



**Faculty of Information and Communication Technology**

**AN EFFICIENT SIEVE TECHNIQUE IN MOBILE MALWARE  
DETECTION**

**Mohd Zaki bin Mas'ud**

**Doctor of Philosophy**

**2018**

**AN EFFICIENT SIEVE TECHNIQUE IN MOBILE MALWARE DETECTION**

**MOHD ZAKI BIN MAS'UD**

**A thesis submitted  
in fulfilment of the requirements for the degree of Doctor of Philosophy**

**Faculty of Information and Communication Technology**

**UNIVERSITI TEKNIKAL MALAYSIA MELAKA**

**2018**

## DECLARATION

I declare that this thesis entitled, “An Efficient Sieve Technique in Mobile Malware Detection” is the result of my own research work 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 : Mohd Zaki Bin Mas'ud

Date : .....

**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 Doctor of Philosophy.

Signature : .....

Supervisor Name : Prof. Datuk Ts. Dr. Shahrin Bin Sahib

Date : .....

## **DEDICATION**

*This thesis is dedicated with  
Deepest love and affections to my beloved parents,  
**Hj. Mas'ud Taib and Hajjah Siti Juariah Hamdan***

*My son*

***Muhammad Hifzhan Irfan Mohd Zaki***

*Brother and sisters*

***Rozita, Roslina and Nazri***

*Their love, patience, guidance, wisdom and strength  
Have inspired me throughout these years in  
Universiti Teknikal Malaysia Melaka  
To be the best that I can be.*

## ABSTRACT

Proliferation of mobile devices in the market has radically changed the way people handle their daily life activities. Rapid growth of mobile device technology has enabled users to use mobile device for various purposes such as web browsing, ubiquitous services, social networking, MMS and many more. Nowadays, Google's Android Operating System has become the most popular choice of operating system for mobile devices since Android is an open source and easy to use. This scenario has also ignited possibility of malicious programs to exploit mobile devices and consequently expose any sensitive transaction made by the user. A malware ability to quickly evolve has made mobile malware detection a more complex. Antivirus and signature based IDS require a constant signature database update to keep up with the new malware, thus exhausting a mobile device's resources. Even though, an anomaly-based detection can overcome this matter, an anomaly detection still produces a high amount of false alarms. Therefore, this research aims to improve Mobile Malware Detection by improving the accuracy, True Positive and True Negative as well as minimizing the False Positive rate using an n-gram system call sequence approach and a sieve technique. This research analyses the behaviour and traces of mobile malware application activity dynamically as mobile malware is executed on a mobile platform. Analysis done on mobile malware activity shows behaviour and traces of benign and malicious mobile applications are able to be distinctively classified through invocation of system call to a kernel level system by a mobile application. However, an n-gram system call sequence generated by this approach can contribute to a large amount of logged features that can consume a mobile device's memory and storage. Hence this research, introduces a sieve technique in Mobile Malware Detection process in order to search for an optimum set of n-gram system call. In order to evaluate the performance of the proposed approach Accuracy, True Positive Rate, True Negative Rate, False Positive Rate and Receiver Operating Characteristic curve are measured with dataset of mobile malware from Malware Gnome Project and benign mobile application from Google Play Store. The experiment finding indicates the 3-gram system call sequence is capable of improving Mobile Malware Detection performance in terms of accuracy as well as minimizing the false alert. Whereas the sieve technique is able to reduce number of n-gram system call features and providing an optimize 3-gram system call sequence features. The outcome indicate that a Mobile Malware Detection using 3-gram system call sequence as features and sieve technique is able to be used in improving a Mobile Malware Detection in classifying the benign and malicious mobile applications. The evaluation and validation shows that a Mobile Malware Detection using 3-gram system call sequence with sieve technique improve the classification performance. As a conclusion the 3-gram system call sequence Mobile Malware Detection with sieve technique is capable of classifying the benign and malicious mobile application more accurately and at the same time minimizing the false alarm.

## ABSTRAK

Perkembangan peranti mudah alih di pasaran telah mengubah cara kita mengendalikan aktiviti kehidupan seharian. Pertumbuhan pesat teknologi mudah alih telah membolehkan pengguna menggunakannya untuk pelbagai perkara seperti melayari web, perkhidmatan merata, rangkaian sosial, khidmat pesanan multimedia dan banyak lagi. Kini, Sistem Pengendalian Android Google telah menjadi sistem operasi pilihan utama untuk peranti mudah alih disebabkan ia adalah dari sumber terbuka dan mudah digunakan. Senario ini juga memunculkan kemungkinan perisian hasad yang boleh mengeksploitasi peranti mudah alih dan seterusnya mendedahkan sebarang transaksi sensitif pengguna. Keupayaan perisian hasad untuk berkembang pantas telah menjadikan pengesanan perisian hasad mudah alih rumit. Sistem Pengesanan Pencerobohan berasaskan antivirus dan kaedah tandatangan memerlukan kemas kini pangkalan data tandatangan secara tetap bagi setiap penemuan perisian hasad baru. Ini menyebabkan sumber peranti mudah alih cepat penuh. Walaupun pengesanan berasaskan anomali dapat mengatasi isu ini; ia masih menghasilkan jumlah penggeraan palsu yang tinggi. Oleh itu, penyelidikan ini bercadang menambahbaik Pengesanan Perisian Hasad Mudah Alih dengan meningkatkan ketepatan, Positif Benar dan Benar Negatif serta meminimumkan kadar Positif Palsu dengan menggunakan pendekatan urutan Sistem Panggilan *n*-gram dan teknik penyaringan. Penyelidikan ini menganalisis tingkah laku dan jejak aktiviti aplikasi Perisian Hasad Mudah Alih secara dinamik. Hasil analisis menunjukkan tingkah laku dan jejak aktiviti aplikasi mudah alih yang benigna dan hasad dapat diklasifikasikan melalui sistem panggilan yang dipanggil oleh aplikasi mudah alih dari sistem kernel. Walau bagaimanapun, urutan Sistem Panggilan *n*-gram yang dihasilkan menyumbang kepada pengumpulan log yang besar dan menyebabkan penggunaan sumber peranti memori dan storan yang tinggi. Oleh itu, teknik penyaringan diperkenalkan dalam Pengesanan Perisian Hasad Mudah Alih untuk mencari set ciri Sistem Panggilan *n*-gram yang optimum. Untuk menilai prestasi kaedah pendekatan yang dicadangkan, pengukuran penilaian Ketepatan, Kadar Positif Benar, Kadar Negatif Benar, Kadar Positif Palsu dan lengkung Ciri Pengendali Penerima digunakan diatas set data aplikasi perisian hasad mudah alih daripada Projek Gnome Malware dan aplikasi mudah alih yang bersih dari Google Play Store. Penemuan awal menunjukkan urutan sistem panggilan 3-gram mampu meningkatkan prestasi pengesanan Perisian Hasad Mudah Alih dari segi Ketepatan, serta meminimumkan Kadar Positif Palsu. Manakala teknik penyaringan dapat mengurangkan jumlah ciri yang perlu dilog seterusnya menyediakan urutan Sistem Panggilan 3-gram yang optimum. Hasil penemuan menunjukkan urutan sistem panggilan 3-gram Pengesanan Perisian Hasad Mudah Alih dengan teknik penyaringan dapat mempertingkatkan Pengesanan Perisian Mudah Alih dalam dalam mengelaskan aplikasi mudah alih yang benigna dan hasad. Ujian Penilaian dan pengesanan menunjukkan yang urutan sistem panggilan 3-gram Pengesanan Perisian Hasad Mudah Alih dengan teknik penyaringan dapat mempertingkatkan prestasi pengelasan. Sebagai kesimpulan urutan sistem panggilan 3-gram Pengesanan Perisian Hasad Mudah Alih dengan teknik penyaringan mampu mengelaskan aplikasi benigna dan hasad dengan lebih tepat dan pada masa yang sama meminimumkan penggera palsu.

## ACKNOWLEDGEMENTS

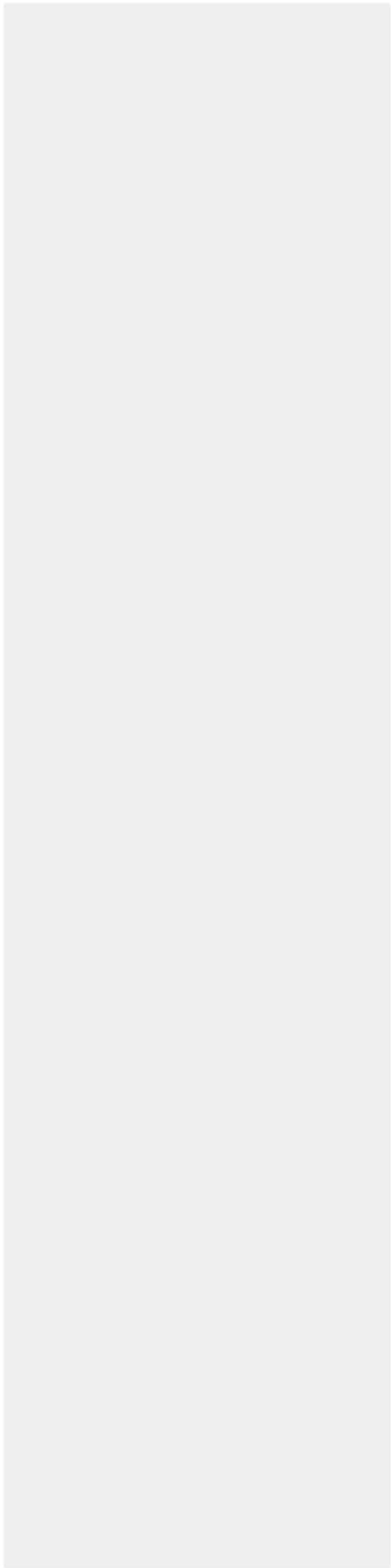
In the name of Allah, the Most Gracious and the Most Merciful

Alhamdulillah, all praises to Allah for His strength and blessing in completing this thesis. I would like to express my gratitude to my respectful supervisor, Professor Datuk Ts. Dr. Shahrin Sahib, whose expertise, understanding, and patience significantly enhanced my graduate experience.

I would like to express my appreciation to my co-supervisor Professor Madya Dr. Mohd Faizal Abdollah, my mentors Dr. Siti Rahayu Selamat and Dr. Robiah Yusof for their support and aid in making my PhD journey a success. My appreciation also go to Universiti Teknikal Malaysia Melaka (UTeM) and Ministry of Education Malaysia for sponsoring this research. My deepest thanks to all the people who had given their support and motivation to make this journey a success especially to all my colleagues in the department of Sistem Komputer dan Komunikasi (SKK) and generally in Fakulti Teknologi Maklumat dan Komunikasi (FTMK) for their constructive discussions and help with the analysis and in thesis writing during the course of this research.

Last but not least, from the bottom of my heart a highest gratitude to my family for their love and caring. Especially to my late father, Haji Mas'ud Hj. Taib and my mother, Hajjah Siti Juariah Hj. Hamdan for their encouragement and blessing, my eternal love to my son, Muhammad Hifzhan Irfan, who has been the pillar of strength in all my endeavours. Finally, to those who indirectly contributed to this research, your kindness has inspired me to embark on this journey





## TABLES OF CONTENTS

## PAGE

<b>DECLARATION</b>	
<b>APPROVAL</b>	
<b>DEDICATION</b>	
<b>ABSTRACT</b>	<b>i</b>
<b>ABSTRAK</b>	<b>ii</b>
<b>ACKNOWLEDGEMENTS</b>	<b>iii</b>
<b>TABLES OF CONTENTS</b>	<b>v</b>
<b>LIST OF TABLES</b>	<b>viii</b>
<b>LIST OF FIGURES</b>	<b>x</b>
<b>LIST OF APPENDICES</b>	<b>xiv</b>
<b>LIST OF ABBREVIATIONS</b>	<b>xv</b>
<b>LIST OF PUBLICATIONS</b>	<b>xvi</b>
<b>CHAPTER</b>	
<b>1. INTRODUCTION</b>	<b>1</b>
1.1 Introduction	1
1.2 Research Problem	3
1.3 Research Questions	6
1.4 Research Aim and Objectives	7
1.5 Research Scope	9
1.6 Research Contributions	10
1.7 Thesis Organization	11
1.8 Summary	13
<b>2. LITERATURE REVIEW</b>	<b>15</b>
2.1 Introduction	15
2.2 Chapter Objective	15
2.3 Chapter Outline	16
2.4 Overview of Mobile Malware Issues	17
2.4.1. Mobile Malware Evolution	18
2.4.2. Mobile Malware Threat and Impact	19
2.4.3. Android Architecture and Security Framework	21
2.5 Mobile Malware Detection (MMD)	23
2.5.1. Mobile Malware Detection	25
2.5.2. Mobile Malware Analysis Approach	27
2.5.3. Mobile Malware Audit Data Source and Detection Technique	30
2.5.4. Mobile Malware Detection Classification Analysis	36
2.5.5. Data Acquisition	38
2.5.6. Feature Selection Process	46
2.5.7. Evaluation Process	49
2.6 The Proposed Mobile Malware Detection	52
2.6.1. N-gram System Call Sequence	54
2.6.2. Wrapper Feature Selection Method	56
2.6.3. Support Vector Machines	58
2.7 Summary	60

**Comment [JD1]:** Formatting note:

Page 113 should be on right margin of table  
(subheading 4.5.4)

<b>3. RESEARCH METHODOLOGY</b>	<b>62</b>
3.1 Introduction	62
3.2 Chapter Objective	62
3.3 Chapter Outline	63
3.4 Research Design	63
3.5 Research Approach	64
3.6 Research Framework	65
3.7 Research Process	67
3.7.1. Clarify Research Question	68
3.7.2. Experimental Approach	69
3.7.3. Design Framework	74
3.8 Research Methodology and Research Objectives	80
3.9 Summary	81
<b>4. MOBILE MALWARE BEHAVIOUR THROUGH N-GRAM SYSTEM CALL SEQUENCE</b>	<b>83</b>
4.1 Introduction	83
4.2 Chapter Objective	83
4.3 Chapter Outline	84
4.4 Mobile Malware Behaviour Experimental Approach Overview	85
4.5 Mobile Malware Behaviour Analysis	86
4.5.1. DroidKungfu Behaviour Analysis	86
4.5.2. AnserverBot Behaviour Analysis	90
4.5.3. DroidDream Behaviour Analysis	93
4.5.4. GoldDream Behaviour Analysis	96
4.5.5. Discussion on Malicious Mobile Malware Behaviour	98
4.6 Mobile Malware Behaviour Traces through System Call Sequence	100
4.6.1. Accessing, Reading and Writing to a File	100
4.6.2. Connecting to the External Server	102
4.6.3. Capturing and Logging SMS Received	104
4.6.4. Discussion on Mobile Malware Behaviour and Sequence of System Call	105
4.7 The N-gram System Call Sequence Generator	107
4.7.1. Data Acquisition Process	108
4.7.2. Feature Vector Generation Process	110
4.8 Summary	113
<b>5. 3-GRAM SYSTEM CALL SEQUENCE FEATURES WITH EFFICIENT SIEVING TECHNIQUE</b>	<b>115</b>
5.1 Introduction	115
5.2 Chapter Objective	116
5.3 Chapter Outline	116
5.4 N-gram Evaluation Experiment and Analysis	117
5.5 The Feature Selection Process	121
5.5.1. Feature Selection Method Evaluation Experiment	122
5.5.2. Feature Selection Method Evaluation and Analysis Result	125
5.5.3. Sieve Technique with Wrapper Feature Selection and Best First Search Method	127
5.6 The Evaluation Process	129
5.7 The Proposed 3-gram System call and Sieve Technique in MMD	132

5.8	Summary	134
<b>6.</b>	<b>N-GRAM MOBILE MALWARE DETECTION WITH SIEVING TECHNIQUE EVALUATION AND VALIDATION</b>	<b>136</b>
6.1	Introduction	136
6.2	Chapter Objective	137
6.3	Chapter Outline	137
6.4	Dataset and Experimental Setup	138
6.5	System Call Evaluation	141
6.5.1.	Result and Findings for System Call Evaluation	141
6.6	The 3-gram System Call Evaluation Experiment.	145
6.6.1.	Result and Findings for 3-gram System Call Sequence Evaluation	145
6.7	Sieve Technique Evaluation	150
6.7.1.	Result and Findings for Sieve Technique Evaluation	150
6.8	Summary	154
<b>7.</b>	<b>CONCLUSION AND RECOMMENDATION</b>	<b>157</b>
7.1	Research Recapitulation	158
7.2	Research Contribution	161
7.2.1.	Mobile Malware Behaviour	163
7.2.2.	3-gram System Call Sequence	164
7.2.3.	An Efficient Sieve Technique	164
7.2.4.	Improve effectiveness of Mobile Malware Detection	165
7.3	Recommendation and Future Work	165
7.4	Conclusion	166
	<b>REFERENCES</b>	<b>168</b>
	<b>APPENDICES</b>	<b>188</b>

## LIST OF TABLES

<b>TABLE</b>	<b>TITLE</b>	<b>PAGE</b>
1.1	Research Problem	5
1.2	Summary of Research Questions	7
1.3	Summary of Research Problems (RP), Research Questions (RQ) and Research Objectives (RO)	9
1.4	Summary of Research Contributions	10
2.1	Mobile Malware Analysis Approach Use by Previous Researcher	28
2.2	Summary of Audit Data Source and Mobile Malware Detection Technique	32
2.3	The Advantages and Disadvantages of each MMD Element	36
2.4	Analysis on Feature Vector Generation	44
2.5	Mobile Malware Detection Framework Enhancement Process	52
3.1	Summary of Hardware and Software Used in Experimental Network Design	71
3.2	Previous Research on System Call Feature Selection	79
3.3	Research Methodology Mapping with RO1, RO2 and RO3	80
4.1	A Sample of the System Call Encoding Scheme	109
5.1	The Classifier Performance Evaluation Result	116
5.2	List of Feature Selection Method Evaluated and Analysed in this Research	121

5.3	The Feature Selection Method Performance Evaluation Result	122
6.1	Numbers of Sample for Each Dataset.	135
6.2	1-gram System Call Sequence Evaluation	137
6.3.1	Summary of Classification Accuracy Percentage	140
6.3.2	ANOVA with Single Factor Result	140
6.4	3- gram System Call Sequence Vs 1-gram System Call Sequence Evaluation	142
6.5	Significance T-Test On 3-gram and 1-gram System Call Sequence	145
6.6	3- gram System Call Sequence With Sieve Technique Evaluation	147
6.7	Significance T-Test on 3-gram	149

## LIST OF FIGURES

FIGURE	TITLE	PAGE
1.1	Distribution of Mobile Malware by Platform, 2012	2
1.2	Thesis Outline	11
2.1	The Structure of Chapter Two	16
2.2	The Known Mobile Malware Known Variant 2009-2016	18
2.3	Architecture of Android Platform	22
2.4	Mobile Malware Detection processes	26
2.5	Android Malware Detection Taxonomy	35
2.6	Filter Method	46
2.7	Wrapper Method	47
2.8	General Overview of Feature Selection	56
2.9	SVM Hyperplane Separating Benign and Malicious Mobile Application	57
3.1	Chapter Three Outline	62
3.2	Research Design	63
3.3	Research Phases	64
3.4	Research Process	66
3.5	Experimental Design	69
3.6	Experimental Network Designs for Testbed	70

3.7	Processes of Data Acquisition and Feature Vector Generation.	74
3.8	Experimental Evaluation Process	75
3.9	Process of Feature Selection Evaluation	77
4.1	Chapter Four Outline	82
4.2.1	The Captured System Call in DroidKungfu Execution Associated with an Attempt for Root Access	85
4.2.2	The Captured System Call in DroidKungfu Execution Associated with an Attempt to Communicate to External Server	86
4.2.3	The Captured Network Traffic Communication to External Server in DroidKungfu Execution	87
4.3.1	The Captured System Call from Anserverbot Execution Associated to Access Malicious Payload	88
4.3.2	The Captured System Call from Anserverbot Execution Associated to Connection Made to an External Server	89
4.3.3	The Captured Network Traffic Communication to External Server in Anserverbot Tcpdump Log	90
4.4.1	The Captured System Call from DroidDream Execution Associated to Accessing File	91
4.4.2	The Captured System Call from DroidDream Execution Associated to Connection Made to an External Server	92
4.4.3	The Captured Network Traffic Communication to an External Server in Anserverbot Tcpdump Log	93
4.5.1	The Captured System Call From GoldDream Execution Associated to Capturing the Incoming SMS	94
4.5.2	The Captured System Call From GoldDream Execution Associated to	95



	Connection Made to an External Server	
4.5.3	The Captured Network Traffic Communication to An External Server In GoldDream Tcpdump Log	95
4.6	Malicious Mobile Malware Behaviour	96
4.7	Malicious Mobile Malware Behaviour Execution Flow	97
4.8.1	System Call Sequence for change mode	99
4.8.2	System Call Sequence for Accessing and Writing File	99
4.8.3	System Call Sequence for Access and Rename File	100
4.8.4	System Call Sequence for Delete a File	100
4.9.1	Sequence of System Call Used to do a DNS Query	101
4.9.2	Sequence of System Call Used to Connect to the External Server	101
4.10	Sequence of System Call Used to Capturing and Logging SMS Received	103
4.11	Number of Features Generated Based on n	104
4.12	Mobile Malware Detection	105
4.13	Data Acquisition Process	106
4.14	Example of the System Call Log	107
4.15	Feature Vector Generation Process	108
4.16.1	Algorithm for System Call Encode.	109
4.16.2	Algorithm for Generating n-gram	110
5.1	Chapter Five Outline	114
5.2	n-gram Evaluation Experiment Procedure	115
5.3	Classifier Performance vs n-gram	116
5.4	ROC Curve of n-gram System Call Sequence	117
5.5	Feature Selection Method Evaluation and Analysis	120

	experiment procedure	
5.6	Sieve Technique with WR-BF Process Flow	124
5.7	Grid Search with Cross Validation Experimental Flow	126
5.8	Classification Accuracy Based on Different Penalty Parameter, $C$	127
5.9	MMD with 3-gram System Call Sequence and Sieve Technique	129
6.1	Chapter Six Outline	133
6.2	Evaluation and Validation Experiment Procedure	136
6.3.1	ROC curve for Dataset 1, D1	138
6.3.2	ROC curve for Dataset 2, D2	138
6.3.3	ROC curve for Dataset 3, D3	139
6.4.1	ROC curve 3-gram vs 1 gram for Dataset 1, D1	142
6.4.2	ROC curve 3-gram vs 1 gram for Dataset 2, D2	143
6.4.3	ROC curve 3-gram vs 1 gram for Dataset 3, D3	143
6.5.1	ROC curve Best 3-gram vs all 3-gram for Dataset 1, D1	147
6.5.2	ROC curve Best 3-gram vs all 3-gram for Dataset 2, D2	148
6.5.3	ROC curve Best 3-gram vs all 3-gram for Dataset 3, D3	148
7.1	The Objectives and Contributions Mapping	158

## LIST OF APPENDICES

<b>APPENDIX</b>	<b>TITLE</b>	<b>PAGI</b>
A	System Call Description	188
B	System call Representation	191

## LIST OF ABBREVIATIONS

**Comment [JD2]:** This list should be in alphabetical sequence, beginning with:  
AB Anomaly Base

Ending with:  
TPR True Positive Rate

API	-	Application Program Interface
RP	-	Research Problem
RQ	-	Research Question
RO	-	Research Objective
IDS	-	Intrusion Detection System
MMD	-	Mobile Malware Detection
TP	-	True Positive
TN	-	True Negative
FP	-	False Positive
FN	-	False Negative
SB	-	Signature Base
AB	-	Anomaly Base
SPB	-	Specification Base
SVM	-	Support Vector Machine
TPR	-	True Positive Rate
FPR	-	False Positive Rate
TNR	-	True Negative Rate
ROC	-	Receiver operating characteristic
IG	-	Information Gain
CHI	-	Chi-Square test,
CFS	-	Correlation-based feature
BF	-	Best First
GA	-	Genetic Algorithm
EA	-	Evaluation Algorithm
PSO	-	Particle Swam Optimization
CNC	-	Command and Control

## LIST OF PUBLICATIONS

Mohd Zaki Mas'ud, Shahrin Sahib, Mohd Faizal Abdollah, Siti Rahayu Selamat and Choo Yun Huoy, 2017. A Comparative Study on Feature Selection Method for n-gram Mobile Malware Detection. *International Journal of Network Security*, 19(5), pp. 727-733.

Comment [JD3]: See above

Mohd Zaki Mas'ud, Shahrin Sahib, Mohd Faizal Abdollah, Siti Rahayu Selamat and Robiah Yusof, 2016. An Evaluation of n-gram System Call Sequence in Mobile Malware Detection. *ARPN Journal of Engineering and Applied Sciences*, 11(5), pp. 3122-3126.

Comment [JD4]: See above

Mohd Zaki Mas'ud, Shahrin Sahib, Mohd Faizal Abdollah, Siti Rahayu Selamat and Robiah Yusof, 2014, May. Analysis of Features Selection and Machine Learning Classifier in Android Malware Detection. In *2014 IEEE International Conference on Information Science and Applications (ICISA)*, pp. 001-005.

Mohd Zaki Mas'ud, Shahrin Sahib, Mohd Faizal Abdollah, Siti Rahayu Selamat and Robiah Yusof, 2014. Android Malware Detection System Classification. *Research Journal of Information Technology*, 6(4), pp. 325-341.

Comment [JD5]: Article titles should have minimal capitalisation

Mohd Zaki Mas'ud, Shahrin Sahib, Mohd Faizal Abdollah, Siti Rahayu Selamat, Robiah Yusof & Rabiah Ahmad, 2013. Profiling Mobile Malware Behaviour through Hybrid Malware Analysis Approach. In *9th IEEE International Conference on Information Assurance and Security (IAS)*, 2013, pp. 78-84

## CHAPTER 1

### INTRODUCTION

#### 1.1 Introduction

The popularity of mobile devices over recent years has been continuously growing, with functionality similar to a personal computer. Mobile device users can do more than just making calls and handling Short Message Service (SMS). According to the International Telecommunication Union (ITU) (2016), at the end of 2016 there are almost 7.5 billion mobile users with more than 3.8 billion mobile-broadband subscriptions worldwide. The rise of mobile devices which have full functionality of a personal computers and support of latest communication technology has enabled users to always get connected to the Internet anywhere at any time. A mobile device can be used for various purposes such as web browsing, ubiquitous services, social networking, Multimedia Messaging Service (MMS) and many more. Robust Operating System (OS) Technology supporting mobile devices has also contributed to the rapid development of mobile applications on the mobile devices.

Currently, there are several mobile device OSs namely iOS from Apple, Blackberry, Symbian, Windows mobile and Android by Google. Among these OSs, Google's Android OS is widely consume in the mobile devices market shares;, Gartner Inc. stated that 84.1% smartphone sales during the first quarter of 2017 is on Android platform (Forni and Meulen, 2017). Android OS open source nature, credibility, performance and ease of customizing has made most mobile users choose mobile devices supported by Android OS from the others. Despite a rapid growth of Android-based mobile devices in the market, ahead of the other competitors, it also has

**Comment [JD6]:** Zaki Masud – it appears that the subheading numbering was updated while I was proofreading (this was not intentional on my part). Apologies for this; however, it should be corrected after all other changes are updated (after you accept or reject each change throughout the thesis) and Track Changes is then turned off. Then click on update Table of Contents – this should correct it.

**Comment [JD7]:** recent

**Comment [JD8]:** ...ahead of...

become an ideal place for malware writers. An increase in mobile applications in Android has also ignited the possibility of malicious programs which can exploit mobile devices. These malicious program are targeting the mobile devices because of the devices are used for online banking, online shopping or any sensitive transaction.

In early 2000, malicious software or malware has been only associated mainly with Desktop Computers but as the mobile technology evolved the malware has now proliferated the mobile space. Proliferation of malware on mobile technology exposed mobile user's sensitive information to malicious actions. Since 2010 new mobile malware is appearing at a regular interval. In 2012, Kaspersky Security Bulletin (Denis and Yuri, 2012) has reported that Android-based malware is growing at an alarming rate. Figure 1.1 shows that 98.96% of newly found mobile malware is targeting the Android-based platform. Mobile malware effect is lethal, mobile malware can steal credential information from the device, sniffed user activity and location, overbilled users by sending random SMS and MMS to contacts, launched denial of services attack from user devices and overloaded device resources such as memory, battery and storage (La Polla et al., 2013).

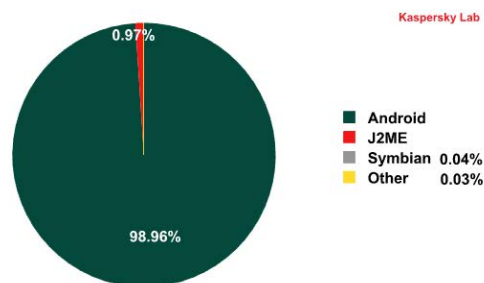


Figure 1.1: Distribution of Mobile Malware by Platform, 2012 (Denis and Yuri, 2012)

In 2010, Malware in general has cost consumer in United States USD 2.3 billion and caused 1.3 million personal computers to be replaced. In addition, in the same year, malware infection has cause USD 118 billion financial impact worldwide. As mobile devices technology

is now adapting all the personal computer capabilities, mobile devices are going to have similar effects. According to Juniper Network Mobile Threat Center (2012), the effect of mobile malware includes exploiting vulnerabilities in mobile payment gateway that can provide the attacker an immediate USD 10 million profit. Mobile malware has become an emerging threat in cyber security and some countermeasures need to be taken to overcome mobile malware infections. Therefore, developing and improving mobile devices security to the same level as computer security is important, especially in finding a mechanism to protect the system and data resources from any kind of intrusion (Sundaram, 1996).

## 1.2 Research Problem

Malicious software or called malware, is purposely written to exploit the vulnerabilities found in a computer system. Malware developers write malware code for different purposes which mostly are used for malicious intention (Robiah et al., 2009). The rapid evolution of malware signature and behaviour have made it difficult to stop. Previously, malware such as Virus, Trojan, Worm and Botnet are synonym to personal computers and rarely found in a mobile device. However, as the mobile devices are become increasingly complex and can support complex OS, mobile devices has become the malware's next target. The worldwide epidemic of malware infections has given malware authors a generous financial benefit through their activities in stealing credential information and gaining access to financial accounts. At present, in response to the emergence of mobile malware, security companies have released mobile antivirus applications as a defence mechanism.

Anti-malware applications, known as Antivirus for mobile, have a similar function as the one on the Desktop version; mobile version antivirus still detects malware based on the known malware signature and is useful for cleaning up the device after it has been infected. With a more advanced malware introduced, the signature is kept on changing from one variant to

**Comment [JD9]:** ADD TO REFERENCE LIST

**Comment [JD10]:** Do you mean IT ? If so, add IT to list of abbreviations. Otherwise, replace with specific noun.

**Comment [JD11]:** Revise word choice (associated with?)

**Comment [JD12]:** reward / benefit