

**DYNAMIC ANALYSIS OF ANDROID MALWARE USING DROIDBOX**

A Thesis  
Submitted to the Graduate School  
of  
Tennessee State University  
In  
Partial Fulfillment of the Requirements  
for the Degree of  
Master of Science  
in  
Computer, Information and Systems Engineering

Graduate Research Series No. \_\_\_\_\_

Priya Chaurasia

November 9, 2015

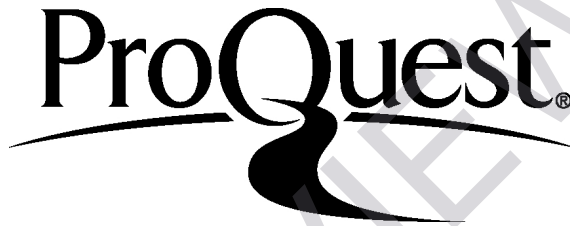
ProQuest Number: 10003138

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 10003138

Published by ProQuest LLC (2016). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code  
Microform Edition © ProQuest LLC.

ProQuest LLC.  
789 East Eisenhower Parkway  
P.O. Box 1346  
Ann Arbor, MI 48106 - 1346

# **DYNAMIC ANALYSIS OF ANDROID MALWARE USING DROIDBOX**

A Thesis  
Submitted to the Graduate School  
of  
Tennessee State University  
In  
Partial Fulfillment of the Requirements  
for the Degree of  
Master of Science  
in  
Computer, Information and Systems Engineering

Priya Chaurasia

November 9, 2015

To the Graduate School:

We are submitting a thesis by Priya Chaurasia entitled “DYNAMIC ANALYSIS OF ANDROID MALWARE USING DROIDBOX”. We recommend that it be accepted in partial fulfillment of the requirements for the degree, Master of Science in Computer, Information and Systems Engineering.

Sachin Shetty

---

Chairperson

Tamara Rogers

---

Committee Member

Liang Hong

---

Committee Member

Accepted for the Graduate School:

Alex Sekwat

---

Dean of the Graduate School

## **DEDICATION**

I dedicate this thesis to my family and to my advisor Dr. Sachin Shetty for their valuable support. Thank you for believing in me. It is also dedicated to the faculty and staff of Electrical and Computer Engineering. I am deeply grateful for the support. Thank you all.

-e.a.t-

PREVIEW

## **ACKNOWLEDGEMENTS**

First and foremost I want to thank God for giving me strength and courage to successfully complete this research work.

I express my deepest gratitude to my thesis advisor and mentor Dr. Sachin Shetty, for his valuable guidance and support. He has always given a very beneficial direction to my research work. He has always motivated, corrected and encouraged me as need be.

I would also like to present my sincere appreciation to my committee members Dr. Liang Hong and Dr. Tamara Rogers for their support. I am also grateful to Dr. Devgan, for his assistance and inputs in writing this thesis. My sincere gratitude goes to the entire Electrical and Computer Engineering Department, for offering their support and time whenever I approached them.

Moreover, I would like to express gratitude to my fellow students and researchers in the Computer, Information and Systems Engineering Lab for helping me in different aspects. They were big source of hope and courage in the times of my despair.

Last but not the least; I want to say special thanks to National Science Foundation for funding my research work.

-e.a.t.-

## **ABSTRACT**

PRIYA CHAURASIA. Dynamic analysis of Android Malware using DroidBox (under the direction of DR. SACHIN SHETTY)

With advent of technology, Smartphones have become an integral parts of our lives. Android is one of the most popular open source operating system used in Smartphones. It is also used by technology companies which require ready made, low cost and customizable operating system. Android's open nature has not only invited large community of developers but hackers as well. According to Forbes report, 97% of Mobile Malware in the year 2014 was on Android. Dynamically analyzing the Android Malware using DroidBox will not only provide an insight to this problem but also help in combating.

Mobile sandboxes are gaining popularity as they are able to overcome deception by executing malware in an isolated environment. DroidBox is an excellent mobile sandboxing tool to dynamically analyze the malware. We will utilize it as a base for porting to the recent version of Android. Porting would not only help in effective detection but also putting defenders way ahead in combating evasive mobile malware through an improved version of DroidBox.

Dynamic analysis of Android malware would collect the output generated by the DroidBox consisting of file system access, network activity, interacting with operating system, data sent, data received logs. We will implement porting of DroidBox which will not only help in combating malware but will be effective against all the existing sandboxes. In this thesis, defense strategies applied by Android Malware to thwart dynamic analysis is also investigated.

## TABLE OF CONTENTS

DEDICATION .....	IV
ACKNOWLEDGEMENTS .....	V
ABSTRACT .....	VI
LIST OF FIGURES .....	X
LIST OF TABLES .....	XII
LIST OF ACRONYMS AND ABBREVIATIONS .....	XIII
I INTRODUCTION AND NEED ANALYSIS .....	2
1.1. INTRODUCTION .....	2
II. REQUIREMENTS ANALYSIS AND CONCEPTUAL DESIGN .....	11
2.1. SYSTEM REQUIREMENT .....	11
2.1.1 Functional Requirements of the System.....	11
2.1.2 Operation Requirements of the System.....	12
2.1.3 User Requirements of the System .....	12
2.1.4 Maintenance Concept of the System .....	12
2.2. ASSUMPTIONS, LIMITATIONS AND CONSTRAINTS .....	13
2.3. ALTERNATIVE APPROACHES .....	13
2.3.1 Alternative 1 – Static Analysis .....	14



2.4.	PROPOSED APPROACH .....	15
2.5	SYSTEM OF INTEREST .....	17
2.6.	SYSTEM ENGINEERING MANAGEMENT PLAN (SEMP) .....	18
2.7.	WORK BREAKDOWN AND LIFECYCLE COST ESTIMATION .....	20
III.	PRELIMINARY SYSTEM DESIGN .....	22
3.1	DATA COLLECTION SUBSYSTEM .....	23
3.1.1.	Functional Requirements for Collection Subsystem .....	24
3.1.2.	Operational Requirements for Data Collection Subsystem.....	24
3.1.3.	Maintenance Requirements for Data Collection Subsystem.....	24
3.1.4.	Log Messages Collection Method.....	25
3.2	MODEL IMPLEMENTATION SUBSYSTEM .....	28
3.2.1	Functional Requirements for Model Implementation Subsystem...	29
3.2.2	Operational Requirements for Model Implementation Subsystem .	30
3.2.3.	Maintenance Requirements for Model Implementation Subsystem	30
IV.	DETAILED DESIGN, IMPLEMENTATION AND TESTING.....	31
4.1	DATA COLLECTION SUBSYSTEM .....	31
4.2	MODEL IMPLEMENTATION SUBSYSTEM .....	34
4.2.1	EXTRACTION AND CLASSIFICATION .....	37
V.	SYSTEM OPERATION MAINTENANCE AND PHASE OUT.....	41
5.1.	OPERATION AND MAINTENANCE PLAN .....	42
5.1.1	System Maintenance .....	42

5.1.2	Cost Allocation for System Operation and System Maintenance ...	43
5.2.	SYSTEM COST ANALYSIS .....	43
5.3.	CASH FLOW AND BREAK EVEN ANALYSIS.....	44
5.3.1	Cash Flow Matrix and Diagram .....	45
5.3.2	Break Even Analysis .....	45
VI.	CONCLUSION AND RECOMMENDATIONS .....	49
6.1.	CONCLUSION .....	49
6.2.	ANALYSIS AND LIMITATIONS.....	60
6.3.	FUTURE WORK AND RECOMMENDATIONS .....	60
	REFERENCES.....	51
	APPENDIX.....	53
	A. CODE TO COLLECT LOG FILES USING DROIDBOX AND LOGCAT .....	53
	B. CODE TO GENERATE OUTPUT USING JSON AND IPYTHON .....	53

## LIST OF FIGURES

<b>Figure No</b>	<b>Description .....</b>	<b>Page No</b>
1.1.	Mobile phone users Vs Desktop users .....	3
1.2.1.	Android System Architecture .....	5
1.3	Android build process overview .....	6
2.3	System of interest and its interaction with environment.....	18
2.4	System Engineering Management Plan .....	19
3.1	Subsystems of the Proposed System .....	23
3.2	Components of the Data Collection subsystem .....	24
3.3	Alternatives for collecting messages of DroidBox.....	25
3.4.	Components of the Model Implementation subsystem.....	28
4.1	Porting of DroidBox to latest version of Android .....	31
4.2.	Log Collection Process.....	32
4.3.	Execution of DroidBox to generate output log files.....	33
4.4.	Logcat file generation .....	33
4.5.	Model Implementation Subsystem.....	34
4.6	Parameters in JSON format generated by DroidBox .....	35
4.7	Format of file using IPython.....	35
4.8.	File System Activities.....	36
4.9	Network Activities .....	36

### Table of Figure (Continue)

4.10.	Send Data.....	37
4.11.	Received Data .....	37
4.12.	Crypto Activities .....	38
4.13.	SMS sent by malware.....	38
4.14.	Activities Chart.....	38
4.15.	Graph generation .....	39

## LIST OF TABLES

Table No.	Description.....	Page No.
2.1.	System alternatives evaluation measures .....	16
2.2.	Decision matrix for system alternatives .....	17
2.3.	Work breakdown and life cycle cost estimation .....	21
3.1	Alternatives and evaluation for Log collection.....	27
3.2.	Decision matrix for Log collection.....	27
5.1.	Cost estimation for system life cycle maintenance .....	44
5.2.	Cash flow matrix.....	45

## LIST OF ACRONYMS AND ABBREVIATIONS

IT	Information Technology
SMS	Secure Messaging Service
OS	Operating Systems
SQL	Structured Query Language
DVM	Dalvik Virtual Machine
UI	User Interface
APK	Application Package
JAR	Java Archive
SEMP	System Engineering Management Plan
IPC	Inter Process Communication
SDK	Software Development Kit
JSON	JavaScript Object Notation

# **CHAPTER I**

## **INTRODUCTION AND NEED ANALYSIS**

### **1.1 INTRODUCTION**

Internet has become an integral part of our lives. There are over 3.1 billion Internet users in the year 2015 which would be approximately 40% of the total world's population. Right from shopping to watching movies, we are completely dependent on it [1]. The past decade has seen an unprecedented growth in the field IT and Internet. E-commerce sites have also grown exponentially which in turn has invited Cyber criminals. Traditionally, IT supported business in terms of increasing efficiency and performance but now it being has being transformed into independent business. It is predicted that number of jobs in IT is going to rise by 22% till 2020. US had a market of over 348.9 billion in 2013 of E-commerce. Such a huge turnover of this industry has brought increase in malicious activities. Cyber criminals have used this tremendous opportunity to gain profit financially from this newly established cyber world. As per the Internet Security report by Symantec, one of the leading organizations in the area of malware security. There has been over 60,000 recorded vulnerabilities over the two decades from over 19,000 vendors representing 54,000 or more products which is alarming. Over the years as Internet has advanced and as more and more people have started using it, it is not confined to computers but there's shift in the paradigm. We are gradually moving away from computers to mobile devices [2].

Smartphones have started replacing computers and are gaining huge popularity. As the technology is advancing, the smartphones are not confined just for making calls or sending messages. The sky is the limit, we use it browsing web, social networking, online banking, online shopping, Starbucks coffee, storing credit card details, personal information and the list is endless [3]. You can be a naïve user and still order everything with the touch of a screen. It is estimated that by the end of year 2014-15, the number of smartphones users are going to surpass the number of desktop/computers users worldwide. The graph bel

shows a steady rise in the number of Smartphone users worldwide. The popularity of smartphones has evoked interest from Cyber criminals and made it a haven for malicious activities.

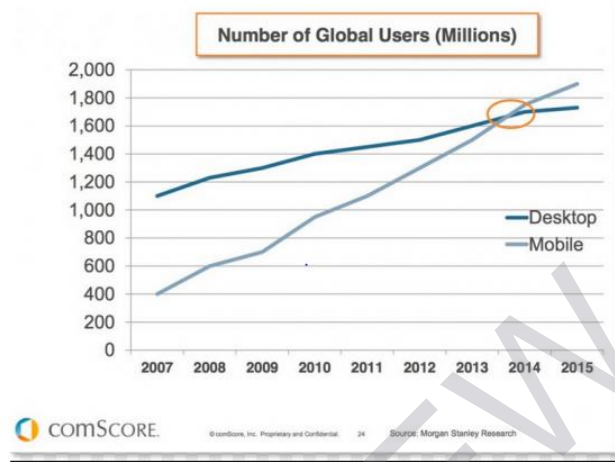


Figure 1.1 Mobile phone users Vs Desktop users [4]

There has been numerous research in the field of smartphone security. Researchers have been studying mobile phone security for years. Initially, the analysis was confined to proof of concept but the malware now has become a real threat. A recent study conducted to categorize the malicious malware activities in smartphones involved stealing user credentials, sending premium rate SMS messages, credential theft, SMS spam, search engine optimization and ransom [5]. As the smartphones are getting more advanced in the software and hardware, the malwares are also evolving. The first malware named Trojan-SMS.Android.fakeplayer was detected by Kaspersky in the year 2010 on Google's Android phone [6]. There's no looking back after that incident.

The traditional method use to detect and analyze malware samples are error prone and time consuming. Therefore, it is required to automate the analysis technique. Dynamic analysis of malware samples would allow us to detect suspicious, possibly malicious applications more effectively. Compared to static analysis, dynamic analysis results are not hindered by obfuscation techniques used by the application which in turn provides more accurate results [7].