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

© Universiti Teknikal Malaysia Melaka

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

© Universiti Teknikal Malaysia Melaka

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

© Universiti Teknikal Malaysia Melaka

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 : ………………….…………

© Universiti Teknikal Malaysia Melaka

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 : …………………………...………

© Universiti Teknikal Malaysia Melaka

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.

© Universiti Teknikal Malaysia Melaka

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.

© Universiti Teknikal Malaysia Melaka

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.*

© Universiti Teknikal Malaysia Melaka

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.

© Universiti Teknikal Malaysia Melaka

TABLE OF CONTENTS

PAGE

© Universiti Teknikal Malaysia Melaka

© Universiti Teknikal Malaysia Melaka

© Universiti Teknikal Malaysia Melaka

LIST OF TABLES

© Universiti Teknikal Malaysia Melaka

© Universiti Teknikal Malaysia Melaka

© Universiti Teknikal Malaysia Melaka

© Universiti Teknikal Malaysia Melaka

© Universiti Teknikal Malaysia Melaka

LIST OF FIGURES

© Universiti Teknikal Malaysia Melaka

© Universiti Teknikal Malaysia Melaka

© Universiti Teknikal Malaysia Melaka

LIST OF APPENDICES

© Universiti Teknikal Malaysia Melaka

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.

© Universiti Teknikal Malaysia Melaka

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 interupt user security and privacy. Studies which resulted in SMS attack

© Universiti Teknikal Malaysia Melaka

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

© Universiti Teknikal Malaysia Melaka