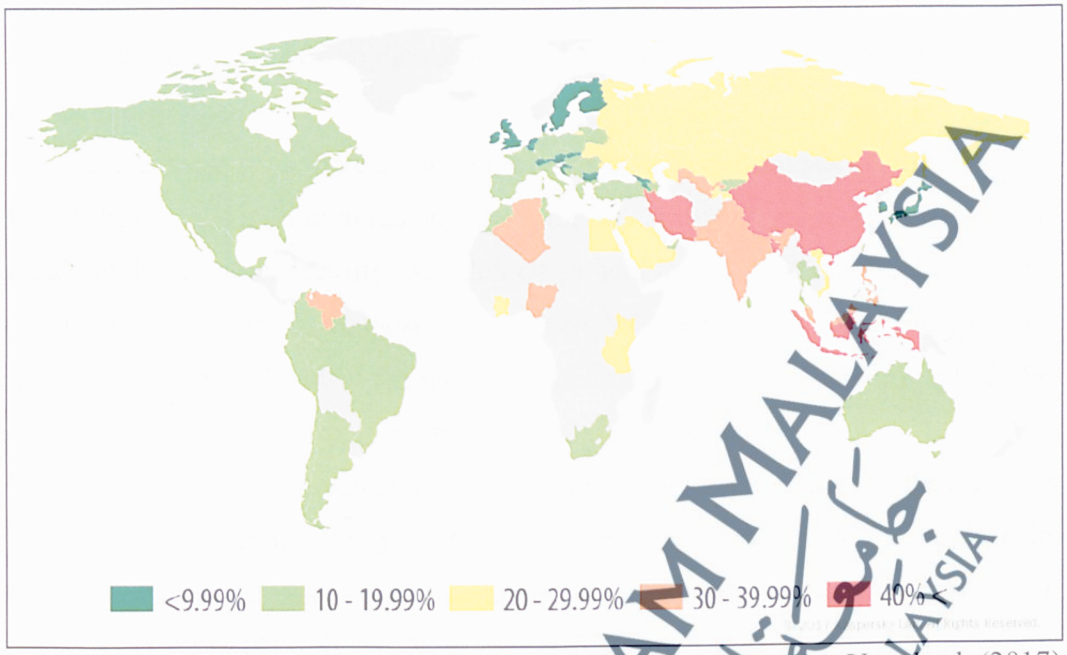
CHAPTER 1

INTRODUCTION

1.1 Motivation Of The Research

Spam has been intruding our daily life since decades ago. This threat has become more mature and even sophisticated in terms of its form in spam dissemination (Natris, 2014) and its unpleasant impact (Alotaibi, Furnell, & Clarke, 2017). It also corresponds to the innovation in mobile technology, booming together with the Internet advancement (Theoharidou, Mylonas, & Gritzalis, 2016). Many reports have been documented due to this threat's impact loss and its advent seems unending (Ladjana, 2018; Roshidi, 2018; Cloudmark, 2013), even getting almost robust and resistance from any mechanism of anti-spam solutions.

Currently, available research focuses on finding solutions from the aspect of spam detection and filtering which later is further implemented either at the telecommunication service providers (Mahmoud & Mahfouz, 2012a) or at the users' end, the mobile phone (Narayan & Saxena, 2013). The solutions have been developed and implemented in many forms but this particular nuisance is still kept arising and even evolved in avoidance to any detection mechanism. Nowadays, the impact of spam is not only it is an inconvenience and annoying but it also has becoming a more serious problem that gives rise to catastrophic effects, such as criminal activities. A survey done by Kaspersky that is reported in Unucheck (2017) discovered that in 2016, mobile threats have gone even wilder globally which includes Short Messaging Services or SMS spam as one of the tools of spreading the threats. For instance, SMS phishing, also known as SMiShing is an activity which enables criminals to get access to classified data. Through this technique, cyber criminals use mobile phone to manipulate innocent users into taking various actions, which can lead to them being defrauded (Sonowal & Kuppusamy, 2018; Goel & Jain, 2017).



Source: Umucheck (2017)

Figure 1.1: The Geography Of Mobile Threats By A Number Of Attacked Users In 2016

Receiving and responding to spam is one aspect that is directly related to users' behaviour. The way users' respond will determine how far the spam will affect them in many ways; such as money loss, time waste, and data loss. This has made human factors as one of the constant risks in information security cycle and regularly identified as the weakest link (Yeboah-Boateng & Amanor, 2014). Users need to be trained and kept updated with the latest method of scamming, especially when it comes to the matter of how they are supposed to handle and manage their private data technologically. Users need smart assistance and need to be told of the danger that they are facing. Unlike in the field of intrusion attacks, a complete cycle of management that include phases of detection, assessment for its potential risk and response against attack has been maturely developed (Anwar et al., 2017).

A survey conducted in 2017 with 500 respondents, certain important factors concerning SMS spams and mobile users awareness has been identified (Haritha, Kumar, & Krishnan, 2017). The survey study shows that there are users attracted mostly by promotional messages and they are aware that they are lack of knowledge about the

SMS spam and it's after effects. Other than that, most of the users are not aware of SMS spam characteristics, mobile applications for spam filtering and the generating point of SMS spam messages.

Hence, these days there are a number of works have been started that consider human behavior and culture to manage the information risk. For instance, a workshop attended by information security executives was held in 2011 to examine information security risks and challenges posed by human behavior. Throughout this session, surprisingly over 80% feel that the human-related risks are more troublesome than the technical challenges in securing information (Mong & Team, 2011).

Complementing the potential solution for the aforementioned issue, the maturity theory of human defence body system is reasonably attractive to be applied in designation and development of the computer security solution. Named as Artificial Immune Systems (AIS), this has been successfully applied to a range of real-world problems such as intrusion detection which includes anomaly and fault detection (Lokesh & Kumaraswamy, 2016; Kumar, 2015), optimization (L. J. Peng & Ruchuan, 2015), and classification (Al-Hassan & EL-Alfy, 2015). Inspired by one of the prominent theory in AIS, Danger Theory, this research work proposed to design, develop and implement an implicit decision maker suggesting to users the potential risk or impact that could cause them by the received spam. The Danger Theory is applied and enhanced in assessing the risk concentration in form of numerical value. The characteristics of antigen presenting and intelligence of dendritic cells in the human body that is compelling have inspired in finding a possible solution for this problem. This particular cell is not only able to detect danger in the human body, but also had biologically proven to be able to measure the severity concentration. By applying this key feature, the theory is highly expected to be able to distinguish spam and assess the potential impact level that could occur if a spam message is deemed trustworthy by the recipient. It makes sense to use this biological metaphor to deal with the problem of spam. A further study to identify reliable characteristics and how it can be applied in this risk assessment task is executed.

In a nutshell, this study is necessary to make an informed decision by identifying and deciphering the severity level of the text spam message to provide an appropriate

approach or a wise response that the user needs to take against the received spam message. Explicit information about the possible risk in a text message would assist users in taking wise action against spam such as delete, escalate the message to an authoritative body or just simply ignore it. The right action might keep the users away from the risk that could be highly severe to them because every choice of action has its consequence.

1.2 Problem Statement

There has been a lot of research executed to tackle the problem of spam in the email system. But, unfortunately, this threat is now seriously flooding in another mediums besides email such as SMS, web page and social media. The focus of the research now becomes wider and extended to a recent format, in line with the technology update. In finding a research gap particular for this issue, a number of constraints have been identified. Firstly, the current available research is mostly only able to classify the messages, which are to differentiate messages: either they are spam or valid message (Trivedi & Dey, 2016; Al-Talib & Hassan, 2013; Parimala & Nallaswamy, 2012). Some research is also available in clustering or categorizing spam messages (Mosquera, Aouad, Grzonkowski, & Morss, 2014; Tziortzis, 2013). But, previous studies have not dealt with labelling these spam messages with the appropriate tag of risk level, in order to assist the message recipients in identifying the potential risk that comes with the spam message. On the other side of information security, the research on risk labelling or ranking the severity level of threats is already conducted majorly for intrusion attacks and malware outbreak. Knowing the risk concentration in an identified threat is critical in making the next move, as in how to respond against it to mitigate the potential impacts that might occur (Adewole, Anuar, & Kamsin, 2016).

In addition to that, there is a very small number of research found for the application of Danger Theory in classifying SMS spam messages, but quite a lot in email systems. At the time of this writing, there is very minimal number of research found that objectively assess the risk or hazardous level in detected text spam messages with an employment of Danger Theory (M.El-Alfy & Alhasan, 2016). With the

capability of Danger Theory in detecting danger and assess the dangerousness of antigen in human body has inspired this study to apply the theory in spam risk assessment (P. Zhang & Tan, 2014). With an assumption that SMS message is the antigen, the concept of Danger Theory is applied in measuring the potential risk that SMS message may have.

Making the spam impact worse, the issue of awareness lacking among users of the risk of spam that may occur is one of the most vulnerable and critical issues that lead to numerous cases of fraud, scam and financial loss. The temptation to respond to the received spam is very risky and has high potential to cause damages, especially without their knowledge (Haritha et al., 2017; Yeboah-Boateng & Amanor, 2014). These factors have been marked as an interesting problem for further and advance research. Many research has been executed with intemperate findings that human factors are the most vulnerable element in securing the technology deployment.

There are numerous study found for SMS spam filtering (Lota & Hossain, 2017; Choudhary & Jain, 2017; Abdulhamid et al., 2017; M.El-Alfy & Alhasan, 2016; Sulaiman & Jali, 2015; Mosquera et al., 2014) which differentiation of spam and ham (valid) messages. However, there are no study found so far in assessing the risk of SMS spam messages. A study done by Theoharidou et al. (2012) shared a risk that may occurred for many type of threats in a smart phone which this study is more focus for threat assessment. Other than that, author in Adewole, Anuar, & Kamsin (2016) suggested a framework to assess the risk of microblogging comments into six (6) different levels; extremely normal, normal, low risk, medium risk, high risk, extremely harmful. This framework that combined multinomial NB and kNN consist of spam identification and risk assessment. However, this study did not provide proposal for SMS messages risk measurement.

Therefore, a tool that is automated in identifying the risk level of spam messages is desperately needed in order to assist spam recipients in making a decision. By considering the risk level assessed by this tool, recipients may act rationally whether to ignore, delete or even escalate these spam to an authority body for further action.

This research is focusing on the designation, development and implementation of an implicit decision maker for users to be well aware of the hidden dangers that lie

within a spam message. An understanding of the characteristics and implying what is the spam content all about is important in helping research communities to find a way to avoid or at least to minimize the tremendous effect brought by spam messages.

1.3 Research Questions

There are some inquiries arise from the problem identification that lead to the following questions for this study. Establishment of these questionnaires has been applied in constructing the objectives of the research.

i. How text spam messages can be assessed for its risk value?
ii. What is the possible mechanism to apply Danger Theory in risk assessment of text spam messages?
iii. How can the proposed risk assessment model for text spam messages identified as a reliable solution?

1.4 Research Objectives

Specifically, the objectives of this research are derived from the aforementioned research questions. These objectives will be as follows, also to fulfil the needs identified in the problem statement:

i. to study and evaluate the Danger Theory of AIS for application in risk identification and assessment on text spam messages;
ii. to propose and develop a model that is related to the assessment of spam risk level using an integration of the Danger Theory, text mining and risk assessment methodology; and
iii. to evaluate the accuracy of the proposed model with the aim of more than 90% accuracy rate.

1.5 Scope Of Research Works

This research is fully intended to focus on the following activities to facilitate the accomplishment of the research objectives:

i. to study the principle of Danger Theory introduced in Artificial Immune System (AIS) for the design of the model;
ii. to identify and formulate the link of Danger Theory to be implemented as a medium in threat and risk detection in text spam messages;
iii. to incorporate text mining and risk assessment technique for determination of spam risk level;
iv. to deploy a reliable data set (in English language) for the purpose of simulation in the developed model;
v. to test the proposed model with the SMS text data format; and
vi. to design, develop and evaluate the performance of the proposed and employed model in assessing the risk level of text spam messages.

As aforementioned, the main objective of this research is to study the Danger Theory abstraction as an alternative to be applied as a risk assessment medium for spam text messages. In addition to that, an identified term weighting scheme is believed to be a reliable method for the derivation of the input signals, has also contributed to this work specifically and in the field of text mining generally. As a result, the implemented immune classifier of the Danger Theory variants has been meticulously abstracted and developed. The findings from series of experiments in Chapter 3 and 5 are able to strengthen the claim via results of evidence.

1.6 Framework Of Research Methodology

Besides implementing the algorithm of the Danger Theory in this work, an integration of text mining for the derivation of a weight value intended as input signals seems as an added value to the model of spam risk assessment. The concept of text mining that has been utilized is incorporated into the risk assessment process via the calculation of danger concentration.

Prior to having these outcomes produced and constructed to align with the research objectives, there are a few phases of tasks that need to be implemented. A plan on how to execute the research is critically required. The plan is commonly known as a methodological framework that consists of many phases of research. This essentially assists and guides the whole journey of research. This methodological framework is depicted in the Figure 1.2.

The applied methodological framework is coinciding with the process proposed in Atila & Jones (2014). With reference to Imenda (2014), methodological considerations refer to the research design and the process of addressing a given research problem including the approach to the literature review, the nature of the collected data, analysed and interpreted.

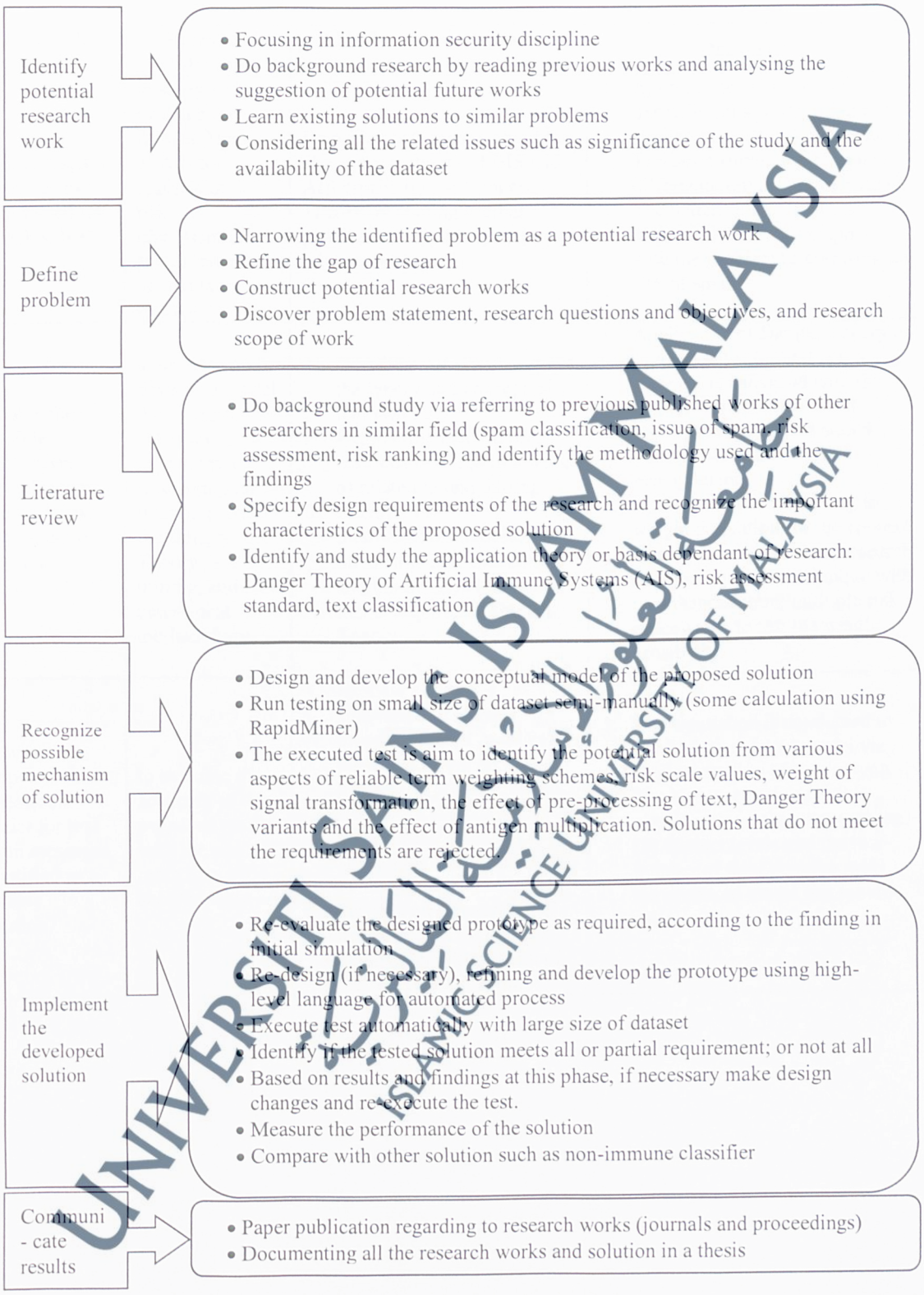


Figure 1.2: Methodological Framework

Table 1.1: Interrelationship Between Research Questions, Objectives, And Outcomes

Research Questions	Research Objectives	Methods and Activities	Deliverables
How text spam messages can be assessed for its risk value?	to study and evaluate the Danger Theory of AIS for application in risk identification and assessment on text spam messages	Do a literature review about the concept and theory of BIS and AIS especially the Danger Theory by reading journal papers, articles, books and related website.	1. Appropriate behaviour of dendritic cells is identified and applied in spam risk assessment. 2. The basic simulation for spam differentiation is comprehended. 3. The underlying theory from biological aspects is mapped with the problem in assessing the risk of spam.
What is the possible mechanism to apply Danger Theory in risk assessment of text spam messages?	to propose and develop a model that is related to the assessment of spam risk level using an integration of the Danger Theory, text mining, and risk assessment methodology	1. Do a literature review about the theory and concept of text mining. 2. Explore the method of previous research in a work of related to text mining. 3. Identify the interrelationship between the methods of text mining and input signals derivation that is required in Danger Theory.	1. Application of Danger Theory as risk evaluator for text spam message is enhanced with the application of suitable text mining technique and added with the standard of risk management. 2. Application of text mining in weight derivation for the context is studied from previous research and employed the technique with an enhanced item; multiple risk range value to test the weight sensitivity.
How can the proposed risk assessment model for text spam messages identified as a reliable solution?	to evaluate the accuracy of the proposed model with the aim of more than 90% accuracy rate	1. Execute a series of simulation to test the proposed model in terms of the effect of the pre-processing process, weight sensitivity and test with two (2) different types of algorithms (Danger Theory variants). 2. Perform an accurate measurement for the true positive and false alarm of the risk classification.	1. The simulation is conducted to verify the proposed model via implementation of a prototype. 2. The high accuracy of risk classification according to three (3) distinct levels and the low number of false positive make the proposed model is a reliable solution.

1.7 Thesis Outline

Chapter 2 elaborates insight of the literature that has been applied in this research. This chapter is divided into two (2) main sections that cover on the continuous issue of spam and how humans behave in reacting against it, as the initial content. The employed mechanism of Artificial Immune System or AIS that is integrated with the data mining methodology has been further studied as the subsequent section in this chapter. This section articulates how the AIS can be modelled in designing the conceptual framework of the spam risk assessment, which is basically constructed and guided by the understanding of the applied theoretical framework.

A further detail on risk assessment is described in Chapter 3. The concept of risk management has been identified to be integrated with the Danger Theory of AIS, which is inspired by the human cell in protecting and defending from any identified potential danger that possibly could cause damage. The consolidated mechanism then established in a form of conceptual model. The proposed model is designed and tested with a small-size of dataset to identify its reliability. A series of manual and automated experiments are simulated to demonstrate the proposed solution. In addition to that, the required parameters are recognized during this testing prior the prototype development.

In Chapter 4, the established model is further scrutinized computationally by designing and developing its algorithm, assisted with the pseudo-code and flow diagram. Parameters for the algorithm identified in the preliminary findings in Chapter 3 are applied in this prototype development.

A series of experiments are executed in Chapter 5 using the developed prototype. The testing is simulated separately in order to identify the best specification for the parameters required in the immune classifier. For comparative analysis, an experiment using non-immune classifiers are also conducted. The results are then comparatively studied to verify its findings.

Finally, Chapter 6 sums up this thesis with concluding remarks and identified potential future works in this area are discussed. The main contributions of this research to the knowledge are also elaborated.