International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 11 Issue 08, August-2022

Implementation of Random Forest Algorithm for Android Malware Detection

Jyoti Dnyandeo Lokhande
Guide: Prof. Pramod Gosavi
Godavari College of Engineering, Jalgaon.

***Abstract*:- Android is a mobile operating system based on a modified version of the Linux kernel and other open-source software, designed primarily for touchscreen mobile devices such as smartphones and tablets. Android application is a software running on this framework. Android application development has brought ease of functionalities in our day-to-day life. Android has become the part of everyone's life, so as of hackers and attackers too. The increased use of android applications and popularity of this framework has increased cyberattacks through malicious applications. We propose an Android malware detection system for such malicious applications.**

Keywords: Android, Android APK, Malware Detection, Random Forest Algorithm, Ransomware.

INTRODUCTION

We propose an Android malware detection system for such malicious applications Android is open-source system as it can be downloaded from anywhere and from any source. The use of this open-source framework is increased in some decades. Anyone can download this framework and can develop their applications using the functionalities provided by this framework. The Google play store has provided many reasonable facilities for such applications to be uploaded to store. These applications are easily been downloaded from google play store by anyone who is using android smartphones. The malicious attacks or cyber-attacks through these android applications has increased nowadays. Many applications are restricted for their versions and unsupported application requirements of devices. Attackers mostly change the source code for the applications and insert the malicious code into it, so that when such application is installed on any device, immediately that device is been attacked and all the required information of that device is been hacked.

Google Play In May 2017, reported its new implicit malware defence for Android, Play Protect, which verifies applications and APK files whenever they are downloaded utilizing the Google official store or third-party stores. Since August 2017 and afterward, it has been accessible on all Android devices with Google Play Services 11 or above, and is set up on devices with Android 8.0 and above. However, when Play Protect is tested, it's only able to detect 51.8% of the test cases [1]. Attacks to these android applications can be done with different manners, Multiple steps with different sequence of procedures can be executed for single attack. Android antiviruses are available to detect and avoid various malwares. Many antiviruses are developed considering signature-based databases. Any new virus attack beyond the scope is not detected by these antiviruses.

PROPOSED METHOD

2.1 Reverse Engineering the application

The Android application apk is reverse engineered to get the packages and functionality code.

Reverse engineering (also known as backwards engineering or back engineering) is a process or method through which one attempts to understand through deductive reasoning how a previously made device, process, system, or piece of software accomplishes a task with very little (if any) insight into exactly how it does so.

Reverse engineering is applicable in the fields of computer engineering, mechanical engineering, design, electronic engineering, software engineering, chemical engineering, and systems biology.

2.2 Verified API Calls

Only verified API calls are been included in the application. Verified API call is a call from a trusted server with the implemented protocol. The untrusted protocol which is calling the API from the application is blocked. The trusted server has their predefined protocols implemented. During the malware detection, if any API is hitting with different protocol and untrusted API call it is immediately blocked from execution, as it can be suspicious for the malware attack.

2.3 signed APK / Trusted certificates

Android requires that all APKs be digitally signed with a certificate before they are installed on a device or updated. The signed and unsigned APK are exactly the same except the signed APK has some extra files that indicates the APK is signed. To generate signed APK, you just run the JDK jar signer tool on the unsigned APK, the results is a new APK file but contains some new files under the folder META-INF.

2.4 Applications allowed permissions combination

Android application one manifest file which has metadata of the application. The manifest file contain the all information about classes, services, permissions, broadcast receiver, versions, Gradle etc. All the required permissions of the application are return inside manifest file. If any of the permissions combination written inside this file is not related to application functionality, or it is found that those permissions are never been used in the application runtime, then such permissions are blocked permanently and are deleted from the code base.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 11 Issue 08, August-2022

REVIEW TABLE

Sr No.	Title	Techniques	Future Scope	Conclusion
1	Reducing Android Malware And Inefficiency By Detecting Defective And Dummy Applications Using Neural Networks	- Data visualization and Pre-processing - Model Creation - Implementation	-	Due to android operating system being open source, it is highly anticipated that attackers will keep finding loop holes in the system and the private data will always be at their disposal but with improving accuracy of detection and classification algorithm, it can be true in near future that android operating systems come with an in-built malware detection scheme which might be able top the existing detection models.
2	A Review on The Use of Deep Learning in Android Malware Detection	- Android Application Components - Android Malware Detection Techniques - Static Analysis - Dynamic Analysis - Hybrid Analysis	Future work may consider dynamic research techniques or utilizing hybrid analysis techniques. Sharing research datasets and tools between researchers lingered unaddressed except in a few cases. Hardening deep learning models against different adversarial attacks and detecting, describing and measuring concept drift are vital in future work in Android malware detection.	In this work, we presented a thorough review of the use of deep learning in Android malware detection. A comparison of existing work with respect to certain criteria was presented. The review uncovered knowledge gaps in the existing work and underscores major challenges and open issues that will direct future research Abdelmonim Naway et al, International Journal of Computer Science and Mobile Computing, 56 efforts.
3	Android Application Security Scanning Process	- Feature extraction - Static analysis - Dynamic analysis - Ransomware Detection -	-	This chapter highlighted the booming of Android technologies and their applications which make them more attractive to security attackers. Recent statistics of Android malwares and their impact were presented. Additionally, this chapter has provided the main phases required to apply security scanning to Android applications. The purpose is to protect Android users and their devices from the threats of different security attacks. These phases include the way of downloading Android apps, decoding them to generate the source code, and how this code is screened to extract the required features to apply either static analysis or dynamic analysis or both
4	Android Malware Detection by Using Random Forest Algorithm	- APK unzipping - Extracting permissions - Applying random forest algorithm - Matching with data set		We can conclude that applications can be altered easily by changing their permissions. Hence we need to test any android .APK file before installing it. This application will test every .APK file and give results on the basis of the algorithm.
5	Mobile Malware Detection: A Survey	- Signature based techniques - Static Analysis - Dynamic Analysis - Android Malware Detection Techniques	In future, work a detailed study with the most effective tools to detect mobile real-time threads.	With the developing utilization of Smartphone, the quantity of assaults and dangers are additionally on increment. It is important to give security to end clients from dangers. In this paper, we represent a full picture about malware environment as discussing malware classes and techniques there are different techniques have been discussed and listed. Papers also mention Android malware detection types, methods, technologies and proposed techniques. In above section we have studied various algorithms,

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 11 Issue 08, August-2022

				which restrict the detection of attacks.
6	Optimizing Android Malware Detection Via Ensemble Learning	- Data collection - Feature Extraction - Model testing and performance evaluation		Random Forest produced the best base detection model, having a true positive detection rate of 97.9%, false positive detection rate of 0.19%, accuracy of 98%, and a detection error rate of 0.2%. The Majority Vote combination rule produced an ensemble model with a true positive malware detection rate of 98.1%, false positive detection rate of 0.18%, a detection accuracy of 98.2%, and a detection error rate of 0.18%. The ensemble Model outperformed the single model with a relative difference of 0.2% on the true positive detection rate. The ensemble model has a very low false alarm rate of 0.18% and the lowest error rate of 0.18%. The study therefore concludes that a supervised ensemble model is an effective approach for the anomaly detection of Android malware.

CONCLUSION

During the process of Android malware detection different techniques are used to detect the malware in android application, and amongst those techniques it has found that Application malware detection using random forest algorithm gives the best results.

REFERENCES

[1] AV-Comparatives, "Mobile Security Review 2018." [Online]. Available: https://www.avcomparatives.org/tests/mobile-security-review2018/#google-play-protect.
[2] Fawcett, T. (2006). An introduction to ROC analysis. Journal of Pattern Recognition Letters, 27(8), 861–874. https://doi.org/10.1016/j.patrec.2005.10.010
[3] Powers, D. M. W. (2011). Evaluation: From Precision, Recall and F-Measure to Roc, Informedness, Markedness & Correlation. Journal of Machine Learning Technologies, 2(1), 37–63.
[4] M. Sun, M. Li, and J. C.S. Lui, "DroidEagle: Seamless detection of visually similar Android apps," in Proceedings of the 8th ACM Conference on Security & Privacy in Wireless and Mobile Networks, 2015, pp. 9
[5] C. Hasegawa and H. Iyatomi, "One-dimensional convolutional neural networks for Android malware detection," in 2018 IEEE 14th International Colloquium on Signal Processing & Its Applications (CSPA), 2018, no. March, pp. 99–102.
[6] L. Shiqi, T. Shengwei, Y. Long, Y. Jiong, and S. Hua, "Android malicious code Classification using Deep Belief Network," KSII Trans. Internet Inf. Syst., vol. 12, no. 1, pp. 454–475, 2018.
[7] E. M. B. Karbab, M. Debbabi, A. Derhab, and D. Mouheb, "MalDozer: Automatic framework for Android malware detection using deep learning," Digit. Investig., vol. 24, no. March, pp. S48–S59, 2018.
[8] Xiang Li, Jianyi Liu, YanyuHuo, Ru Zhang, Yuangang Yao "An Android malware detection method based on AndroidManifest file" 2016 4th International Conference on Cloud Computing and Intelligence Systems, 19 Aug 2016
[9] Hyo-Sik Ham, Mi-Jung Choi "Analysis of Android malware detection performance using machine learning classifiers" 2013 International Conference on ICT Convergence, 16 Oct. 2013
[10] Patrick P. K. Chan, Wen-Kai Song "Static Detection Of Android Malware by Using Permissions and Api Calls" 2014 International Conference on Machine Learning and Cybernetics, 16 July 2014
[11] https://www.idc.com/promo/smartphone-market-share/os [Accessed 6 September 2018]
[12] Joshi, P.; Jindal, C.; Chowkwale, M.; Shethia, R.; Shaikh, S. A. & Ved, D. Protego: A passive intrusion detection system for Android smartphones Computing, Analytics and Security Trends (CAST), International Conference on, 2016, 232-237
[13] Mohata, V. B.; Dakhane, D. M. & Pardhi, R. L. Mobile Malware Detection Techniques International Journal of Computer Science & Engineering Technology (IJCSET), 2013, 4, 2229-3345
[14] Statista. (2018). Smartphone OS global market share 2009-2018 | Statistic. Retrieved June 26, 2018, from https://www.statista.com/statistics/266136/global-market-share-held-bysmartphone-operating-systems/
[15] Narudin, F. A., Feizollah, A., Anuar, N. B., & Gani, A. (2016). Evaluation of machine learning classifiers for mobile malware detection. Soft Computing, 20(1), 343–357.https://doi.org/10.1007/s00500-014-1511-6
[16] Feng, Y., Anand, S., Dillig, I., & Aiken, A. (2014). Apposcopy: Semantics-Based Detection of Android Malware Through Static Analysis. In Proceedings of the ACM SIGSOFT International Symposium on Foundations of Software Engineering (FSE'14) (pp. 1622)https://doi.org/10.1145/2635868.2635869
[17] Lueg, C. (2018). Malware figures for Android rise rapidly. Retrieved August 24, 2018, from https://www.gdatasoftware.com/blog/2018/07/30937-malware-figures-for-androidrise-rapidly