ANDROID MALWARE ANALYSIS USING APPLICATION PERMISSIONS

ZAID KHALID HAMADI

MASTER OF COMPUTER SCIENCE

(INTERNETWORKING TECHNOLOGY)

2015

Faculty of Information and Communication Technology

ANDROID MALWARE ANALYSIS USING APPLICATION PERMISSIONS

Zaid Khalid Hamadi

Master of Computer Science (Internetworking Technology)

2015

ANDROID MALWARE ANALYSIS USING APPLICATION PERMISSIONS

ZAID KHALID HAMADI

A thesis submitted
in fulfillment of requirements for the degree of Master of Computer Science
(Internetworking Technology)

Faculty of Information and Communication Technology

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

2015

DECLARATION

I declare that this thesis entitle "Android Malware Analysis using Application Permissions" is the result of my own research except as cited in the references. The thesis has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.

Signature : ……..……

Name : Zaid Khalid Hamadi

Date : February 02, 2015

APPROVAL

I hereby declare that I have read this thesis and in my opinion this thesis is sufficient in terms of scope and quality for the award of the degree Master of Computer Science (Internetworking Technology).

Signature	:	…….
Supervisor Name	:	Assoc. Prof. Dr. Burairah Bin Hussin
Date	:	February 02, 2015

DEDICATION

I would like to dedicate my hard work fruit to those who did not stop their daily support since I was born. They never hesitate to provide me all the facilities to push me forward as much as they can.

To the symbol of love who suckled me her love and compassion:

My ***MOTHER…*** my precious diamond

(God's mercy on her soul, and make her lives in his eternal paradise).

To the person who spend many years working hard to make me grown up:

My ***FATHER…*** who has the big heart

Assoc. Prof. Dr. Khalid Hamadi Sharaf.

To the woman who provided all the suitable circumstances for me that lead me to this success that is made by my own hands:

My ***WIFE…*** my soul mate,

the first and last love.

To my ***BROTHERS, SISTERS***, I thank them all for their support, and I thank them for being my family.

To my precious kids **"Usama & Rokaya"**, whose smiles give me passion and strength.

This work is a simple and humble reply to their much kindness and goodness I have taken over during my study time.

Zaid Khalid Hamadi

إهـــداء

أود أن أهدي ثمرة عملي الشاق لأولئك الذين لم يكفوا عن دعمهم اليومي منذ أن ولدت. لأنهم لم يترددوا في أن يقدموا لي كل التسهيلات لدفعي إلى الأمام بقدر ما يستطيعون.

إلى رمز المحبة التي أرضعتني الحب والرحمة:

أمي ... ياجوهرتي الثمينة

(رحمها الله واسكنها جنة الخلد)

إلى الشخص الذي قضى الكثير من سنوات عمره بالعمل الشاق ليجعلني أكبر:

أبي ... ياصاحب القلب الكبير

أ. م. د. خالد حمادي شرف

إلى المرأة التي لم تكف عن تهيئة كل الظروف التي قادتني الى هذا النجاح الذي هو من صنيع يدي:

زوجتي ... يا توأم روحي

وحبي الأول والأخير.

إلى ***أخوتي*** و ***أخواتي***، أنا أشكركم جميعاً لكل الدعم الذي قدمتوه لي واشكركم لأنكم عائلتي.

وإلى طفلي الغاليين ***أسامة*** و ***رقية*** اللذان اعطياني بابتسامتهما القوة و الاصرار على مواصلة دراستي.

فهذا العمل هو رد بسيط ومتواضع لكم مقابل ما حصلت عليه من الدعم والخير واللطف الكثير منكم خلال فترة دراستي.

زيد خالد حمادي

ABSTRACT

Smartphones are the most useful devices nowadays because they offer a lot of useful services besides the aspect of mobility that benefit the user even more. In addition, the most popular platform is Android, because it offers verity of thousands free applications and also because the platform is open source. In this case anybody can develop an application and then publishing it on the store. In this research, we are aiming to analyze 400 Android application samples taken from Google's play store, in order to determine the percentage of having the malware behavior within the collected samples. A confirmed malware dataset will be collected as well and the analysis will be done in order to derive malware patterns (permissions) and then comparing the 400 application samples with the malware derived malware patterns based upon the permissions requested. However, a certain combination of some Android user permissions could create a malware behavior such as the ability to read user contacts and the permission of using the web browser. At this point we can determine that this application has a malware behavior, which can send the user contacts to a third-party server without the knowledge of the user, but this is needed to be confirmed by analyzing the application's source code. After doing the analysis, we will be able to propose a framework to protect the user private data that will benefit the users and the application developers to avoid designing an application that request such dangerous permissions combination if possible.

ABSTRAK

Telefon pintar adalah peranti terkini yang menawarkan banyak perkhidmatan yang berfaedah di samping aspek mobiliti yang memberi manfaat kepada pengguna. Platform yang paling popular adalah Android, kerana ia adalah platform sumber terbuka dan menawarkan pelbagai aplikasi telefon pintar secara percuma. Dalam hal ini sesiapa sahaja boleh membangunkan aplikasi telefon pintar dan kemudiannya memuatnaik aplikasi ini. Oleh yang demikian, kemungkinan untuk seseorang memuatnaik aplikasi malware yang bertujuan jahat adalah sangat tinggi. Kajian ini menyasarkan untuk menganalisis sampel aplikasi android melalui corak malware berdasarkan kebenaran yang diminta oleh aplikasi telefon pintar. Ini adalah kerana gabungan tertentu beberapa kebenaran pengguna Android boleh mewujudkan tingkah laku malware seperti kebolehan untuk membaca maklumat data pengguna. Kajian ini menganalisis 400 sampel aplikasi Android yang diambil dari kedai permainan Google untuk menentukan peratusan yang mempunyai tingkah laku malware dalam sampel. Walaubagaimana pun analisa ini tidak dapat membuktikan bahawa terdapat malware di dalam aplikasi telefon pintar sehingga perlu disahkan oleh menganalisis kod sumber aplikasi. Selepas melakukan analisis, kita akan dapat mencadangkan satu rangka kerja untuk melindungi data peribadi pengguna yang akan memberi manfaat kepada pengguna dan pembangun aplikasi telefon pintar.

ACKNOWLEDGEMENT

First and foremost, praise be to Allah, for giving me this opportunity, the strength and the patience to complete my thesis finally, after all the challenges and difficulties. I would like to thank my supervisor, ***Assoc. Prof. Dr. Burairah Bin Hussin*** for his high motivation and most significant contribution in this thesis.

I would also like to thank Ministry of Higher Education and Scientific Research of IRAQ, all UTeM staff and Malaysian people and ***Dr. Abdul Samad Bin Shibghatullah*** and ***Dr. Mohd Sanusi bin Azmi.*** Furthermore, I want to thank my friends who have helped and motivated me throughout. May Allah reward them all abundantly, Sincere thanks to all.

Zaid Khalid Hamadi

TABLE OF CONTENTS

PAGE

DECLARATION
DEDICATION
ABSTRACT i
ABSTRAK ii
ACKNOWLEDGEMENT iii
TABLE OF CONTENTS iv
LIST OF TABLES vii
LIST OF FIGURES viii
LIST OF EQUATIONS xi
LIST OF SYMBOLS xii

CHAPTER 1 **1**
1 INTRODUCTION **1**
1.0 Background 1
1.1 Introduction 1
1.2 Research background 2
1.3 Research Problem 3
1.4 Research Questions 4
1.5 Research Objectives 4
1.6 Research Scope 5
1.7 Research Significance 5
1.8 Thesis texture 6
1.9 Summary 8

CHAPTER 2 **9**
2 LITERATURE REVIEW **9**
2.0 Introduction 9
2.1 Malware History Review 9
2.1.1 DOS and Windows Malwares 9
2.1.2 First Mobile Malware on Symbian OS 11
2.1.3 First Mobile Malware on iOS 12
2.1.4 First Mobile Malware on Android OS 13
2.2 Threats objectives 15
2.3 Vectors of the attacks 17

2.4 Android sensitive data sources 21
2.5 Android Application Components 23
2.6 Types of Android Threats 25
2.6.1 Spyware 25
2.6.2 Trojans 26
2.6.3 Backdoor 27
2.6.4 Drive-by-download 27
2.6.5 Premium-Rate SMS sender 28
2.6.6 Root exploits 28
2.6.7 Botnet 29
2.6.8 Virus 30
2.6.9 Worm 31
2.7 Types of Android Security Mechanisms 32
2.7.1 Sandbox 32
2.7.2 Application signing 33
2.7.3 Anti-virus applications 34
2.7.4 Google Bouncer 34
2.7.5 Remote kill switch 35
2.7.6 Permissions 35
2.8 Categorizing Android Permissions 36
2.9 Summary 37

CHAPTER 3 38
3 METHODOLOGY 38
3.0 Introduction 38
3.1 Research Methodology 38
3.1.1 Stage 1: Data collection 40
3.1.2 Stage 2: Data analysis 42
3.1.3 Stage 3: Derive malware patterns 43
3.1.4 Stage 4: Proposing the framework 43
3.1.5 Stage 5: discuss the limitations and future work 44
3.2 Summary 45

CHAPTER 4 **46**
4 DATASETS COLLECTION & ANALYSIS **46**
4.0 Introduction 46
4.1 Data collection 46
4.1.1 Google's play store market dataset collection 47
4.1.2 Malware dataset collection 50
4.2 Permission types and filtering 52
4.3 Storing collected datasets 53
4.4 Data analysis 57
4.4.1 Statistical analysis of both datasets 57
4.4.2 Analyzing requested permissions 62
4.4.3 Discovering dangerous permission patterns 64
4.5 Summary 71

CHAPTER 5 **74**
5 A FRAMEWORK FOR ANDROID TO PROTECT THE USER **74**
5.0 Introduction 74
5.1 Proposing the framework 74
5.1.1 User steps by step framework 75
5.1.2 Illustrating the framework 75
5.2 Summary 77

CHAPTER 6 **78**
6 CONCLUSION & FUTURE WORK **78**
6.0 Introduction 78
6.1 Problems and Limitations 78
6.2 Recommendations and future work 81
6.3 Conclusion 82

REFERENCES **83**

APPENDIX **91**

LIST OF TABLES

TABLE	TITLE	PAGE
Table 2.1:	History of mobile malwares from 2004 until 2012.	15
Table 4.1:	Average, maximum, and minimum number of requested permissions from our market and malware datasets.	58
Table 4.2:	The 14 highest requested permissions in our market dataset.	60
Table 4.3:	The 14 highest requested permissions in our malware dataset.	61
Table 4.4:	Permissions requested and used by our malware dataset only.	63
Table 4.5:	Statistics of Counter-Clank malware family in our datasets.	65
Table 4.6:	Statistics of Geinimi malware family in our datasets.	66
Table 4.7:	Statistics of Gold-Dream malware family in our datasets.	67
Table 4.8:	Statistics of Droid-Dream malware family in our datasets.	68
Table 4.9:	Three applications detected in our market dataset with the same combination of permissions as Droid-Dream malware family.	69
Table 4.10:	Statistics of SMS-advertiser malware family in our datasets.	70
Table 4.11:	The all five malware families' table shows the relation and shared permissions between them. The numbers 1 to 5 represents the malware families, 1: Counter-Clank, 2: Geinimi, 3: Gold-Dream, 4: Droid-Dream, and 5: SMS-advertiser.	72

LIST OF FIGURES

FIGURE	TITLE	PAGE
Figure 1.1:	A sample of an Android application permissions screenshot.	2
Figure 1.2:	Worldwide Smartphone OS Market trend.	4
Figure 2.1:	Boza malware message screenshot.	10
Figure 2.2:	effect of Marburg on Windows.	11
Figure 2.3:	Cabir notification window screenshot.	12
Figure 2.4:	Cabir notifications at the victim device screenshot.	12
Figure 2.5:	Screenshot of Ikee worm on iOS.	13
Figure 2.6:	First malware infected application on Android OS and its requested permissions screenshots.	14
Figure 2.7:	288 malware side effects on smartphones.	17
Figure 2.8:	Mobile-to-Computer attack using USB.	20
Figure 2.9:	Computer-to-Mobile attack using USB.	20
Figure 2.10:	Mobile-to-Mobile attack using crafted USB.	20
Figure 2.11:	Crafted USB cable with 2 Micro B USB connector dongles.	21
Figure 2.12:	Types of user's sensitive data that accessed by an application.	23
Figure 2.13:	A screenshot shows an option to enable unknown sources applications to be installed.	26
Figure 2.14:	Drive-by-download update malware installer screenshot.	27

Figure 2.15: What happened when installing an application that request a super user access. 29
Figure 2.16: (Bowling Time) game screenshot infected by Droid-Dream malware. 30
Figure 2.17: Mobile threats type percentages from 2004 to 2011. 32
Figure 3.1: Proposed methodology stages chart. 39
Figure 3.2: A screenshot of the Google's play store categories. 40
Figure 3.3: Contagio Mobile malware samples website screenshot (http://contagiominidump.blogspot.com). 42
Figure 4.1: Flow chart for data collection process summarization. 47
Figure 4.2: Raccoon V3.1 tool user interface. 48
Figure 4.3: Viewing the details of applications in Raccoon tool. 49
Figure 4.4: How apktool works screenshot. 51
Figure 4.5: Contents of Manifest file screenshot. 52
Figure 4.6: Snapshot of different types of permissions presented by the Raccoon tool. 52
Figure 4.7: Stored market dataset screenshot in MS office word 2013. 54
Figure 4.8: A screenshot of our market apps information database in MS office Access 2013. 55
Figure 4.9: A screenshot of our market apps permissions database in MS office Access 2013. 55
Figure 4.10: A screenshot of the MS office word 2013 malware dataset. 56
Figure 4.11: A screenshot of the malware apps permissions database in MS office Access 2013. 57
Figure 4.12: Permissions analysis chart of our market dataset. 58
Figure 4.13: Permissions analysis chart of our malware dataset. 59

Figure 4.14: Statistics of both datasets malware and market in terms of the highest requested permissions. 62

Figure 4.15: Illustrating the relation between the five malwares in terms of requested permissions. 73

Figure 5.1: Framework chart for Android users to protect their private data. 76

Figure 6.1: Virussign.com website for malware samples. 79

Figure 6.2: My email messages trying to get the malware samples from Virussign.com. 80

LIST OF EQUATIONS

EQUATION	TITLE	PAGE
Equation 4.1:	percentage of each requested permission in our datasets.	59

LIST OF SYMBOLS

SYNOPSIS		STAND FOR
3G	:	Third generation network.
4G	:	Fourth generation network.
C&C	:	Command and Control Server.
DoS	:	Denial-of-Service.
DOS	:	Disk Operating System.
GPS	:	Global Positioning System.
IMEI	:	International Mobile Equipment Identity.
IMSI	:	International Mobile Subscriber Identity.
IP	:	Internet Protocol.
MAC	:	Media Access Control.
MMS	:	Multimedia Message Service.
PUP	:	Potentially Unwanted Programs.
SMS	:	Short Message Service.
SSH	:	Secure Shell.
UI	:	User Interface.
UID	:	User Identifier.
URL	:	Uniform Resource Locator.

CHAPTER 1

INTRODUCTION

1.0 Background

In this chapter, we are permitted to provide a clear picture about this research by listing a bunch of important things such as the introduction, research problem, research objectives, research questions, and so on, as we will see further. Each one on this list will be described and explained clearly.

1.1 Introduction

Smartphones are the most useful devices nowadays because they offer a lot of useful services besides the aspect of mobility that benefit the user even more. The uses of these devices, increasing now a days by the people. It helps the user to input and store their critical information into their devices which creates an obvious privacy issue that can be happened through the device stolen or by trying to hack the user and collect his critical data such as the contacts, banking information, messages, and important user data… etc... In this work we are attempting to study the security types as well as the types of threats that can be made in the literature review chapter.

1.2 Research background

Referring to figure 1.1 below that shows a sample screenshot of an Android application permissions. This screenshot is taken before installing Facebook application that is published at Google's play store. As we can see, the first permission that required to accept is to give the ability to access the messages by this app which can read both SMS and MMS. The second permission is to give the app the ability to access your personal information which can add, modify, and delete calendar events and also has the ability to send emails to anyone without the knowledge of the host. The third permission is to give the ability to access and connect or disconnects the Wi-Fi network access and so on (Matsudo et al. 2012).

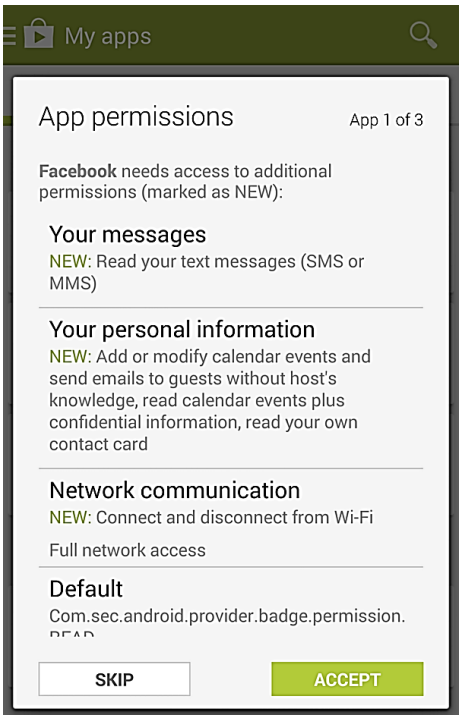

Figure 1.1: A sample of an Android application permissions screenshot.

1.3 Research Problem

Android platform OS application development getting popular by apps developer communities because of the increasing trend of using the Android platform. Referring to figure 1.2, we can identify the increasing of the percentage trend of using Android platform comparability with other platforms, such as iOS, Windows phone, and so on (International Data Corporation, 2014). This is why we choose the Android platform to work on in this study. Also, there is another reason to choose this platform to work on is the increasing of the malware applications since the first detected malware in 2010, then the percentage rise to 13% in only 14 months (Pieterse & Olivier 2012).

Android applications involved with HTML web apps, Facebook, and android apps included, in this case a third-party development are allowed. These third party apps are prone to malicious behavior, such as malware that treats the security of this platform. The security threats are coming from the permissions that given from the user for these third-party apps when they permitting to download and install on the android platform, in this case we can say that a third-party applications could include a malware behavior that integrated with the app. There are approximately 137 permissions (Holavanalli et al. 2013) can be used for these apps which are very challenging to the users and the developers as well. A certain combination of some permissions could create a malware behavior such as the ability to read user contacts and the permission of using the web browser. At this point we can notice that this app can send the user contacts to an external server that is predefined in the application source code. In this case it sends them without the user's knowledge or he is not notified to do this action. In this case, these activities considered as a user's privacy attack (Stevens et al. 2013).

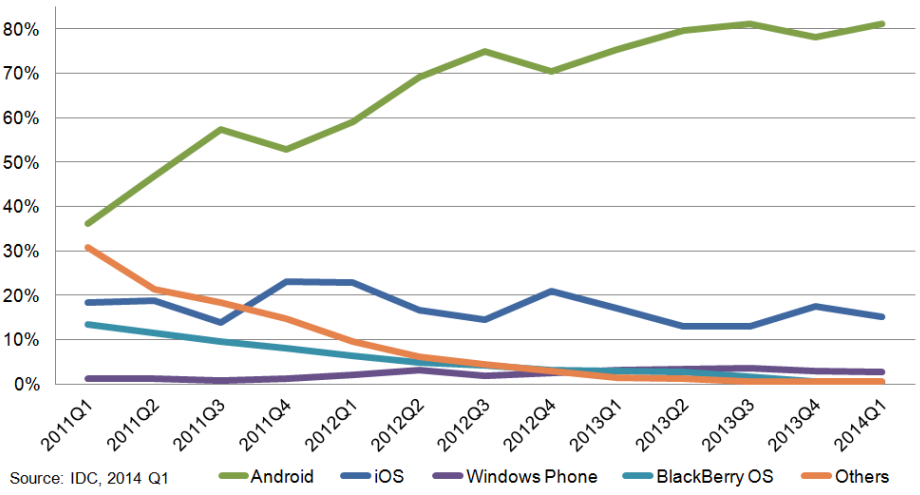

Figure 1.2: Worldwide Smartphone OS Market trend.

1.4 Research Questions

The study is permitted to clarify a number of questions that need to be solved in order to gain the achievements that listed above

1. How to conduct a security study on android platform to secure the user?
2. How to evaluate the results of the analysis?
3. How to identify the malware dataset?

1.5 Research Objectives

This study aims to achieve a number of objectives that clearly described in the following:

1. Identify the types of threats and types of security for the Android platform.
2. Analyze the dataset using permissions base aspect to detect the malware applications.
3. Classify the threats for Android security.

1.6 Research Scope

This research has a scope on the permissions that are presented to the user before downloading and installing any application from Google's play store. These permissions were given from the user to an application are the access rights to hardware or software of the smartphone components. In this research, we will collect a dataset from Google's play, store and then making an analysis based on the permissions for each application and then we will collect other dataset which belongs to confirmed malware applications by security companies, which will be analyzed also to determine the permissions that are requested from a malicious application. Then we will make a comparison between the two datasets in order to identify the percentage of possible malwares in the collected dataset from Google's play store. The next step is to propose a framework based on the generated knowledge from the analysis and the comparison between the two datasets to prevent any privacy threat which benefit the user as well as application developers to avoid building an application that request such malware permissions.

1.7 Research Significance

The importance of this study will be as the following:

1. Enhance the consistency of this research by proposing a methodology to do this study in terms of organizing the work flow.
2. In chapter two which review the related works and the purpose of the study will provide a very good knowledge of this research importance.
3. Providing very well practices of collect and analyze such datasets.
4. Presenting the knowledge of malware application behavior on this platform which benefits the users and the application developers as well.