Manju Bala et al. International Journal of Recent Research Aspects ISSN: 2349-7688, Vol. 4, Issue 4, Dec 2017, pp. 25-27

Android Malware Classification using Neural Nets

Manju Bala

I. P. College for Women, University of Delhi, Delhi, India

***Abstract*— As indicated by AV merchants vindictive programming has been developing exponentially years ago. One of the principle purposes behind these high volumes is that all together to sidestep discovery, malware creators began utilizing polymorphic and transformative procedures. Therefore, conventional mark based ways to deal with recognize malware are being lacking against new malware and the classification of malware tests had turned out to be basic to know the premise of the conduct of malware and to battle back cybercriminals. Amid the most recent decade, arrangements that battle against pernicious programming had started utilizing machine learning approaches. Tragically, there are few open source datasets accessible for the scholarly group. One of the greatest datasets accessible was discharged a year ago in an opposition facilitated on Kaggle with information gave by Microsoft to the Huge Information Trailblazers Social event (Huge 2015). This proposition presents two novel and adaptable methodologies utilizing Neural Systems (NNs) to dole out malware to its comparing family. On one hand, the principal approach makes utilization of CNNs to take in a include pecking order to segregate among tests of malware spoke to as dark scale pictures. Then again, the second approach utilizes the CNN engineering acquainted with order malware tests concurring their x86 guidelines. The proposed strategies accomplished a change of 80.86% and 81.56% as for the equivalent likelihood benchmark.**

Keywords—component; formatting; style; styling; insert (key words)

I. INTRODUCTION

Malware is nothing but just malicious software. It is intended to spy someone's personal Data without permission of owner. Nowadays due to increase in Smartphone's around the World we need to customize over device's security in order to prevent it from attackers. There is various kind of malware including key loggers, spyware, Trojan horse, virus, worms, Ransom-ware etc. Malware is basically designed for damaging or any illegal action on a system such as collecting sensitive information, to get access control, interrupting CPU operations, advertisement and display that information to distant hacker.

Generally, malware's design is based on its creator intent rather than actual features. Nowadays, Malware creation is on boom due to lure of money in this work through organized internet crime. Nowadays, malware is fabricated to get advantages using advertisement (adware), stealing sensitive information (spyware), email spam or child pornography (zombie computers), to extort money (ransom ware).

The sole point of this proposition is to detect & classify the unknown malware into its respective categories. In this thesis we are using neural network back propagation algorithm in order to classify the respect malware . The main objective is to analyze the .apk file of android application. The .apk file is passed through different disassemble tools to find its manifest.xml file and source code.dex file, after that we use another tools to parse the manifest.xml file and .dex file. Here we are employing two basic analysis strategies. Static examination and Dynamic investigation, where static examination manages the required consents in show document and touchy Programming interface utilized as a part of that application, though powerful investigation manage the dynamic conduct of use like stream of data, Capacities utilized and Programming interface utilized. To concentrate dynamic conduct, we have to run the application in a controlled execution condition.

We have as of now observed many machine learning systems to distinguish and characterize the malware like regression, decision trees, SVM, CNN. We can improve execution in the event that we utilize deep learning notwithstanding neural net strategies. The calculation execution increment up to 96%, much superior to different methods.

II. METHODOLOGY

We have utilized the distinctive classification strategies and utilize the Android application Permissions and APIs as the elements for order display. We have gathered the 440 test of the malware and the generous application, utilized the Androguard and Mosf to extricate the consent and Programming interface from the android application bundle in particular .apk document and made a dataset for the characterization model, for example, Decision tree, Neural Systems.

III. IMPLEMENTATION

The android application bundle .apk record are handled with Androguard to remove the highlights Android authorization and APIs and gathered the consent to make a dataset what's more, aggregate have the 330 elements. We have utilized the Quick Excavator apparatuses the python Scikit-learn and Perfect python to play out the order on the element dataset of the authorizations and APIs. Presently we will talk about outcomes and assessment.

IV. RESULTS

We had connected the Decision tree, Neural Systems characterization strategies, we had utilized the dataset with respect to the preparation and testing is finished 10 cross

overlap. We talk about the aftereffect of every method one by one and they analyze them with the Flawless neural systems.

Decision Trees

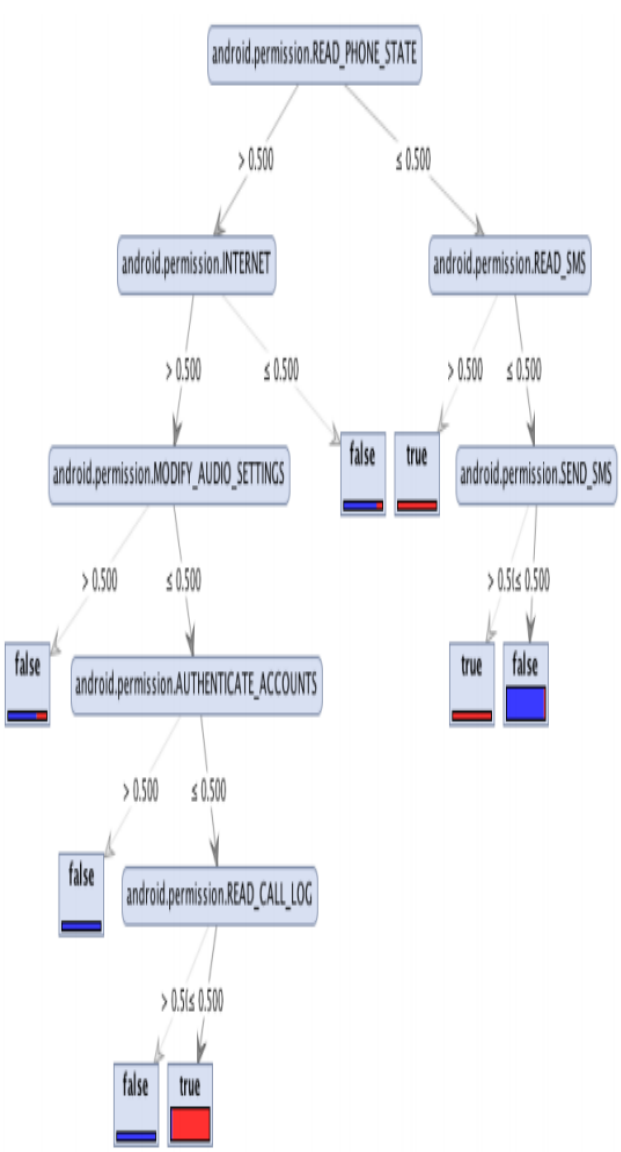

Fig. 1. Decision Trees with Gini Index

TABLE 1. RESULTS USING DECISIO TREES USING GINI INDEX

Item NO.	Accuracy	Precision	Recall	F1_Score
1	74.4	69.9917557932	74.4	70.3844038718
2	74.4	61.3957375479	74.4	64.5601616162
3	62.0	62.1556709957	62.0	57.0238222971
4	77.5	78.8874895572	77.5	77.1946742532
5	71.3	66.3924006075	71.3	64.3761920373
6	83.7	82.068226817	83.7	75.769352166
7	62.0	60.1655723906	62.0	55.311356503

Support Vector Machines

TABLE 2. RESULTS USING LINEAR SUPPORT VECTOR MACHINES

Item No.	Accuracy	Precision	Recall	F1_score
1	71.3	72.1147005772	71.3	70.472194211
2	74.4	82.2106165171	74.4	70.2916233766
3	71.3	55.9106636156	51.3	62.024477738
4	65.1	60.4371505687	65.1	61.1400462154
5	65.1	54.4060121038	65.1	58.0587163268
6	77.5	66.7503839467	77.5	68.7402424286
7	62.0	63.2480519481	62.0	59.7253316503

Neural Systems

TABLE 3. RESULTS USING NEURAL NETWORKS

Item No.	Accuracy	Precision	Recall	F1_score
1	88.0	101.523809524	88.0	89.6727272727
2	56..0	50.4761904762	56.0	50.819047619
3	72.0	93.9487179487	72.0	76.2666666667
4	32.0	45.6857142857	32.0	35.0099865047
5	72.0	61.3333333333	72.0	57.8517482517
6	72.0	85.3333333333	72.0	75.7333333333
7	96.0	99.3841269841	96.0	95.9327731092

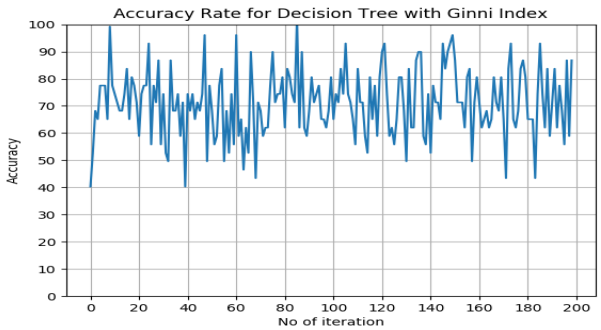

Fig. 2. Accuracy Rate for Gini Index based Decision Trees

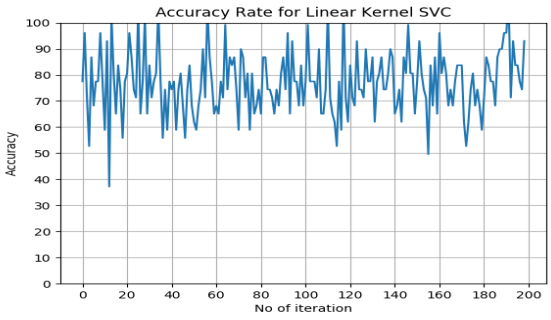

Fig. 3. Accuracy Rate for Linear Kernel Based SVM

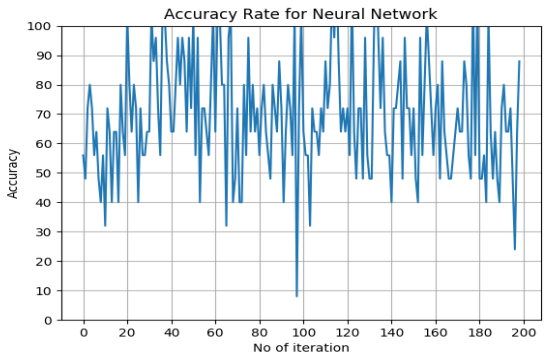

Fig. 4. Accuracy Rate of Neural Network

So based on the provided results, on the predefined dataset neural networks works better and we achieve 81% success by using it. Whereas other techniques like decision tree and support vector machine produces results nearly 60% and 75% simultaneously. Neural networks outperform other algorithms like logistic and linear regression, but in order to increase its efficiency more we have to apply more number of hidden layers. Such that model presented by this neural networks works better.

V. REFERENCES

[1] Wikipedia, "Google Play — Wikipedia, the free encyclopedia." https://en.wikipedia.org/wiki/Google_Play.

[2] Androguard 2017 https://www.darknet.org.uk/2016/11/androguard-reverse-engineering-malware-analysis-for-android/

[3] Android services https://developer.android.com/guide/components/services.html.

[4] Wikipedia, Android application package. https://developer.android.com/reference/android/app/package-summary.html

[5] APK Tool- https://github.com/iBotPeaches/Apktool

[6] Android dataset permissions –Kaggle https://www.kaggle.com/xwolf12/datasetandroidpermissions

[7] IEEE malign or benign dataset permission android . https://ieee-dataport.org/documents/dataset-malwarebeningn-permissions-android

[8] Yearly analysis for the sale of android and windows sales http://www.zdnet.com/article/canalys-predicts-a-billion-android-smartphone-sales-in-2017-and-rapid-growth-for-windows-phone/

[9] Usage share of android operating systems https://en.wikipedia.org/wiki/Usage_share_of_operating_systems

[10] Security report over malware by alcatel leucent https://resources.alcatellucent.com/theStore/files/Kindsight_Security_Labs_Q112_Malware_Report_EN.pdf

[11] Android developer security best training https://developer.android.com/training/articles/security-tips.html

[12] X. Jiang, "An Evaluation of the Application ("App") Verification Service in Android 4.2" https://www.csc2.ncsu.edu/faculty/xjiang4/appverify/

[13] G. Canau, M. Buhu, C. Oprisa, "Malware Classificatios using Filesytem Foot Prints", IEEE conf. on Automation, Quality and Testing, Robotics (AQTR), May 2016.

[14] A. Mylonas, S. Dritsas, B.Tsoumas, D. Gritzalis, "Smartphone Security Evaluvation The Malware Case", in Proc. Of Internaltion Conf. on Security and Cryptography (SECRYPT), Spain, Feb. 2014.

[15] D.G.N.B. Mejia, G.S. Perez, L.K.T. Medina, " Third Intl. conf. on Digital Information Processiing, Data Mining, ad Wireless Communications (DIPDMWC), Russia, Aug. 2016

[16] J. LI, L. Zhai, X. Zang, D. Quan, "Research of Android Malware Detection on Network Traffuc Monitoring", China, Oct. 2014.

[17] W.C. Hiseh, C.C. Wu, Y.W. Kau, "A study of Android Malware Detection Technology Evolution", Intl. Conf. on Security Technology, Taiwan, Jan 2016.

