



A SIMULTANEOUS SPAM AND PHISHING ATTACK
DETECTION FRAMEWORK FOR SHORT MESSAGE SERVICE
BASED ON TEXT MINING APPROACH

CIK FERESA BINTI MOHD FOOZY

DOCTOR OF PHILOSOPHY

2017



Faculty of Information and Communication Technology

**A SIMULTANEOUS SPAM AND PHISHING ATTACK DETECTION
FRAMEWORK FOR SHORT MESSAGE SERVICE BASED ON TEXT
MINING APPROACH**

Cik Feresa binti Mohd Foozy

Doctor of Philosophy

2017

**A SIMULTANEOUS SPAM AND PHISHING ATTACK DETECTION
FRAMEWORK FOR SHORT MESSAGE SERVICE BASED ON TEXT MINING
APPROACH**

CIK FERESA BINTI MOHD FOOZY

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

Faculty of Information and Communication Technology

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

2017

APPROVAL

I hereby declare that I have read this thesis and in my opinion this thesis is sufficient in term of scope and quality for the award of Doctor of Philosophy.

Signature :

Supervisor Name : **PROFFESOR DR. RABIAH AHMAD**
.....

Date :

DECLARATION

I declare that this thesis entitled “A Simultaneous Spam And Phishing Attack Detection Framework For Short Messaging Services Based On Text Mining Approach” 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 : CIK FERESA BINTI MOHD FOOZY
.....

Date :

DEDICATION

To my beloved mother; Noraity Nordin, father; Mohd Foozy Ghazali, husband;
Ahmad Shahir Abdul Rahman, daughters; Arissa Sofea and Arna Shafina and also family.

ABSTRACT

Short Messaging Service (SMS) is one type of many communication mediums that are used by scammers to send persuasive messages that will attract unwary recipients. In Malaysia, most sectors such as telecommunication, banking, government, healthcare, and private have taken the initiative to educate their clients about SMS scams. Unfortunately, many people still fall victim. Within the field of SMS detection, only the framework for a single attack detection for Spam has been studied. Phishing has never been studied. Existing detection frameworks are not suited to detect SMS Phishing because these attacks have their own specific behaviour and characteristic words. This gives rise to the need of producing a framework that is able to detect both attacks at the same time. This thesis addresses SMS Spam and Phishing attack detection framework development. 3 modules can be found in this framework, of which are Data Collection, Attack Profiling and Text Mining respectively. For Module 1, the data sets used in this research are from the UCI Machine Learning Repository, the Dublin Institute of Technology (DIT), British English SMS and Malay SMS. The Phishing Rule-Based algorithm is used to extract SMS Phishing. For Module 2, the SMS Attack Profiling algorithm is used in order to produce SMS Spam and Phishing words. The Text Mining module consists of several phases such as Tokenization, Lemmatization, Feature Selection and Classifier. These phases are done with the use of Rapidminer and the Weka data mining tool. Three (3) types of features are used in this framework, which are the Generic Features, Payload Features and Hybrid Features. All of these features are examined and the resulting performance metric used to compare the results is the rate of True Positive (TP) and Accuracy (A). There are four (4) set of results that were successfully obtained from this research. The first result shows that the extraction of SMS Phishing from the SMS Spam class contributes to four (4) enhanced datasets of the UCI Machine Learning Repository, the Dublin Institute of Technology (DIT), British English SMS and Malay SMS. The second results are the SMS Spam and Phishing attack profiling from the enhance UCI Machine Learning Repository, the Dublin Institute of Technology (DIT), British English SMS and Malay SMS. The third and fourth results are obtained from Feature Selection and Classifier phase where Eighty (80) experiments were done to examine the Generic Feature, Payload Features and Hybrid Features. There are five (5) Classification techniques used such as Naive Bayes, K-NN, Decision Tree, Random Tree and Decision Stump. The result of Hybrid Feature accuracy using Rapidminer and Naive Bayes technique is 77.47%, for K-NN: 78.56%, Decision Tree: 57.16%, Random Tree: 57.24% and Decision Stump: 57.16%. Meanwhile, by using Weka the Naive Bayes accuracy rate get 71.45%, K-NN: 81.64%, Decision Tree: 57.10%, Random Tree: 70.64% and Decision Stump: 60.19%. The experiments done using Rapidminer and Weka data mining tool because this is the first survey to detect SMS Spam and Phishing attack at the same time and the results are acceptable. Additionally, the proposed framework also can detect the attack simultaneously using text mining approaches.

ABSTRAK

Khidmat Pesanan Ringkas (SMS) adalah salah satu medium yang digunakan oleh penipu untuk menghantar mesej memujuk yang akan menarik penerima. Di Malaysia, sektor-sektor seperti telekomunikasi, perbankan, kerajaan, penjagaan kesihatan dan perniagaan telah mengambil inisiatif untuk mendidik pelanggan mereka mengenai penipuan SMS. Malangnya, ramai masih terperangkap. Dalam bidang pengesanan SMS, hanya rangka kerja pengesanan serangan Spam yang dikaji. Serangan Phishing masih belum. Rangka kerja pengesanan sedia ada tidak sesuai untuk mengesan SMS Phishing kerana serangan Phishing mempunyai tingkah laku dan ciri perkataan yang tersendiri. Ini menimbulkan keperluan untuk menghasilkan satu rangka kerja yang mampu untuk mengesan kedua-dua serangan pada masa yang sama. Tesis ini menangani SMS Spam dan Phishing serta pembangunan rangka kerja pengesanan. Terdapat 3 modul dalam rangka kerja ini, iaitu Pengumpulan Data, Profil Serangan dan Perlombongan Teks. Bagi Modul 1, set data yang digunakan dalam penyelidikan ini ialah UCI Machine Learning Repository, Dublin Institute of Technology (DIT), British English SMS dan SMS Melayu. Untuk Modul 2, algoritma Profil Serangan SMS digunakan untuk menjana perkataan SMS Spam dan Phishing. Modul Perlombongan Teks mempunyai beberapa fasa seperti Pemecahan ayat, Pengumpulan perkataan, Pemilihan Ciri dan Pengelasan. Fasa-fasa ini dilakukan oleh perisian perlombongan data Rapidminer dan Weka. Terdapat Tiga(3) jenis pemilihan ciri yang digunakan dalam rangka kerja ini, iaitu Ciri Generik, Muatan dan Hibrid. Kesemua ciri ini diteliti dan metrik prestasi yang digunakan untuk membandingkan keputusan adalah Kebenaran Positif dan Ketepatan. Terdapat empat(4) set keputusan yang telah diperolehi daripada kajian ini. Keputusan pertama ialah pengeluaran SMS Phishing dari kelas SMS Spam yang menghasilkan empat(4) set data penambahbaikan daripada UCI Machine Learning Repository, Dublin Institute of Technology (DIT), British English SMS dan SMS Bahasa Melayu. Keputusan kedua adalah profil serangan untuk SMS Spam dan Phishing dari set data penambahbaikan UCI Machine Learning Repository, Dublin Institute of Technology (DIT), British English SMS dan SMS Bahasa Melayu. Keputusan ketiga dan keempat ialah Pemilihan Ciri dan fasa Pengelas dari Lapan puluh (80) eksperimen yang dilakukan berdasarkan Ciri Generik, Muatan dan Hibrid. Terdapat lima(5) teknik Pengelasan iaitu Naive Bayes, K-NN, Decision Tree, Random Tree dan Decision Stump. Keputusan ketepatan bagi Ciri Hibrid menggunakan Rapidminer untuk Naive Bayes adalah 77.47%, K-NN: 78.56%, Decision Tree: 57.16%, Random Tree 57.24% dan Decision Stump: 57.16%. Ketepatan menggunakan Weka bagi Naive Bayes ialah 71.45%, K-NN: 81.64%, Decision Tree: 57.10%, Random Tree: 70.64% dan Decision Stump: 60.19%. Eksperimen ini dilakukan dengan alat perlombongan data Rapidminer dan Weka kerana pertama kali mengesan serangan SMS Spam dan SMS Phishing pada masa yang sama dan menghasilkan keputusan yang memuaskan. Selain itu, rangka kerja yang dicadangkan ini juga boleh mengesan serangan secara serentak menggunakan kaedah perlombongan teks.

ACKNOWLEDGEMENTS

First and foremost, I would like to take this opportunity to express my sincere acknowledgement to my supervisor Professor Dr. Rabiah Ahmad from the Faculty of Information and Communication Technology, Universiti Teknikal Malaysia Melaka (UTeM) for her essential supervision, support and encouragement towards the completion of this thesis.

I would also like to express my greatest gratitude to my co supervisor Associate Professor Dr. Mohd Faizal bin Abdollah from the Faculty of Information and Communication Technology, Universiti Teknikal Malaysia Melaka (UTeM) for his supervision, advice and suggestions in this research.

Thank you to Faculty of Computer Science and Information Technology for guidance and moral support. Thanks also to University Tun Hussein Onn Malaysia (UTHM) and Ministry of Higher Educations (MOHE) for the study sponsorship.

TABLE OF CONTENTS

	PAGE
DECLARATION	
APPROVAL	
DEDICATION	
ABSTRACT	i
ABSTRAK	ii
ACKNOWLEDGEMENTS	iii
TABLE OF CONTENTS	iv
LIST OF TABLES	vii
LIST OF FIGURES	xii
LIST OF APPENDICES	xv
LIST OF PUBLICATIONS	xvi
CHAPTER	
1. INTRODUCTION	1
1.1 Problem Background	1
1.2 Research Problem	3
1.2.1 RP 1: Unavailable Framework to Detect both SMS Spam and Phishing	4
1.2.2 RP 2: Insufficient SMS Phishing Corpus for Detection Framework	4
1.2.3 RP 3: Deficiency of SMS Phishing Profiling for Detection Framework:	5
1.3 Research Question	6
1.4 Research Objectives	7
1.5 Research Scope	8
1.6 General Methodology	8
1.7 Research Significance	10
1.8 Thesis Outline	11
1.9 Chapter Summary	13
2. LITERATURE REVIEW	14
2.1 Introduction	14
2.2 Introduction to Short Message Services (SMS)	15
2.3 Introduction to Spam	16
2.3.1 Spam Attack Characteristics	17
2.3.2 SMS Spam Attack	18
2.3.3 Effect of SMS Spam	19
2.3.4 The Existing Solution for SMS Spam Attack	20
2.4 Introduction to Phishing	29
2.4.1 Phishing Attack Characteristic	30
2.4.2 SMS Phishing Attack	31
2.4.3 Effect of SMS Phishing	33
2.4.4 The Existing Solution for SMS Phishing Attack	35
2.5 The Attack Detection Framework	36
2.5.1 The Single Attack Detection Framework	38
2.5.2 The Simultaneous Detection Framework	40

2.6	Existing Module in SMS Attack Detection Framework	41
2.6.1	SMS Datasets Collection	42
2.6.1.1	Extract SMS Phishing from Spam Class	45
2.6.2	Tokenization	46
2.6.3	Lemmatization	48
2.6.4	Feature Selection	49
2.6.4.1	Feature Selection Technique	49
2.6.5	Classification	51
2.6.5.1	Text Classification Technique	52
2.7	Chapter Summary	56
3.	METHODOLOGY	57
3.1	Introduction	57
3.2	Research Design	57
3.2.1	Phase 1: Data Collection and Generate Dataset	58
3.2.2	Phase 2: Framework Development and Experiment	69
a)	Module 1: Data Collection	71
b)	Module 2: SMS Attack Profiling	72
c)	Module 3: Text Mining	74
3.2.3	Phase 3: Testing and Validating	81
3.2.4	Performance Metric	81
3.3	Chapter Summary	85
4.	RESULT AND DISCUSSION	86
4.1	Introduction	86
4.2	Programming Setup	87
4.2.1	Eclipse Java Neon	87
4.2.2	RapidMiner 5 and Weka 5	87
4.3	Result Experiments	90
4.3.1	Data Collection Result	90
a)	Extracting SMS Phishing from Spam Class	91
4.3.2	Result of SMS Attack Profiling	92
a)	Summary of the Attack Profiling Result	93
4.3.3	Text Mining Result	96
a)	Generic, Payload, Hybrid Feature and Classification Techniques Experiment Result	96
b)	Summary Accuracy Rate of RapidMiner and Weka tool	110
c)	Testing the Generic Feature with SMS Malay	113
d)	Testing the Classification Model to Different SMS Dataset	118
4.4	Discussion	121
4.4.1	Data Collection	121
4.4.2	SMS Attack Profiling	122
4.4.3	Feature Group	122
4.4.4	Classification	123
4.5	Chapter Summary	123

5. CONCLUSION AND RECOMMENDATION	124
5.1 Introduction	124
5.2 Summary of the Research	124
5.3 Conclusion on the Research Objectives	126
5.3.1 RO1: To Develop a Simultaneous SMS Spam and Phishing Detection Framework	126
5.3.2 RO2: To Extract SMS Phishing Class from SMS Spam using a SMS Phishing Rule-Based Algorithm	127
5.3.3 RO3: To Develop SMS Attack Profiling and Identify Suitable Features for Enhanced Dataset	128
5.3.4 RO4: To Validate SMS Spam and Phishing Framework by Comparing the Performance in Terms of True Positive (TP) and Accuracy (A) Rate	128
5.4 Discussion on the Findings of the Research	129
5.4.1 Extracting SMS Phishing from the SMS Spam Class	129
5.4.2 Simultaneous Detection in a Framework Successfully Detects SMS Spam and Phishing at the Same Time.	130
5.4.3 SMS Spam and SMS Phishing Contains Different Words Profiling	130
5.4.4 Suitable Feature Set Increased the Detection Accuracy Rate	131
5.5 Limitation of Research	131
5.6 Future Research	132
REFERENCES	133
APPENDICES	159

LIST OF TABLES

TABLE	TITLE	PAGE
1.1	Summary of research objectives	7
2.1	Processes in SMS Spam studies from 2006 to 2014	26
2.2	Literature on SMS corpus	44
2.3	SMS Classification Technique	55
3.1	Result from Raosoft Model by Krejcie and Morgan method	62
4.1	The Existing Dataset for this research	89
4.2	Result for the variable used in the phishing rule-based algorithm	90
4.3	The enhanced SMS Dataset after implemented Phishing Rule-Based algorithm	91
4.4	The most popular SMS Spam and Phishing words from the enhanced UCI Machine Learning repository dataset	93
4.5	The most popular SMS Spam and Phishing from the enhance DIT Spam Dataset (Delany et al., 2012)	93

4.6	The most popular SMS Spam and Phishing Profiling from the enhanced British Spam Dataset	94
4.7(a)	Accuracy result for the existing SMS dataset using RapidMiner tool	98
4.7(b)	Accuracy result for the enhance SMS dataset using RapidMiner tool	99
4.7(c)	Accuracy result for the existing SMS dataset using WEKA tool	99
4.7(d)	Accuracy result for the enhance SMS dataset using WEKA tool	100
4.8(a)	Accuracy result for the existing SMS dataset using RapidMiner tool	100
4.8(b)	Accuracy result for the enhance SMS dataset using RapidMiner tool	101
4.8(c)	Accuracy result for the existing SMS dataset using WEKA tool	101
4.8(d)	Accuracy result for the enhance SMS dataset using WEKA tool	102
4.9(a)	Accuracy result for the existing SMS dataset using RapidMiner tool	103
4.9(b)	Accuracy result for the enhance SMS dataset using RapidMiner tool	103
4.9(c)	Accuracy result for the existing SMS dataset using WEKA tool	104

4.9(d)	Accuracy result for the enhance SMS dataset using WEKA tool	104
4.10(a)	Accuracy result for the existing SMS dataset using RapidMiner tool	105
4.10(b)	Accuracy result for the enhance SMS dataset using RapidMiner tool	105
4.10(c)	Accuracy result for the existing SMS dataset using WEKA tool	106
4.10(d)	Accuracy result for the enhance SMS dataset using WEKA tool	106
4.11(a)	Accuracy result for the existing SMS dataset using RapidMiner tool	107
4.11(b)	Accuracy result for the enhance SMS dataset using RapidMiner tool	108
4.11(c)	Accuracy result for the existing SMS dataset using WEKA tool	108
4.11(d)	Accuracy result for the enhance SMS dataset using WEKA tool	109
4.12(a)	Summary accuracy result for the existing SMS dataset using Classification technique.	109
4.12(b)	Summary accuracy result for the enhanced SMS dataset (Class Spam and Phishing)	110
4.13(a)	Summary accuracy result for the existing SMS dataset using Classification technique.	111

4.13(b)	Summary accuracy result for the enhanced SMS dataset (Class Spam and Phishing)	111
4.14(a)	Summary accuracy result for SMS Malay Ham and Spam class using Naive Bayes	112
4.14(b)	Summary accuracy result for SMS Malay Ham and Spam class using K-NN	113
4.14(c)	Summary accuracy result for SMS Malay Ham and Spam class using Decision Tree	113
4.14(d)	Summary accuracy result for SMS Malay Ham and Spam class using Random Tree	114
4.14(e)	Summary accuracy result for SMS Malay Ham and Spam class using Decision Stump	114
4.15(a)	Summary of the accuracy result for Malay SMS Spam and Phishing class using Naive Bayes	115
4.15(b)	Summary of the accuracy result for Malay SMS Spam and Phishing class using K-NN	116
4.15(c)	Summary of the accuracy result for Malay SMS Spam and Phishing class using Decision Tree	116
4.15(d)	E Summary of the accuracy result for Malay SMS Spam and Phishing class using Random Tree	117

4.15(e)	Summary of the accuracy result for Malay SMS Spam and Phishing class using Decision Stump	117
4.16(a)	Example of testing result on DIT Spam Dataset (Ham and Spam)	118
4.16(b)	Example of testing result on British Spam Dataset (Ham, Spam)	118
4.16(c)	Example of testing result on DIT Spam Dataset (Spam and Phishing)	118
4.16(d)	Example of testing result on British Spam Dataset (Spam and Phishing)	118
4.17	Comparison result with other studies	119

LIST OF FIGURES

FIGURE	TITLE	PAGE
2.1	SMS message structure(Mahmoud and Mahfouz, 2012)	15
2.2	Number of Email Spam Detections, Q1-Q3 2016(Darya Gudkova et al., 2016)	17
2.3	Example of SMS Spam in Malay sent from Washington DC phone numbers to numerous phones in Malaysia (Cloudmark, 2014)	19
2.4	Spam message filtering (Guzella and Caminhas, 2009)	29
2.5	Number of Phishing Site and email from October 2015 to March 2016	29
2.6	Example flow of an SMS Phishing attack when to steal private information	31
2.7	Example of SMS Phishing in the Malay language	32
2.8	Phishing category (SOPHOS, 2010)	33
2.9	Top Phishing hosted by country (Malaysia, 2014)	34
2.10	Social engineering attack statistics for the year 2012 to 2014(Raj Samani and McFarland, 2014)	39
2.11	Spam message filtering (Guzella and Caminhas, 2009).	42

2.12	Phishing is a subsection of Spam (Beck and Zhan, 2010)	46
2.13	Classification	52
3.1	Research methodology model	58
3.2	The flowchart for SMS data collection	61
3.3	Raosoft Software for Sampling Size calculation	63
3.4	Flowchart of extracting SMS Phishing from SMS Spam	64
3.5	The flowchart extracting SMS Phishing from SMS Spam class	65
3.6	Generic flowchart of Phishing rule-based algorithm	66
3.7	Flowchart of Phishing rule-based algorithm for JAVA programming source code	67
3.8	Phishing rule-based algorithm	68
3.9	Main Steps in filtering Email Spam(Guzella and Caminhas, 2009)	70
3.10	Generic detection framework for Simultaneous SMS Spam and Phishing	71
3.11	SMS attack Profiling algorithm	73
3.12	Example of tokenization from sentence to words	74
3.13	Example of SMS lemmatization by grouping words in the dataset	75
3.14	Flowchart task for Training and Testing model using text mining for existing dataset SMS Spam	78

3.15	Flowchart task for Training and Testing model using text mining for enhance dataset SMS Ham, Spam and Phishing	80
4.1	Proposed detection framework for Simultaneous SMS Spam and Phishing.	86
4.2	Summary of the experiments for the research detection framework	88

LIST OF APPENDICES

APPENDIX	TITLE	PAGE
A	SMS Generic Features	158

LIST OF PUBLICATIONS

M Foozy, C Feresa, R Ahmad, MF Abdollah. 2014. A Framework for SMS Spam and Phishing Detection in Malay Language: A Case Study. *International Review on Computers and Software* 9 (7), 1248-1254.

Cik Feresa Mohd Foozy, Rabiah Ahmad, Mohd Faizal Abdollah. 2014. A Practical Rule Based Technique by Splitting SMS Phishing from SMS Spam for Better Accuracy in Mobile Device. *International Review on Computers and Software (IRECOS)* 9 (1), 8.

CHAPTER 1

INTRODUCTION

1.1 Problem Background

There are several services on a mobile device that are used by spammers and phishers to launch attacks. These include services such as Email, Mobile Browser, Short Messaging Service (SMS), Voice Call and other mobile applications. Recently, social engineering attacks such as Spam and Phishing have affected both the security and privacy of mobile phone users. These attacks are attributed to the mistreatment of mobile phones (Balduzzi et al., 2016).

Among these services and applications, SMS is the most widely used all over the world since the cost of sending a message via SMS is considered cheaper than a phone call. Many SMS websites provide services that enables a person to send either a single message or multiple messages in bulk for free to telephone numbers worldwide. This service is not only used for personal means but is also used for business marketing. By using SMS to advertise a product, sellers are better enabled to improve their business profits despite the possibility that their recipients may feel uncomfortable when receiving unwanted advertisement from an unknown sender. This type of unwanted message is known as SMS Spam (Sulaiman and Jali, 2016).

On the other hand, SMS Phishing is SMS messages that provoke a response in its recipients. These type of attacks increase every year (Landesman, 2012) and many people end up losing money because them.

This shows that the original sender had malicious intent as they have manipulated the use of SMS to interrupt user security and privacy. Studies which resulted in SMS attack

detection frameworks for SMS Spam to be developed have already been conducted and fleshed out. However, to date, no detection framework for SMS Phishing exists even though this type of attack is ever increasing.

Existing SMS attack detection frameworks are only able to detect SMS with Spam characteristics and Spam words. As such, these frameworks will give a false alarm when it detects SMS that contain Phishing words.

Various studies on intrusion detection system frameworks have successfully managed to detect several attacks at the same time. This gave rise to a solution to detecting SMS Spam and Phishing attacks simultaneously by developing and tweaking such a framework.

Although SMS Spam and Phishing are similar in their spelling and grammatical errors, the differences lie in the type of words used and the overall textual behavior for each attack respectively. Regarding the differences when it comes to risk, the threat of Phishing attacks have a higher risk than Spam attacks (Xavier et al., 2014). In September 2014, SMS Phishing experienced an increase of over 58% according to a huge growing number of reports in the U.S.A (Landesman, 2014). This is because SMS Phishing is the most popular type of attack used in cyber space (Yeboah-Boateng and Amanor, 2014). This shows that both attacks are considered risky to SMS recipients. Thus, through the integration of SMS Spam and Phishing in a detection framework via the text mining approach, an efficient solution for users to simultaneously mitigate these attacks can be found.

Text mining approach has shown positive results in detecting attacks and is also capable of extracting common patterns from textual data (Aggarwal and Zhai, 2013) (Kumar and Ravi, 2016). There are a few processes involved in the Text Mining approach that will be implemented into this framework to detect SMS Spam and Phishing attacks simultaneously. In this chapter, the problem background, research problem, research question, research