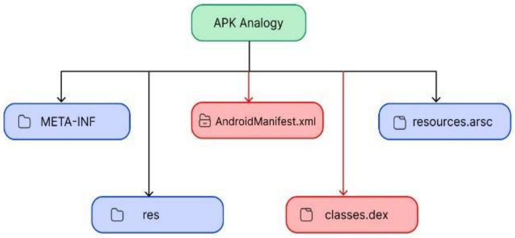
An Analysis of Android Malware Detection Methods

Senthil Kumar V[1], Ravina.C[2], Sneka.V[3], Kavipriya.T[4]
[1]Assistant Professor. Dept of Computer Science and Engineering
[2, 3, 4]Dept of Computer Science and Engineering
[1, 2, 3, 4] Kumaraguru College of Technology [Autonomous], Coimbatore, India

Abstract- *Modern society has made mobile devices an integral part of daily life, and malware is growing at the same time. Users often utilize Android programmes to conduct a variety of tasks, and the security of these applications depends on the rights that users provide to them.Malware attackers now frequently target the Android platform. Before installing an application, it is crucial to utilize a method to identify Android malware. Then, models created using various machine learning approaches by examining the current malware activity patterns and applying the information to identify any similar behavior carried out by unidentified attackers. In-depth investigation of machine learning techniques and methods for detecting malware on Android is provided by this study.*

I. INTRODUCTION

Since its launch in September 2008, Android has started dominating the whole of the mobile industry having around 75 percent global market share today. The features of android applications like easy and faster installation, user-friendly environment, and lower costs make it worthier than other OS available today. Android development has not only brought advancements in every aspect of using mobile phones but has also become a greater platform for the misuse of people's privacy and safety by adding malware to android applications. It is either due to unprofessional development by the developer himself or people purposefully injecting malware into the application. When we look into the composition of an APK (Android Application Package) file, we find two folders and three files, and their representation is shown in the following chart:



The scope of adding malware into the application is possible in two of the above-mentioned contents of the APK file. They are: AndroidManifest.xml file and classes.dexfile.The working of the application can be disturbed by altering the classes.dex file as it has the code for working the application. So, people who try intentionally to invade the working, try to modify the classes.dex file. However, it is a little more complex to do this as it involves editing the core part of the application. The other possibility is to modify the manifest file. The manifest file in general contains the structure of the android application. It includes the version of the application and all the required permissions essential for the unadorned running and functioning of the application. So, there is no doubt that these two contents (classes and the manifest file) in the apk file have the highest chance of being targeted by the one who wants to inject malware purposefully to perform any malicious act.

Any application that is malware-infected can be dealt with in two ways. One is to identify the presence of the malware before its installation in the android device and not proceed with the installation. To do this one should go through all the code in the apk file and find if there are any discrepancies in the code. The other way is to check for some malicious behavior while running the application. However, in this method, the malicious application might have already caused some problems before the user identifies the malware. There is a possibility that personal information and other information disturbing one's privacy has been fetched by any third party because of the malicious behavior of the application.

II. LITERATURE REVIEW

S NO	AUTHOR	TITLE	METHOD	ADVANTAGES	DISADVANTAGES
1	Seif EL Dein Mohamed, Mostafa Ashaf, Amr Ehab, Omar Shereef, Haytham Metwaie, Eslam Amer	Detecting Malicious Android Applications Based on API calls and Permissions Using Machine Learning Algorithms	To compare and analyze different android malware detection systems based on detection techniques, analysis processes and extracted features by using K- Means Machine Learning Methods.	Collecting Features that best reflect malicious activity as machine learning features aids in analyzing malware.	High Error Susceptibility. Risk of running inefficient algorithms and making limited predictions when not trained properly.
2	Long Wen and Haiyang Yu	An Android Malware Detection System Based on Machine Learning	A lightweight system capable of identifying malware. Extracting features with the method of static and dynamic analysis, it detects malicious android application using machine learning	A new Feature Selection algorithm PCA-RELIEF is also proposed to decrease dimensions of the features.	Training time is higher Doesn't perform well when a larger dataset is used.
3	S.J.K.,S. Chakravarty and R.K. Varma P	Feature Selection and Evaluation of Permission based Android Malware Detection.	With the use of information gain, Relief F, and Gain Ration, the dataset is scrutinized for feature reduction. Through feature reduction, the top 5 permissions that significantly influence classification were identified. The Randomizable Filtered Classifier generated great accuracy, according to a comparison between J48 algorithm, Multilayer Perceptron, Random Committee, Sequential Minimal Optimization (S MO), and this method.	In order to improve the malware detection system and lessen the time and space complexity, an attempt was made to select the important rights from the enormous list of 330 permissions that were accessible.	After feature reduction, the model was fed with just 5 permission characteristics. Consequently, the model developed cannot be trusted to identify malware.

4	Tian liang lu, Yanhui Du, Li Ouyang, Qiuyuchen, and Xirui Wang	Android Malware Detection Based On Hybrid Deep Model	Based on the effects of deep learning algorithms that integrate the Gate Recurrent Unit (GRU) and Deep Belief Network (DBN), where DBN processes static features and GRU processes dynamic features. In addition to extracting a static feature, we also built a comprehensive feature set by extracting dynamic features of malware at runtime in order to improve malware detection.	The frequency of this feature extraction is high, which indicates the most obfuscated malware samples are generated through automatic repackaging.	This model only works on extremely fast computers, is time-dependent, incapable of detecting malware, and has a higher risk of overfitting than deterministic algorithms.
5	Rishab, Agrawal Visha Shah, Sonam Chavan.	Android Malware Detection Using Machine Learning	Based on permission analysis and semantic analysis, they have implemented an admin and user panel where the admin panel has access to upload apk files and comments that can be used for semantic analysis in the user panel, where users can select the category and see the percentage of malicious applications.	In addition to determining whether an application has been properly done or not, semantic analysis is used to identify malware based on an application by comparing it to a dataset.	Only one algorithm of ML can't manage it, so requiring from other technologies for high efficiency.

III. INFERENCE

A feature selection strategy is proposed to dispose of the experimental findings and raw features and demonstrates whether the model performs with a high or low detection rate when compared to the conventional detection approaches, such as the detection method based on traditional methods(where an individual looks for bugs or malicious parts in the application).

Relief, information gain, and gain ratio are assessed for feature reduction of the Android malware permissions dataset. Then a few machine learning classifiers have been compared for accuracy. The size of the dataset has been decreased with the use of information gain feature selection.

The random filter classifier has produced the results with the highest accuracy. For the purpose of detecting malware on Android devices, static and dynamic analytic techniques have been used to create a hybrid deep learning model based on DBN and GRU.The dynamic elements of the application programme during runtime are extracted to enhance the Android malware feature set and new features with strong anti-obfuscation capabilities are added in order to combat obfuscation technology.

IV. CONCLUSION

The rate at which the number of android users grows is exponential. And there is no centralized authority given to anyone to govern the android applications so that malicious activities disturbing the code of conduct could be stopped.Instead of digging through the whole working code of the programme, this method of locating the malicious application making use of the rights that it wants from the user and is more easier and simpler. The necessity to download apps from unofficial websites and untrusted third-party sites result in regular individuals losing their personal information, casting doubt on the Android operating system as a whole. That encapsulates the main argument in favor of scanning for dangerous software before installing them.

REFERENCES

[1] Seif EI Dein Mohamed, Mostafa Ashaf, Amr EEhab, Omar Shereef, Haytham Metwaie and Eslam Amer, “Detecting Malicious Android Application Based on API calls and Permissions USing Machine Learning Algorithms”, 2021, in IEEEAccess, doi:10.1109/MIUCC52538.2021.9447594.

[2] Long Wen and HAiyang Yu, An android malware detection system based on machine learning, Aug 2017, doi: 10.1063/1.4992953.

[3] S. J. K., S. Chakravarty and R. K. Varma P., "Feature Selection and Evaluation of Permission-based Android Malware Detection," *2020 4th International Conference on Trends in Electronics and Informatics (ICOEI)(48184)*, 2020, pp.795-799,doi: 10.1109/ICOEI48184.2020.9142929.

[4] Tianliang Lu, Yanhui Du , Li Ouyang, Qiuyu Chen, and Xirui Wang, “Android Malware Detection Based on a Hybrid Deep Learning Model”,2020 , in IEEE Access, doi:10.1155/2020/8863617.

[5] R. Agrawal, V. Shah, S. Chavan, G. Gourshete and N. Shaikh, "Android Malware Detection Using Machine Learning," 2020International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE), 2020, pp. 1-4, doi: 10.1109/ic-ETITE47903.2020.491.

[6] E. Amer, "Permission-Based Approach for Android Malware Analysis Through Ensemble-Based Voting Model," *2021 International Mobile, Intelligent, and Ubiquitous Computing Conference (MIUCC)*, 2021, pp. 135-139, doi: 10.1109/MIUCC52538.2021.9447675.

[7] W. Cho, H. Lee, S. Han, Y. Hwang and S. -j. Cho, "Sustainability of Machine Learning-based Android Malware Detection Using API calls and Permissions," *2022 IEEE Fifth International Conference on Artificial Intelligence and Knowledge Engineering (AIKE)*, 2022, pp. 18-25, doi: 10.1109/AIKE55402.2022.00009.

[8] A. H. E. Fiky, A. Elshenawy and M. A. Madkour, "Detection of Android Malware using Machine Learning," *2021 International Mobile, Intelligent, and Ubiquitous Computing Conference (MIUCC)*, 2021, pp. 9-16, doi: 10.1109/MIUCC52538.2021.9447661.

[9] M. Gohari, S. Hashemi and L. Abdi, "Android Malware Detection and Classification Based on Network Traffic Using Deep Learning," *2021 7th International Conference on Web Research (ICWR)*, 2021, pp. 71-77, doi: 10.1109/ICWR51868.2021.9443025.

[10] W. Qing-Fei and F. Xiang, "Android Malware Detection Based on Machine Learning," *2018 4th Annual International Conference on Network and Information Systems for Computers (ICNISC)*, 2018, pp. 434-436, doi: 10.1109/ICNISC.2018.00094.

[11] A. A. Zaabi and D. Mouheb, "Android Malware Detection Using Static Features and Machine Learning," *2020 International Conference on Communications, Computing, Cybersecurity, and Informatics (CCCI)*, 2020, pp. 1-5, doi: 10.1109/CCCI49893.2020.9256450.

[12] N. C. Lê, T. -M. Nguyen, T. Truong, N. -D. Nguyen and T. Ngô, "A Machine Learning Approach for Real Time Android Malware Detection," *2020 RIVF International Conference on Computing and Communication Technologies (RIVF)*, 2020, pp. 1-6, doi: 10.1109/RIVF48685.2020.9140771.

[13] K. Liu, S. Xu, G. Xu, M. Zhang, D. Sun and H. Liu, "A Review of Android Malware Detection Approaches Based on Machine Learning," in IEEE Access, vol. 8, pp. 124579-124607, 2020, doi: 10.1109/ACCESS.2020.3006143.

[14] R. Vinayakumar, M. Alazab, K. P. Soman, P. Poornachandran and S. Venkatraman, "Robust Intelligent Malware Detection Using Deep Learning," in IEEE Access, vol. 7, pp. 46717-46738, 2019, doi: 10.1109/ACCESS.2019.2906934.

[15] Zhongyuan Qin, Yuqing Xu, Yuxing Di, Qunfang Zhang and Jie Huang, "Android malware detection based on permission and behavior analysis," International Conference on Cyberspace Technology (CCT 2014), 2014, pp. 1-4, doi: 10.1049/cp.2014.1352.

[16] P. Faruki et al., "Android Security: A Survey of Issues, Malware Penetration, and Defenses," in IEEE Communications Surveys & Tutorials, vol. 17, no. 2, pp. 998-1022, Secondquarter 2015, doi: 10.1109/COMST.2014.2386139.

[17] S. Sabhadiya, J. Barad and J. Gheewala, "Android Malware Detection using Deep Learning," 2019 3rd

International Conference on Trends in Electronics and Informatics (ICOEI), 2019, pp. 1254-1260, doi: 10.1109/ICOEI.2019.8862633.

[18] R. Lukas and G. Kołaczek, "Android Malware Detection Using Deep Learning Methods," 2021 IEEE 30th International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE), 2021, pp. 119-124, doi: 10.1109/WETICE53228.2021.00033.

[19] F. Shen, J. D. Vecchio, A. Mohaisen, S. Y. Ko and L. Ziarek, "Android Malware Detection Using Complex-Flows," in IEEE Transactions on Mobile Computing, vol. 18, no. 6, pp. 1231-1245, 1 June 2019, doi: 10.1109/TMC.2018.2861405.

[20] N. Penning, M. Hoffman, J. Nikolai and Y. Wang, "Mobile malware security challenges and cloud-based detection," 2014 International Conference on Collaboration Technologies and Systems (CTS), 2014, pp. 181-188, doi: 10.1109/CTS.2014.6867562.

[21] P. Agrawal and B. Trivedi, "A Survey on Android Malware and their Detection Techniques," 2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT), 2019, pp. 1-6, doi: 10.1109/ICECCT.2019.8868951.

[22] J. Zhang, F. Zou and J. Zhu, "Android Malware Detection Based on Deep Learning," 2018 IEEE 4th International Conference on Computer and Communications (ICCC), 2018, pp. 2190-2194, doi: 10.1109/CompComm.2018.8781037.

[23] A. Saracino, D. Sgandurra, G. Dini and F. Martinelli, "MADAM: Effective and Efficient Behavior-based Android Malware Detection and Prevention," in IEEE Transactions on Dependable and Secure Computing, vol. 15, no. 1, pp. 83-97, 1 Jan.-Feb. 2018, doi: 10.1109/TDSC.2016.2536605.