CHAPTER 3

RISK CONCENTRATION FOR CONTEXT ASSESSMENT (RiCCA)

3.1 Introduction

The purpose of this chapter is to illustrate the design and initial evaluation of the proposed model named as **Ri**sk **C**oncentration for **C**ontext **A**ssessment or RiCCA. The design of RiCCA is evaluated and verified via a series of simulations conducted in Section 3.8 of this chapter. The design presents and illustrates an alignment between natural dendritic cells (DCs) and the artificial DCs used in the Dendritic Cell Algorithm (DCA) to demonstrate the principles behind the RiCCA's algorithm. The DCA is not a classification algorithm, which does not indicate whether the abnormality of an antigens, but the abnormal degree or concentration.

A series of simulation is conducted as a preliminary experiment to exhibit the designed model, RiCCA. This simulation is referred to as small-scale because the number of datasets utilized for testing purpose is just a few numbers of messages. The measurement of the identified risk is executed via manual calculation. Once the result is verified via this testing, then a prototype is developed for this task. A further experiment is executed using this prototype to evaluate the model in an automated way with the deployment of the larger size of the dataset.

This research that is aimed to apply Danger Theory in assessing the risk level of spam text messages identified that there are two (2) focal points to address the research interest; risk concentration and context assessment.

According to Greensmith, Aickelin, & Cayzer (2010), concentration refers to the number of molecules of signal per unit volume. In this case, risk concentration can be translated as the severity density (dangerousness) of a spam text message. Authors Mohsin Mohamad, Bakar, & Hamdan (2017) also suggested that the distinct advantages of DCA over other data mining approaches is the anomalies detection mechanism where it employs the dangerousness of an antigen. Hence, in this study, to differentiate the dangerousness, the malignant level is referring to the risk density based on three (3) levels of categorical data (high, medium, low level of risk) that comes together with a

numerical value (in between 0 to 1). The closer the calculated value to 1, the higher the potential risk is anticipated. This concept of risk measurement is aligned with the Mature Context Antigen Value (MCAV) calculation as suggested in the Danger Theory. In addition to the risk measurement concept, Greensmith & Aickelin (2008) suggested that the generated real value for anomaly score may assist in the polarization of normal and anomalous processes.

As articulated in Chapter 2, the Danger Theory is not only able to detect danger but in addition to that is its capability of measuring the maliciousness. This malignant state is commonly referring to the context of antigens. DCs that become the main role in the Danger Theory are known as professional antigen presenting cells and have the ability to process and collect signals and antigen. Its algorithm, DCA performs filtering of input signals, correlate between signals and antigen, and finally classifies the antigen types as normal or anomalous (Greensmith & Aickelin, 2009). The cumulative output signals are measured and the greater of semi-mature or mature output signal becomes the cell context for antigen assessment. This cell context is used to label all antigen collected by the DCs with the derived context value of 1 or 0 (Greensmith, Aickelin & Cayzer, 2010).

On the other hand, multiple signals that collected by DCs are combined and processed. This produced context information that represents the status of the environment. The state of environment is explicated either as semi-mature that implies a 'safe' context or mature that implies a 'dangerous' context (Greensmith, 2007; Aickelin & Greensmith, 2007). In this study, context assessment is referring to the phase where antigen (content of SMS messages) is being assessed to identify the malicious level or dangerousness.

The RiCCA conceptual model is depicted in Figure 3.1. This model is designed based on theoretical understanding of signal (translated as concentration) and antigen (translated as context) correlation from the Danger Theory. Signals are categorized from differentiation of its strength and every antigen has its own signal value. The state of surrounding will be determined by correlation of this signal and antigen information.

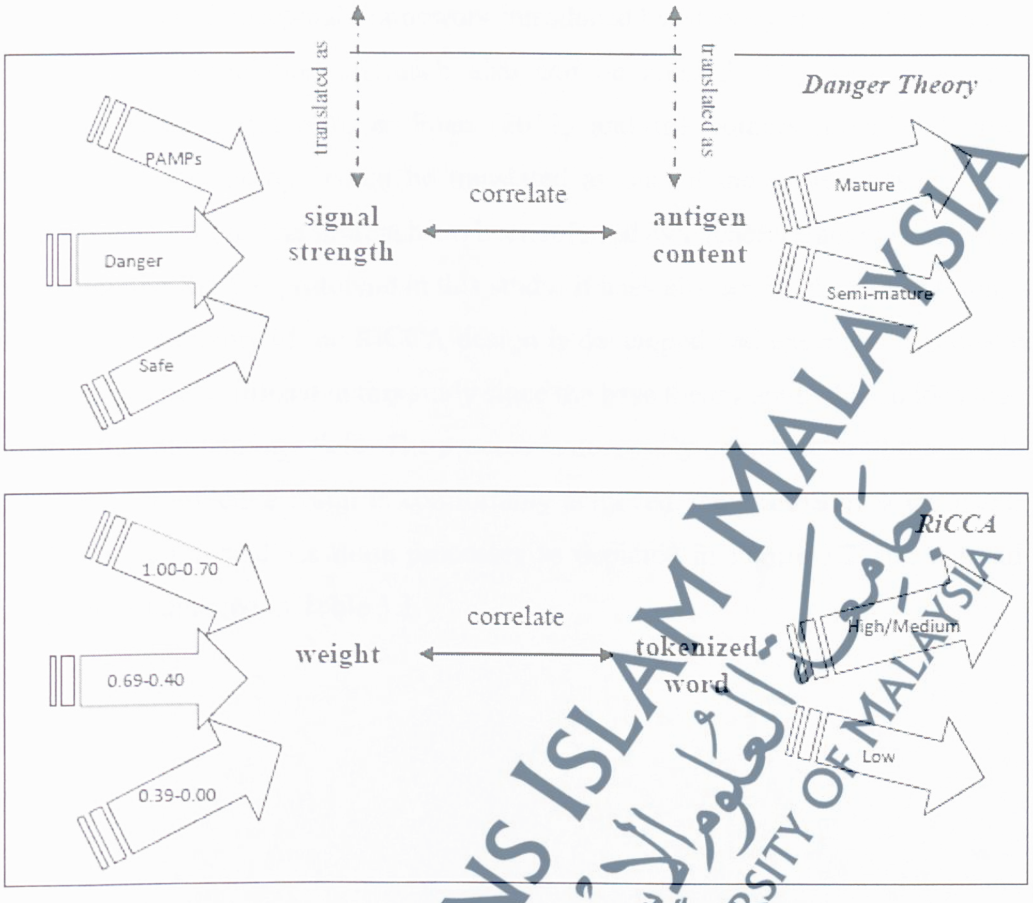


Figure 3.1: The Model of RiCCA

In this study, each research objective has its own particular task to ensure the achievability of the targeted cumulative goal is accomplished. Every phase of the task is consisting of a few sub-tasks that are required to be executed. All of these works are conducted phase by phase in a timely manner. Details of the milestone of the research work and its progress can be found in Appendix D.

3.2 Immunology-based System's Design And Development: The Life Cycle

Other than Conceptual Framework introduced by Stepney et al., (2005), a cycle of immunology simulation approach also can be applied. This is introduced by Figueredo, Siebers, Aickelin, & Foan (2012) and the authors articulated how a simulation in immunology could be translated as one of the approaches applied in problem investigation. The approach has been referred as a general framework to design and develop the RiCCA prototype in this study. It basically covers the initial testing to verify the functionality of the RiCCA design is developed and executed as projected. This guide has been utilized in this study since the base theory applied is an idea that is related to the immunology field. The process is iteratively simulated until the absolute design with an effective result is comfortably achieved. This research is particularly guided and coordinated via these processes as depicted in Figure 3.2 and the entire description is tabulated in Table 3.1.

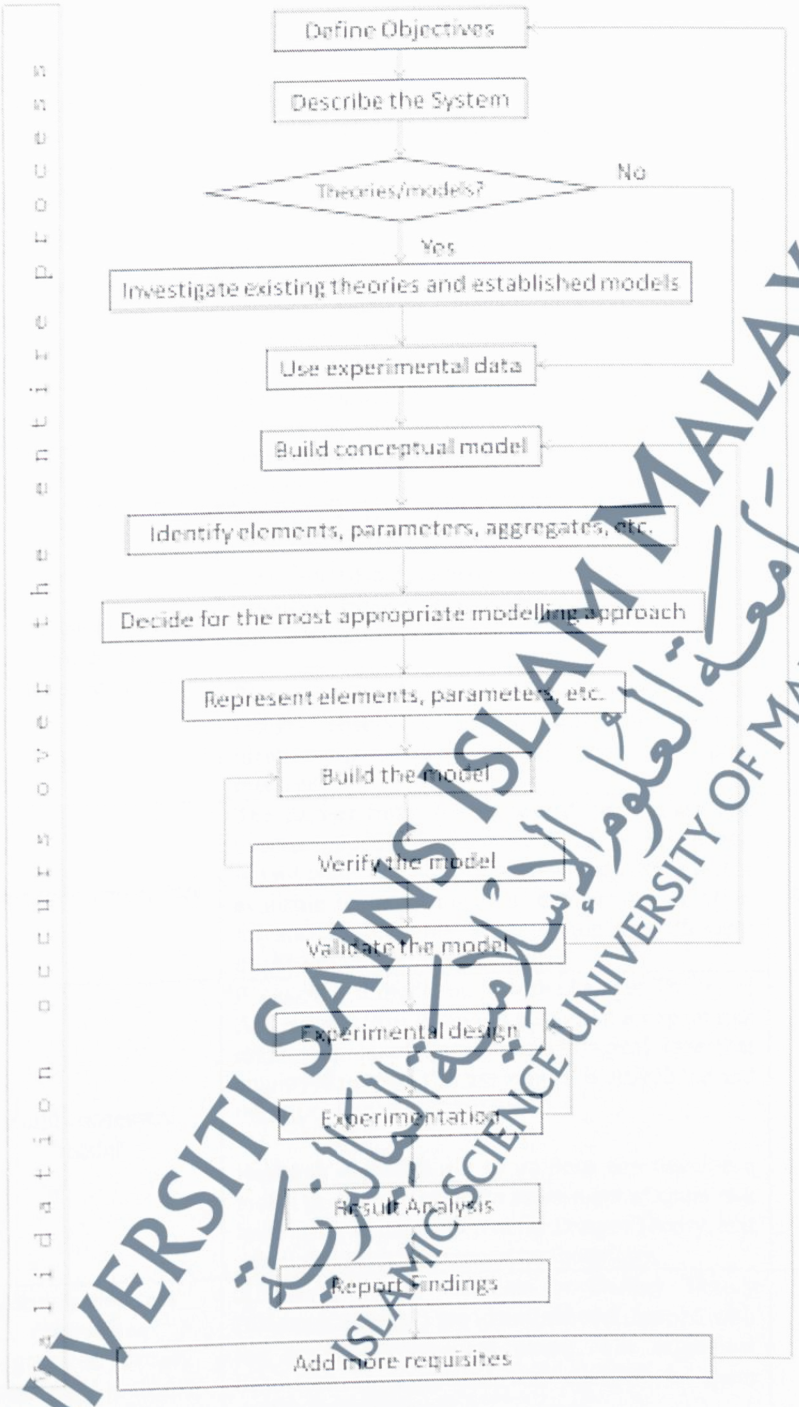


Figure 3.2: The Life Cycle Of Simulation Study Processes

Table 3.1: The Application Of Simulation Study Processes In RiCCA Designation And Development

No	Process in immunological simulation (as depicted in Figure 3.2)	Process for RiCCA implementation	References for further description
1	Define the objectives	Objectives of this research are defined and discussed based on identified scenario as the motivation of the research. According to formulated problem statement, the research objectives are developed with the intention to have a final outcome of risk assessment prototype.	Chapter 1
2	Describe the system	The description of the system is elaborated and defined by the scope of research works. The functionalities and limitations of the system are also identified.	Chapter 1
3	Investigate existing theories and established models	Meticulous reviews of the past literature are conducted. All possible related papers and journals are validated and studied to build a new model or an extension version as an improvement of what has already been established. Research objective #1: to study and evaluate the Danger Theory of AIS for application in risk identification and assessment on text spam messages.	Chapter 2 Chapter 3- Section 3.3, 3.4 and 3.5
4	Use experimental data	The dataset from the real-world observation and experimentation is identified. By utilizing the largest collection and publicly shared corpus that is available online, a potential comparison by other researcher is possible in finding the most efficient method to assess the spam risk.	Chapter 2- Section 2.9, Chapter 3- Section 3.6.1, Chapter 5- Section 5.2.1
5	Build conceptual model	A conceptual model on how the Danger Theory of AIS is capable of developing solution for spam risk assessment is analysed. The biological idea that employed in spam risk assessment is articulated and the RiCCA model is designed. Research objective #2: 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.	Chapter 3- Section 3.5
6	Identify elements, parameters, aggregates. Already established in theory and real-world data	All the potential elements in Danger Theory (theoretical model) are identified and mapped with the designed conceptual model. The biological behaviour is articulated and clarified in spam problem environment.	Chapter 3- Section 3.6, 3.7
7	Decide on the most appropriate simulation approach	All related elements for the proposed conceptual model to function properly are identified and tested. These additional elements include risk assessment process and text mining.	Chapter 3- Section 3.6
8	Represent elements or parameters using the appropriate simulation approach	Series of experiments are executed to ensure the most appropriate simulation approach is effectively applied for the proposed model. The results from this preliminary simulation determined the design precision.	Chapter 3- Section 3.7, 3.8 and 3.9

Table 3.1, continued

No	Process in immunological simulation (as depicted in Figure 3.2)	Process for RiCCA implementation	References for further description
9	Build the simulation model	The findings from series of experiments are used to build the prototype for the purpose of the automation process.	Chapter 3- Section 3.8, 3.9 Chapter 4
10	Verify the model	The computational model (*in silico*) is verified to ensure the algorithms are aligned and as delineated in the original DCA algorithm.	Chapter 4
11	Validate the model with existing theories and if available real-world data	The model is re-validated with the biological theory (Danger Theory) and the source of the dataset is confirmed. The authenticated model is established at this point.	Chapter 4
12	Experimental design	The design of the prototype is re-validate. The source of the dataset for the initial population and testing phase are confirmed and collected. Series of experiments are determined according to identified objectives. Research objective #3: to evaluate the accuracy of the proposed model with the aim of more than 90% accuracy rate.	Chapter 5
13	Experimentation	Once the proposed model is completely developed as the prototype, experiments are run with multiple series to verify various aims of experiments. Research objective #3: to evaluate the accuracy of the proposed model with the aim of more than 90% accuracy rate.	Chapter 5
14	Result Analysis	The results are analysed as an outcome of the model. The efficiency of the model is identified.	Chapter 5
15	Report Findings	The findings and conclusion are reported.	Chapter 5 and 6
16	Validate and add more requisites	The findings are validated and verified if it is well aligned with the biological inspiration idea and future enhancement is identified if any.	Chapter 5 and 6

3.3 Algorithms In Spam Management

According to Brownlee (2013), supervised learning deploy an input data that is called training data and has a pre-defined label or result, for example spam, not spam or a stock price at a time. A model is arranged through a training process where it is obliged to make predictions and is corrected when those predictions are wrong. The training process continues until the model achieves a desired level of accuracy on the training data. In managing spam, spam classification is applying this type of learning. In contrast to supervised learning, unsupervised learning has an input data that is not

labelled and does not have a known result. A model is prepared by deducing structures present in the input data and example problems are association rule learning and clustering. Clustering spam into category of its' content subject-matter is the sample of unsupervised learning.

A preliminary study conducted for this research includes various machine learning algorithms developed for numerous applications. A different algorithm is meant for the different purposes, for instance, an algorithm designed and developed for a classification task is not fit for a clustering task. This is due to the factor that every algorithm is unique in terms of its mechanism and functions in its own specific way. These algorithms usually end up developed as a tool for ease of employment specially to test on a large size of dataset automatically.

The initial test in this study is conducted using data mining tool, RapidMiner and Weka to classify and cluster the same set of SMS messages, a collection from UCI Machine Learning Repository. In this simulation, three (3) different algorithms (Naïve Bayesian (NB), Support Vector Machine (SVM) and k-Nearest Neighbour (kNN)) are applied for spam classification and three (3) others different algorithms (kMeans, Cobweb, Hierarchical) were applied for clustering these spam messages, as depicted in Figure 3.3. The phase for risk assessment is the focal point of this research, shown as dotted line. These algorithms were chosen since these are the well-performed algorithms for SMS spam filtering as discussed in Abdulhamid et al. (2017); Choudhary & Jain (2017); Lota & Hossain (2017); Arago et al. (2016); Sethi & Bhootna (2014); Warade et al. (2014); A. K. Uysal, Gunal, Ergin, & Gunal (2013) and Narayan & Saxena (2013).

This simulation is executed separately using data mining tools, RapidMiner and Weka. The final results showed that SVM is the best algorithm for spam classification, while K-Means is the most suitable algorithm to cluster spam messages. The results are referring to the highest accuracy rate of true classification and clustering task. These results are applied for both in RapidMiner and Weka data mining tools. In addition to that, this simulation also showed that the more the dataset is used in the training phase, the better the algorithm or classifier performed and gave a higher accuracy rate but as the size of datasets increases, the time taken to learn and classify the spam messages also increases in both training and testing phases.

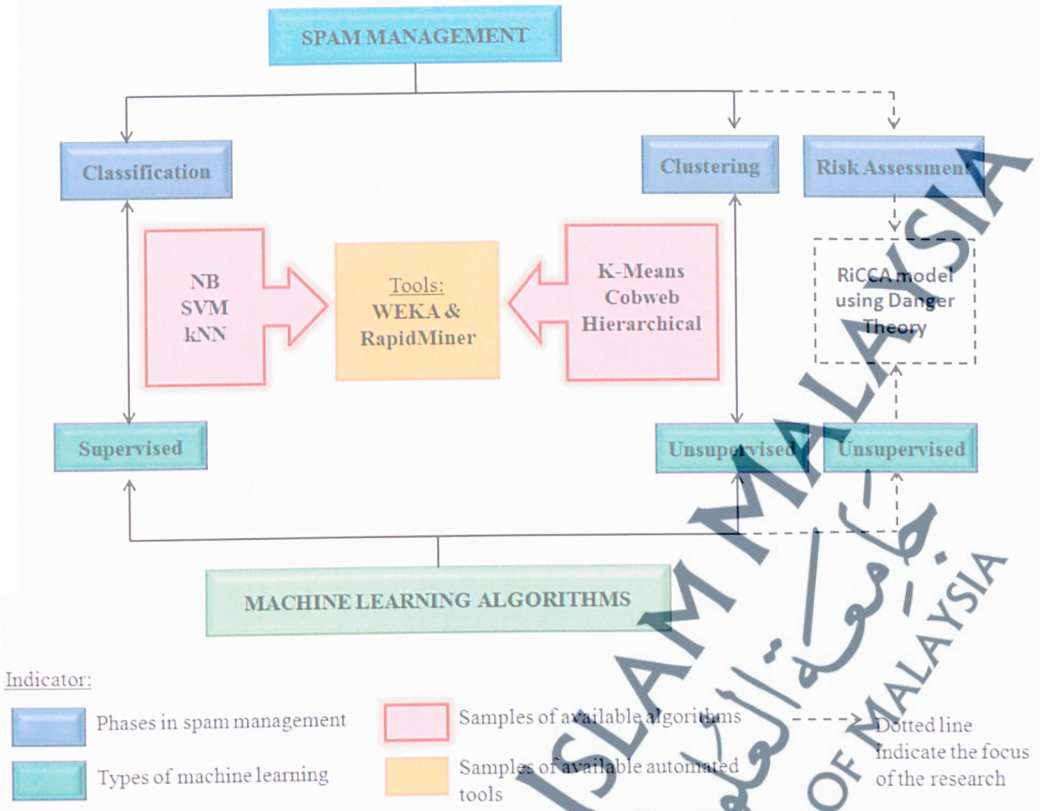


Figure 3.3: Integration Of Spam Management Processes With Machine Learning Algorithms And Experimental Tools

This preliminary study also identified that spam should be treated the same as other mobile or online threats that potentially comes with the effect of risk. An integration of risk management into the system of managing spam is elaborated in Section 3.4.

With reference to Figure 3.3, the phase of assessing risk is using unsupervised learning method. This is due to the element of DCA that there is no dynamic learning is attempted and hence its algorithm is not relying on a training data. Gu et al. (2011) and Gu et al. (2008) justified that the knowledge of normality and anomaly is acquired through basic statistical analysis. For this study, that analysis is retrieved from the application of term weighting scheme and risk scale. In this spam risk assessment, the sampling of initial population is later fed into the testing phase and executed as in the following Figure 3.4.

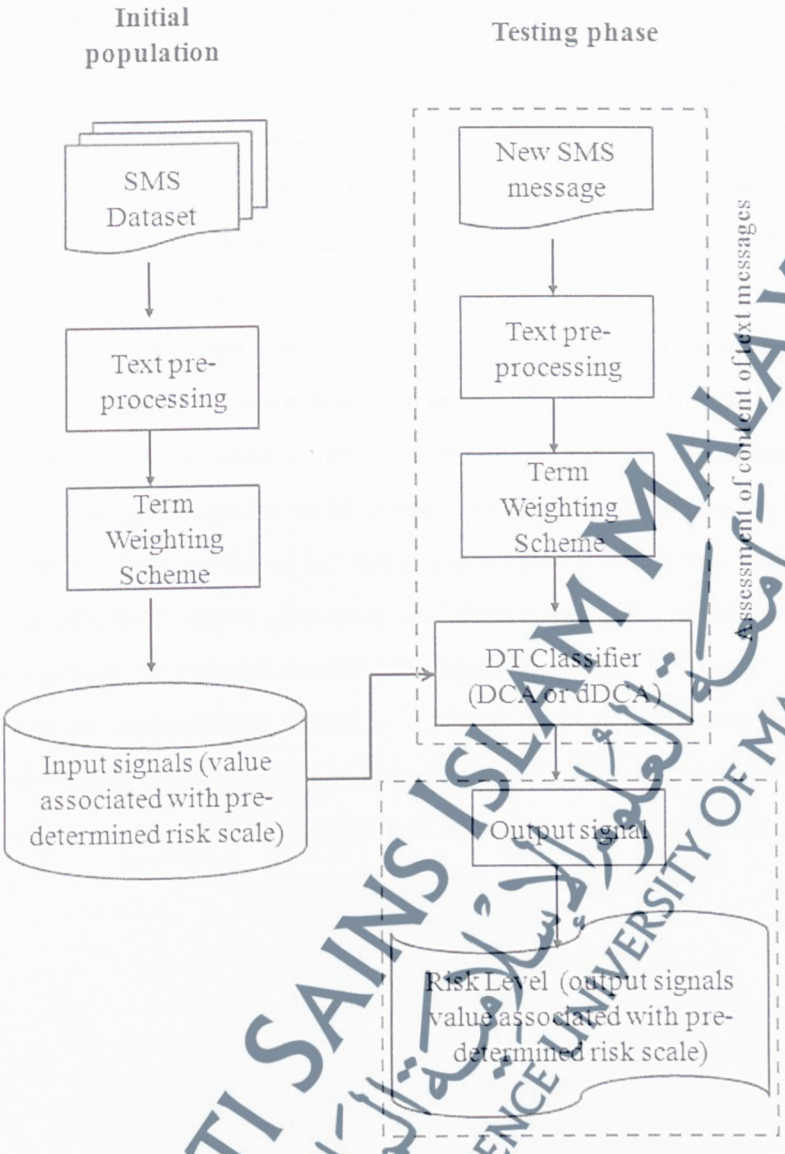


Figure 3.4: Unsupervised Learning Of RiCCA Model

In this study, a pre-determined risk scale is applied as the knowledge expert to distinguish the different levels of input and output signals. In addition to that, a questionnaire has been distributed to develop a list of expert judgment on how to map and label the calculated risk qualitatively (high, medium and low level of risk). This part is elaborated in detail in Section 3.6.3 and 5.3.1 of this thesis.

3.4 Treating Spam As Part Of Risk Management

Blank & Gallagher (2012) produced a guidance document about conducting risk assessments under National Institute of Standards and Technology (NIST). Risk assessments are a key, part of effective risk management to facilitate decision making that includes processes of identification of risk, assessing risk, responding to risk, and risk monitoring.

Theoharidou, Mylonas, & Gritzalis (2016) and Yeboah-Boateng & Amanor (2014) proclaimed that SMS spam has been observed as one of the mobile threats due to its malevolent impacts. Since spam is needed to be treated as other threat, then it is required to be managed as proposed in most of the risk management standards in order to prevail over it. There are three (3) main processes involved in managing spam which are spam classification, spam clustering, and determination level of spam's severity which also includes the options available to the response against spam.

The spam management model is depicted in Figure 3.5 and the integration between risk management and spam management is shown in Figure 3.6. The phase for risk assessment for spam is the focal point of this research, shown in dotted line.

Figure 3.5: The Model Of Spam Management

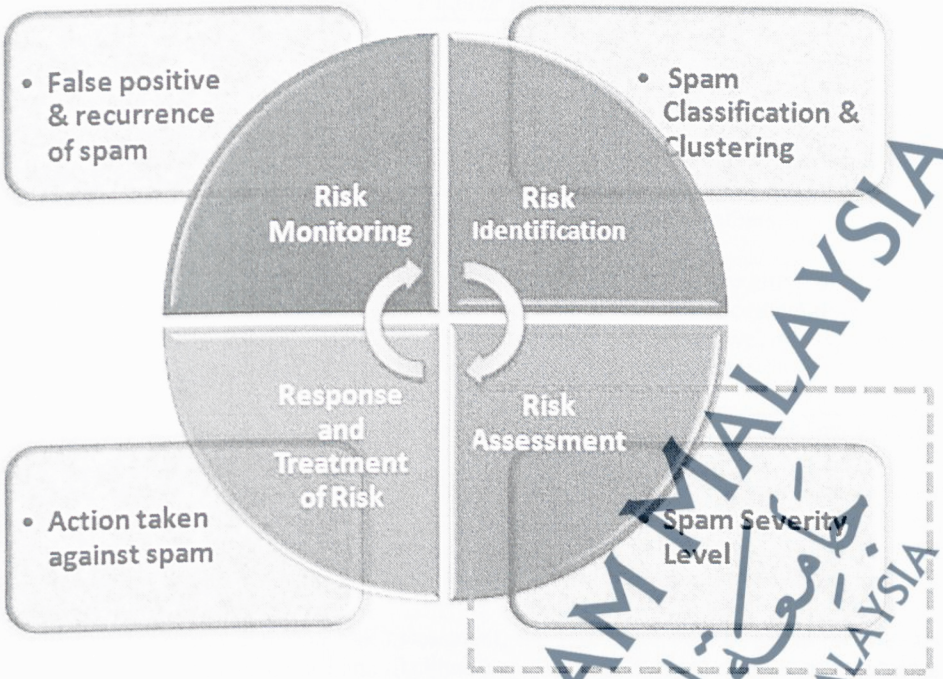


Figure 3.6: The Proposed Of An Integration Between Risk Management And Spam Management

The proposed spam management model consists of a few phases; classification, clustering, severity assessment and potential type of responses are justified in the following Table 3.2.

Table 3.2: Description Of Phases In Spam Management Model

Spam management model	Justification
Classification	In spam management model, classification of spam messages is the initial phase by identifying whether it is legitimate or spam. The detection of spam messages can be executed using various techniques such as machine learning algorithms (supervised and unsupervised), linguistics, graph pattern and many others.
Clustering	The detected spam is clustered according to the respective group or category (based on subject-matter). The spam categories such as competition, free prizes, advertisement, and financial assistance have a different level of risk according to the impact loss of identified threat.
Determination of severity level	The phase of determination of spam's severity level is divided into three (3) main risk category: i. High – spam event could be expected to have a severe or catastrophic adverse effect ii. Medium – spam event could be expected to have a serious adverse effect iii. Low – spam event could be expected to have a limited or negligible adverse effect As in a common risk assessment, the risk level of a threat is measured via the function of impact and likelihood of identified threat that exploiting vulnerabilities. Basically, risk assessment is the critical phase which the evaluation of risk will influence the decision on how to deal with it.
Response	The action taken to response is according to recipient's discretion either to ignore, delete or even escalate the spam message to authority body. The tagged of spam message with high and medium risk labelled implicitly has suggested to the recipient the possible action to response.

3.5 Mapping Biological Facets Of Danger Theory And Spam Management

Prior to designing the model inspired by biological characteristics, an understanding of how this dendritic cell (DC) behaves is vitally important. The consideration theoretically and conceptually is implied in the framework and translated into an algorithm or system through processes of abstraction and modelling. The conceptual framework is then verified via preliminary simulation before the *in silico* (computerized) type or the prototype is developed.

Table 3.3: The Conceptual Mapping Between Immune System Model, Danger Theory And SMS Spam

Biological property	Danger Theory property	Abstract property in risk assessment of detected SMS spam
Bacteria, pathological cells	Antigen	SMS spam or threat
Collecting signals around damaged cells	Dendritic cell	SMS spam folder/terms with weight value
PAMPs alerts – indicate the presence of microbial	PAMPs signals	High risk – signature of likely anomaly
Necrotic alerts – indicate the damage of tissue	Danger signals	Medium risk – indicate potential anomaly
Apoptotic alerts – indicate of healthy tissue	Safe signals	Low risk – indicate an absence of anomaly
Antigen Presenting Cells	Context	Content of SMS spam
Peptides	Peptides	Tokenized spam message
Inflammatory cytokines – indicate general tissue distress	Inflammation	Confirmation of SMS is a spam
Information collection by dendritic cells	Antigen and signal	Tokenized words and its associated weight value
Major Histocompatibility Complex (MHC)	Antigen storage	Spam database library / sampling during the initial population
Represents the dendritic cell of likely safe	Semi-mature signal	Low risk – indicate an absence of anomaly
Represents the dendritic cell of likely danger	Mature signal	High risk – signature of likely anomaly; or Medium risk – indicate potential anomaly

The biological characteristics applied for this proposed model, RiCCA is simplified in the following Figure 3.7. This figure illustrated the mapping of Danger Theory biological abstraction to portray the RiCCA, spam severity assessment proposed model.

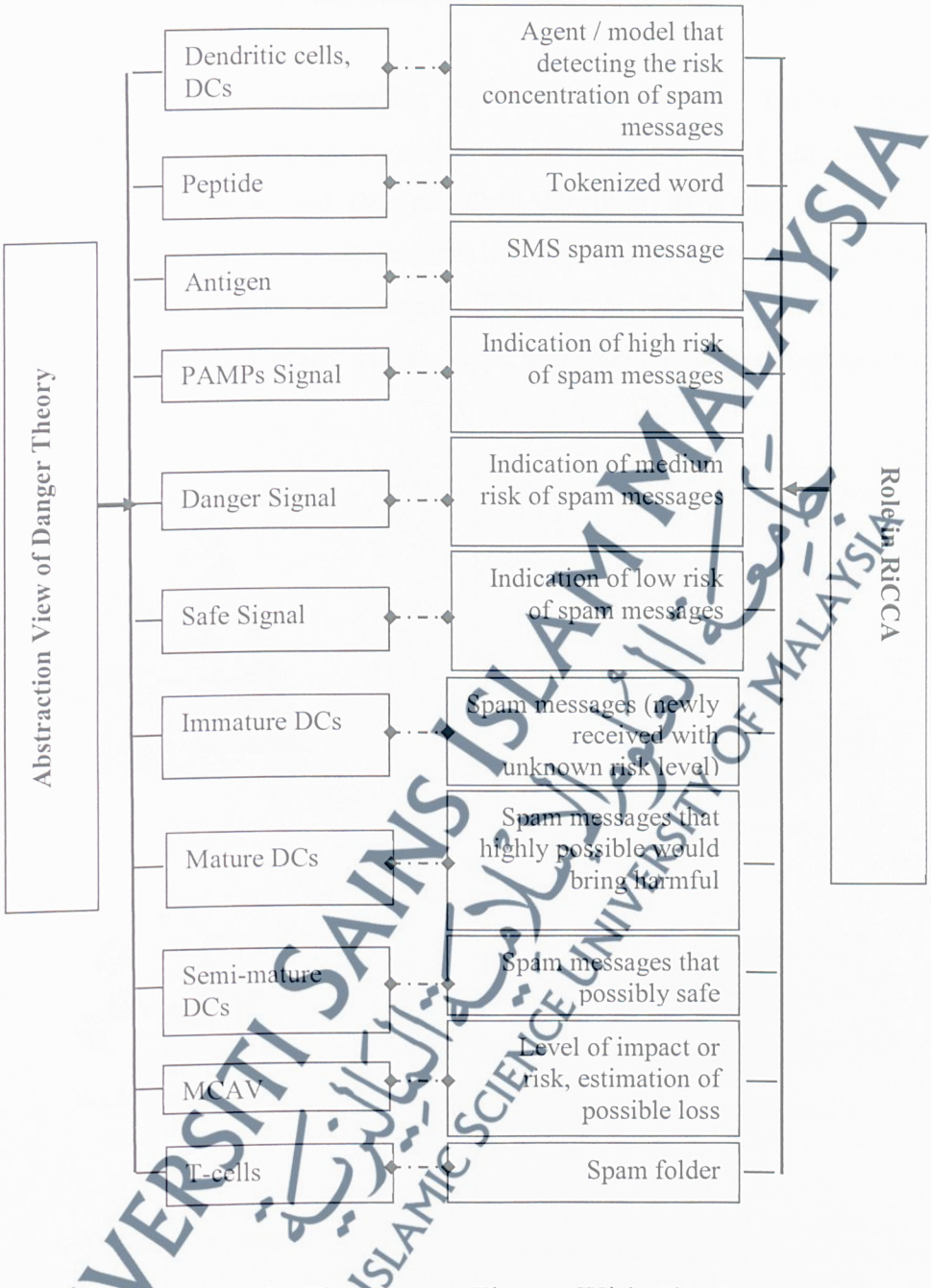


Figure 3.7: Mapping Of The Danger Theory With The Proposed Model Of RiCCA

3.5.1 Theoretical Versus Conceptual Consideration Of The Proposed Model, RiCCA

From the understanding of how the Danger Theory behaves theoretically, a conceptual consideration in the real case of this research is required to be constructed. Dendritic Cell Algorithm (DCA) that emerged from the Danger Theory detects differences in antigen based on the context, which is derived from a signal (Greensmith, 2007). A detailed look on this concept establishment is elaborated via prototype implementation described in Chapter 4.

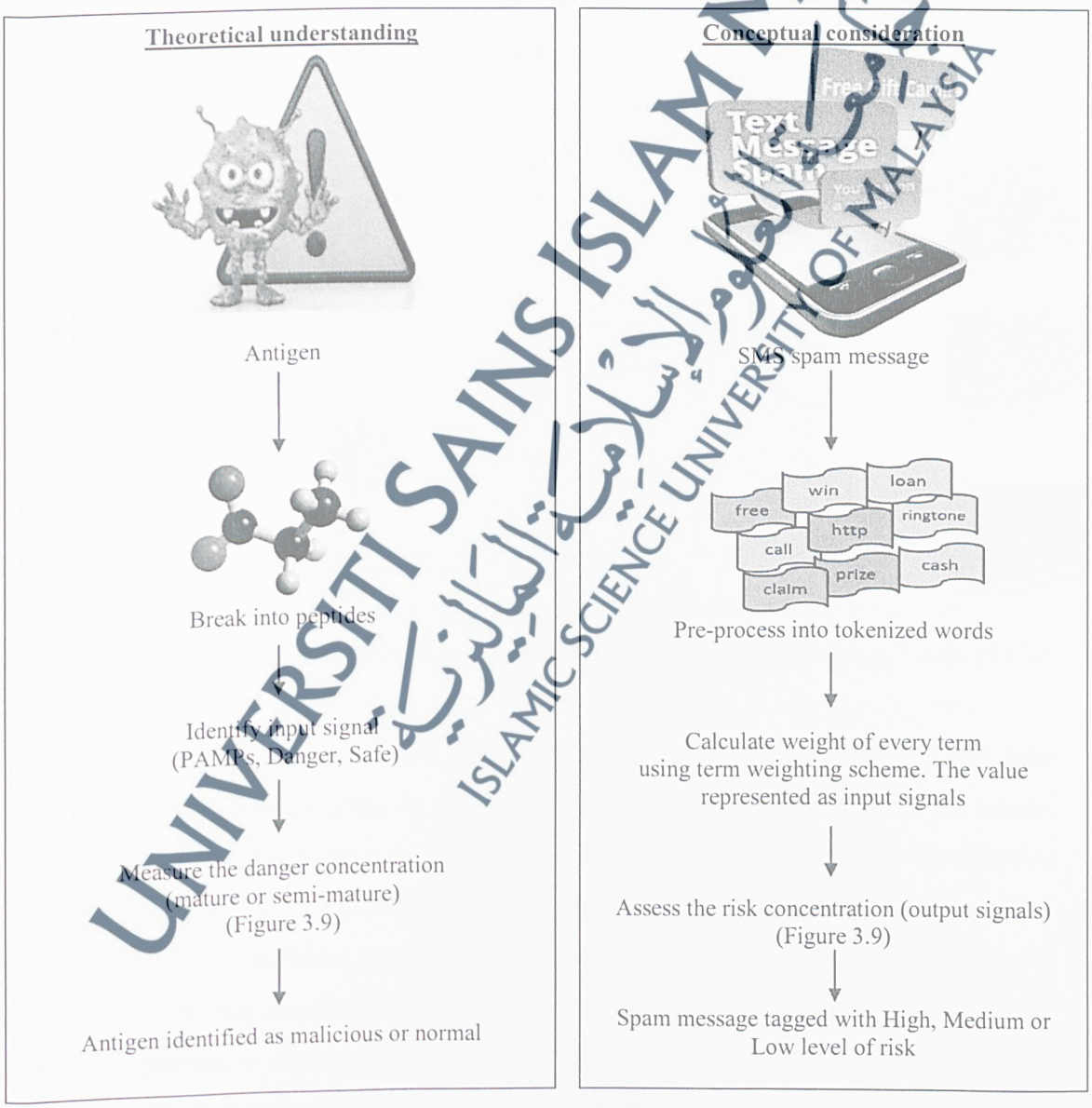


Figure 3.8: Theoretical Understanding In Contrast With The Conceptual Consideration

The theoretical understanding as depicted in Figure 3.8 is further elaborated in Figure 3.9 which identified biological abstraction as reflected in the process of signal transformation in spam risk assessment. In Figure 3.8, the theoretical understanding is portrayed in contrast with the conceptual consideration of applying Danger Theory in assessing the risk concentration for a text spam message.

The key feature in DCA is the application of Mature Content Antigen Value or MCAV defined as the mean value of context per antigen type. In SMS spam environment, this value is the frequency of possible spam term existence according to a category that finally will rank the severity level. This ranking is reflecting the degree of impact or loss.

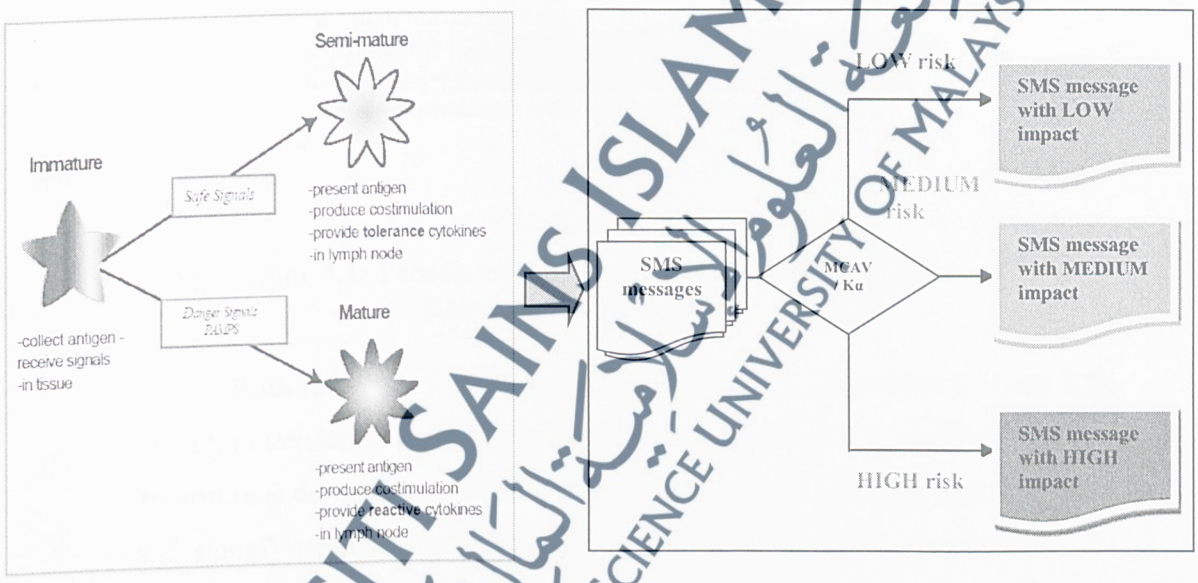


Figure 3.9: The Transformation Between Signals In Dendritic Cells (DCs)

The calculated MCAV is also defined as in numerical value which later is defined according to the range of risk value scale. The potential impact associated with the calculated risk is also defined, which is further described in Section 3.6.3.2.1Risk Scale Value Range.

The initial version of the algorithm for this research is depicted in the following Algorithm 3.1. The detailed algorithm which elaborated the whole process of danger assessment is explained in Chapter 4, projected with the development of the prototype, *in silico* algorithm.

```
1   Input :                 S = set of spam messages to be labelled as high, medium or low risk
2   Output :                L = set of spam messages labelled as high, medium or low risk
3   Begin
4                           Create a database of spam with risk indicator, D (database library)
5                           Create a folder to contain risk-labelled spam, M (folder for spam with risk level
                            indicator)
6   Identified risk →       for all spam messages, S in P do
7                             Create a set of spam messages from sample in D, P (spam folder)
8                             for all spam messages in P do
9                                 Add data item in D
10                                Update information on high, medium and low-risk level
11                                Update risk level of spam cluster with related spam term
12                                Migrate S from P to M and create a new data item in D, if current information
                                  not available
13                                S then become L
14                            end
15                          end
16                          for all L in M do
17                            Label L as to be high, medium or low risk
18                          end
19                          for all spam messages, S do
20  Assess risk level →       Calculate the risk value accordingly with the spam content
21  Prioritize risk →         Label spam messages with the high, medium or low risk
22                            Add spam messages with risk level indicator into M
23                          end
24  end
```

Algorithm 3.1: Pseudo-code Of Spam Risk-labelled Algorithm Based On Generic DCA

With reference to generic DCA in Chapter 2 (Algorithm 2.1 and 2.2), the initial version of the designed algorithm is depicted in Algorithm 3.1. This algorithm is designed and developed based on Figure 3.9, which inputs (antigen and signal) are referring to the incoming messages. While the output of the system is the SMS message that is tagged with the associated risk level that it is potentially carried with. This is shown as a cause-and-effect relationship that is believed to exist between signals and antigen. In this correlation, signals are the explicit effects that potentially result from the implicit cause of antigen (Gu et al., 2011).

The calculation of risk level for every spam message is further defined as to enhance this algorithm. This calculation essentially reflects the MCAV measurement that is introduced in the Danger Theory. Term weighting scheme that applied statistical calculation is employed to assigned related weight to the content of spam messages. This content of spam messages is treated as an

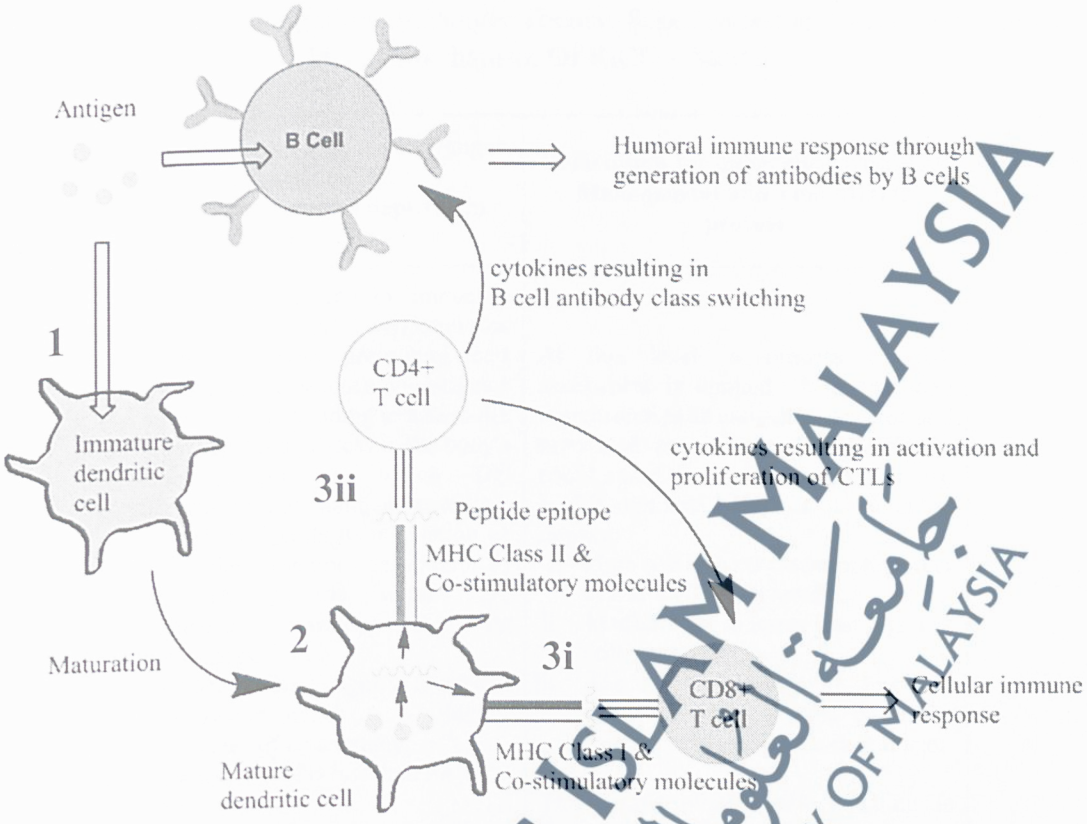
antigen and the weights derived from the schemes and risk scale form the three (3) input signals.

This initial version of the algorithm is further enhanced to measure the hazardous level of identified spam and articulated in Chapter 4. The risk assessment technique is integrated with this algorithm in order to design and develop an enhanced version of the existing Danger Theory algorithm that is applied in this field of study.

The general algorithm depicted in Algorithm 3.1 is enhanced in this proposed model. Integration with other techniques is suggested to fulfil the requirement of the original algorithm, which includes:

i. application of text mining that covers two (2) main tasks:
   - the pre-processing phases
   - the use of term weighting schemes to calculate the importance of relevant words that is represented as an input signal
ii. the application of risk scale to distinguish every level of risk
iii. the mapping of risk assessment process practiced in many standards of risk management of handling spam

In addition to that, the enhancement made in this proposed algorithm includes mature cell that is further divided into high and medium risk level to indicate the maliciousness. This is already shown in Figure 3.9. However, to have a better understanding, the integration between the Danger Theory, risk assessment, and text mining is depicted in Figure 3.10 and articulated accordingly in Table 3.4. For a specific articulation of DCA and dDCA version, a description of this algorithm is explained in Chapter 4.



Source: Zaman & Toth (2013)

Figure 3.10: Induction Of An Immune Response

Table 3.4: The Integration Of Danger Theory, Risk Assessment And Text Mining In The Mechanism Of RiCCA Model

Process no.	Biological process adapted in Danger Theory of AIS (chain of command as depicted in Figure 3.10)	Description for integration with Risk Management and Text Mining process	Generic algorithm (depicted in Algorithm 3.1)
1	The antigen is captured by immature dendritic cell (DC). DC is well known as professional antigen presenting cell (APC) and crime science investigator that survey its surrounding to assess the situation. APC is a process in the body's immune system by which DC capture and process antigens that later recognized by T-cells for activation or suppress the immune response. In Danger Theory, the signals are distinguished based on its anomalousness: i. PAMPs signal signify abnormal behaviour which highly indicates of an anomaly; ii. Danger signal is less than PAMPs; and iii. Safe signal released to indicate safe context.	At this level, a process of risk assessment is applied. Spam message represented as an antigen is captured and processed as a threat that potentially could cause damage. The levels of risk is differentiated based on its potential impact: i. High risk reflects catastrophe effect that is difficult to handle; ii. Medium risk is lesser than high but still dangerous; and iii. The level of hazard for low risk could be very minimal or nearly secure with no significant damage. The details of this risk levelling is described as in Table 3.12.	Line 6
2	The DC then has its maturation process that involves processing of antigen to display peptides on major histocompatibility complex (MHC) class molecules together with co-stimulatory molecules to T cells (Cytotoxic T-cells or Helper T-cells). The main function of MHC molecules is to bind to peptide fragments derived from the pathogens and display them on the cell surface (with receptors) for recognition by the appropriate T-cells. While co-stimulatory molecules are required in addition to the antigen-specific signal from their antigen receptors.	The process of breaking the antigen into small pieces (peptides) is the process of pre-processing in text mining for a spam message. The pre-processed spam messages are known as tokenized word instead of the peptide in biology. Term weighting schemes then are applied to calculate the weight value of every tokenized word which represented as the input signal.	Line 20
3i	Cytotoxic T-cells (killer) or CD8+ recognize and destroy infected cells through recognition peptides from within the cell. These peptides are displayed on MHC class I molecules. Presentation of appropriate peptide epitope on MHC class I molecules by matured DC results in a cellular response. This type of response does not involve antibodies.	The measurement of risk level via MCAV determine the damage level that potentially caused by the antigen. In this case, the antigen is assessed as malicious that finally is killed by the T-killer cells. In RiCCA, the spam message is classified as malicious and marked as high or medium risk (as depicted in Figure 3.9)	Line 21

Table 3.4, continued

Process no.	Biological process adapted in Danger Theory of AIS (chain of command as depicted in Figure 3.10)	Description for integration with Risk Management and Text Mining process	Generic algorithm (depicted in Algorithm 3.1)
3ii	Helper T-cells (TH) or CD4+ cells recognize peptides derived from extracellular proteins displayed on MHC class II molecules on DCs. CD4+ is the commander of the immune response and it detects infection and sounds the alarm for both initiating T-cell and B-cell responses.	At this stage, the antigen is assessed as safe by semi-mature DC. As adaptation in RiCCA model, the spam message is most likely will not cause any harm and the content is safe (negligible risk).	Line 21

3.6 Potential Influence Factors For Spam Classification

Greensmith & Aickelin (2008) declared that the algorithm which emerged from the Danger Theory is potential to be magnified if it is incorporated with other more sophisticated metric. Hence, in this study, a number of potential influence factors that may amplify the results of algorithms are identified. These factors include:

i. The role of pre-processing in reducing dimensionality complexity or data sparseness (Section 3.6.2);

ii. With text mining, each term is assigned with an associated weight value, accordingly to the chosen term weighting schemes. These terms and its weights (antigen and signal value) are represented as the input signals for the algorithm. The weights derived from the schemes also act as signal normalization for the DCA (Section 3.6.3.1);

iii. The risk scale for various range and multi-weights for signal transformation is tested to analyse the sensitivity of different weight value set in the algorithms as proposed in Greensmith (2007). Greensmith, in her thesis, discovered that DCA is sensitive to changes in the weights (Section 3.6.3.2); and

iv. The effect of antigen multiplication to verify its reliability to overcome the issue of antigen deficiency or insufficiency that caused a signal decay (Greensmith, Whitbrook, & Aickelin, 2010; F Gu, Greensmith, & Aickelin, 2008). The signal decay occurred due to the absence of an activating (danger) signal. The DCA required both activating and inhibitory (safe)

signals (Greensmith & Aickelin, 2008) for danger detection. The absence of activating signals causes the danger to impossibly to be detected (Section 3.7.2).

All of these identified factors have been studied in a preliminary testing. The findings are elaborated in the following section that later is considered in the development and implementation of the prototype.

3.6.1 Dataset Of SMS Spam

As elaborated in Chapter 2, there are three (3) main types of dataset sources. For this research, a dataset that is publicly shared and available online seems the best option. Besides, it is the largest collection for the English language, the findings from this work could be compared and analysed empirically with other methods for risk calculation. This dataset known as SMS Spam Collection v.1 can be accessed at the UCI Machine Learning Repository (Almeida & Hidalgo, 2012). This collection consists of a total 5,574 SMS messages which are collected from four (4) different sources.

Table 3.5: Origin Sources For SMS Spam Collection V.1

Source	Ham Messages	Spam Messages
Grumbletext[13]	0	425
National University of Singapore (NUS) Corpus[14]	3,375	0
Tagg's PhD[15]	450	0
SMS Spam's Corpus v0.1 Big[16]	1,002	322
Total	4,827	747

[13] http://co.uk-www.com/grumbletext.co.uk. accessed: 29 March 2016
[14] http://www.comp.nus.edu.sg/entrepreneurship/innovation/osr/corpus/. accessed: 29 March 2016
[15] http://etheses.bham.ac.uk/253/1/Tagg09PhD.pdf. accessed: 29 March 2016
[16] http://www.esp.uem.es/jmgomez/smsspamcorpus/. accessed: 29 March 2016

In another collection of SMS datasets such as British English Corpora (BEC)[23], the collected messages are in a smaller number of total messages which is 875 SMS messages only. It has been collected from the same sources of UCI Machine Learning Repository collection that are Grumbletext and Tagg's Ph.D. But in Dublin Institute of Technology (DIT)[24] SMS corpus, the collection consists of 1,353 SMS spam messages, some is not in the UCI Machine Learning Repository collection. This DIT spam collection has been utilized by Delany, Buckley, & Greene (2012) for research in clustering the spam messages into ten (10) categories of spam clusters. Via checking with WCopyfind[25], 517 out of 1,353 SMS spam messages are identified as unique compared to the UCI Machine Learning Repository collection.

Table 3.6: Statistics Of Instances In SMS Spam Collection V.1

	Number of instances	Percentage
Spam	747	13.4%
Non-spam (ham)	4,827	86.6%
Total	5,574	**100.0%**

Initialization phase is referring to the process of reducing noise, indexing, and to assign a term with the associated weight value via statistical calculation. This is kept as a database for the library. The self-collected dataset is used to verify the reliability of the model in assessing the risk of a previously unseen spam message.

Table 3.7: The Dataset Used For Initialization And Testing Phase

Dataset	Description
SMS Spam Collection v.1	All 5,574 messages used for initialization phase. Then, a very small portion from this dataset (10 spam messages) randomly selected for testing the functionality of the model and also as the preliminary requirement to verify its capability in assessing risk.
Self-collected	Used for testing purpose and not included during the initialization phase. These spam messages are used for testing the reliability of the model for classifying the risk of the unseen pattern (spam message).

[23] https://mtaufiqnzz.wordpress.com/british-english-sms-corpora/. accessed: 29 March 2016
[24] http://www.dit.ie/computing/research/resources/smsdata/. accessed: 29 March 2016
[25] http://plagiarism.bloomfieldmedia.com/wordpress/software/wcopyfind/. accessed: 31 March 2016

However, because spam messages continuously increase, data should be added constantly for a precise analysis (S.-E. Kim et al., 2015). Hence, unseen spam messages earlier can be used to create the initial population database especially when it is classified as false positive.

Other than that, different set of data may present different risks level. This is due to the content of the dataset, which dataset with more spam messages may retrieved numerous number of messages with high and medium risk level.

3.6.2 Data Pre-processing

Greensmith in her thesis (Greensmith, 2007) urged that the DCA demand the raw input data to be normalized and categorized into the corresponding signal categories. This process is vital as the DCA does not consist of a training phase. In text mining, this normalization process can be executed via pre-processing phase.

SMS messages may contain noise such as irrelevant words that do not contribute in mining the context. It is required to clean the noise or execute the text normalization via pre-processing prior to further mine the text contents.

In the process of text mining, pre-processing or also known as the pre-treatment process is one of the important stages. The entire cycle of text categorization that involves all stages includes preparation or collection of data or documents, pre-processing, feature indexing, feature filtering, text classification with algorithm and performance measure. Messages usually consist of various types of words, which are known as parts of speech. These texts may consist of adverbs, articles, conjunctions and many others that are possibly not significant for the context assessment. They are usually considered as irrelevant attributes and eliminated during the pre-processing phase in text mining (Al-Hassan & El-Alfy, 2015; Eshmawi, 2015; S.Kannan & Gurusamy, 2015; Wasilewska et al., 2014; Abbott, 2013). Some sample of parts of speech is tabulated in Table 3.8. Hence, pre-processing is a process that could distinguish between relevant and irrelevant attributes.

Samsudin, Puteh, Hamdan, & Nazri (2012) and Samsudin, Hamda, Puteh, & Nazri (2013) regard pre-processing as the normalization of the noisy

text. This process somehow reduces the high dimensionality of the data that commonly turned out to be the main problem in text mining. This issue can be overcome by executing pre-processing that alleviates the data sparseness problem (L. Zhang, Zhu, & Yao, 2004). This effect of noise is reduced by eliminating any irrelevant words during the stage of pre-processing that is necessary prior to text mining process.

Table 3.8: Part Of Speech

Part of Speech	Examples
adverbs	quickly, as
articles	a, an, the
conjunctions	and, but, however
interjections	hooray, ouch
prepositions	on, over, beside
pronouns	she, you, us

Greensmith & Aickelin (2009) articulated that as the dimensionality or size of the feature space increases, the number of detectors required to fully cover such spaces increases exponentially. In practicing this for a real world problem somehow will increase the computational cost such as storage and memory that are required in data processing.

Table 3.9: The Wordlist Created Using RapidMiner

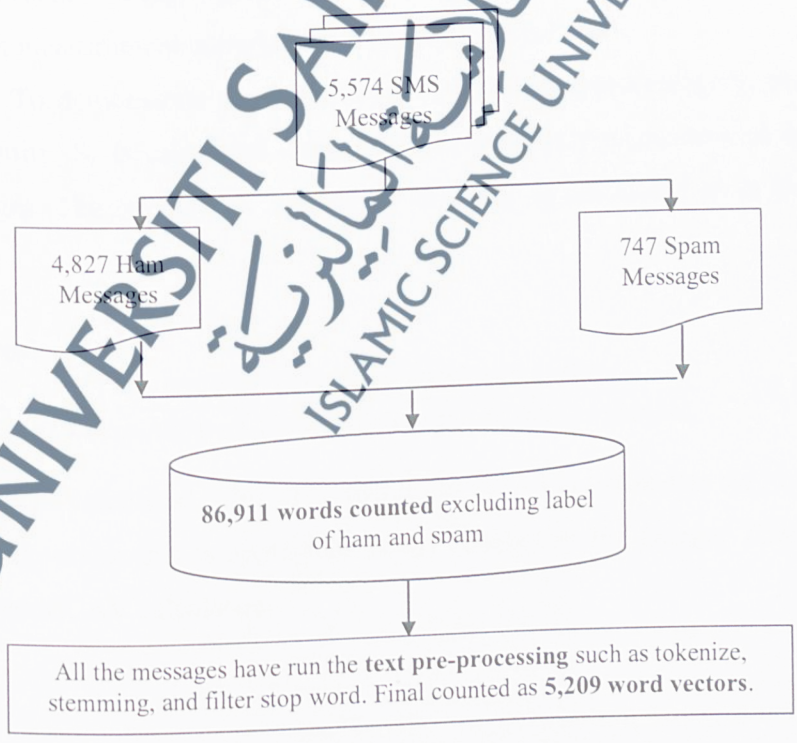


Table 3.9, continued

Spam Probability, *Sp*	Description for Frequency of Term Occurrence in SMS	Number of Tokens Calculated	Sample Tokens (Keyword)
1.00	Term occurred in spam messages only	745	claim, tone, prize, voucher, rington, bonus, http, quiz, winner, freemsg, discount
0.51– 0.99	Frequency of term occurred in spam > ham messages	207	www, award, txt, latest, statement, draw, cash, urgent, auction, sex, win, offer, click, chat, reward, date, free, price, access, profit, sms, sexi, flirt, luck, game, login, receipt, call, congrat
0.50	0.50 indicate that the term frequency equally occurred in both ham and spam messages	84	ton, refund, market, gold, allow, travel, tour, result, news, msg
0.01 – 0.49	Frequency of term occurred in spam < ham messages	401	help, visit, password, sale, download, account, cheaper, vote, trip, request
0.00	0.00 indicate that the term occurred in ham messages only	3,772	accept, admin, afford, blank, cheat, companion, entertain, hate, hug, jealous, visitor
TOTAL Word Vectors		**5,209**	

Referring to Table 3.9, it is shown that the pre-processing for 5,574 messages has tremendously reduced 94% of irrelevant words (81,702 words). This reduction of high dimensionality is expected to be significant in reducing the risk measurement complexity and computational cost

To demonstrate the importance of terms in spam category, the spam probability, S_p is calculated using the term frequency of occurrence in spam messages. The strength or importance of words is calculated as in Equation (3.1).

$$S_p(\textbf{term}) = \frac{frequency\ of\ \textbf{term}\ occurence\ in\ spam}{total\ frequency\ of\ \textbf{term}\ occurence\ in\ all\ messages} \quad (3.1)$$

The calculated value of S_p that is closer to 1 is defined as the higher the risk is detected; this is applicable to the concept in the Danger Theory that involved MCAV calculation.

3.6.3 Signals And Antigen: Identification Of The Reliable Term Weighting Schemes With Various Range Of Risk Value And Weights For Signal Transformation

Other than noise reduction via pre-processing, there are many potential factors that could influence and discriminate results of the simulation. These identified factors need to be verified for its reliability in contributing optimized result for assessing the risk concentration for text spam messages.

3.6.3.1 Input Signals Calculation With Term Weighting Schemes

In text mining, there is a requirement to apply the term weighting schemes. This scheme is assigning an appropriate weights value to the attribute of text contents. Prior to choosing the precise pre-selected schemes for further testing, it is mandatory to identify if the schemes are competent to offer the following criteria:

i. the higher the weight of an attribute, the more relevant it is associated with the spam category; and
ii. the weight calculated is normalized in between 0 to 1, to meet the requirement with the malicious measurement in DCA.

With the aforementioned characteristics, three (3) pre-selected schemes have been decided to be applied in this initial testing. In addition to that, these three (3) schemes have been successfully applied in spam detection, as discussed in Chapter 2 (Table 2.8). These schemes are Term Frequency (TF), Information Gain Ratio (IG Ratio) and Chi-square (CHI^2).

i. Term Frequency (TF) also known as term strength or word frequency. It measures the number of feature (term) in a category that appears in a corpus. TF alone is the unsupervised term weighting. For example, the probability of a word that possibly presents in spam messages can be measured as a ratio or proportional to word frequency in spam messages divided by word frequency in both spam and ham messages.
ii. Information Gain Ratio (IG Ratio) used as one of disparity measures and the high gain ratio for selected feature implies

that the feature is useful for classification. It applies normalization to information gain score by utilizing a split information value.

iii. Chi-square (CHI^2) measures the association between the word feature and its associated class or category.

Using RapidMiner, the same set of SMS messages corpus has been deployed to calculate the term weight employing with this three (3) pre-selected schemes. The top 10 of spam words is tabulated in Table 3.10, executed with the pre-processing and Table 3.11, executed without the pre-processing. This weight value is subsequently applied as input signals in the proposed model of RiCCA using the enhanced DCA algorithm. As previously elaborated, DCA requires both signal and antigen to be correlated for the context assessment. In this study, tokenized term is treated as an antigen and the calculated weight is the input signal.

Table 3.10: The Top 10 Spam Terms, With The Pre-processing

TF		IG Ratio		CHI^2	
Term (antigen)	**Weight (input signal)**	**Term (antigen)**	**Weight (input signal)**	**Term (antigen)**	**Weight (input signal)**
claim	1.0000	claim	1.0000	call	1.0000
prize	1.0000	prize	0.9475	txt	0.8607
tone	1.0000	www	0.9148	free	0.7476
guarante	1.0000	txt	0.8844	claim	0.6518
ppm	1.0000	uk	0.8681	mobil	0.6474
cs	1.0000	tone	0.8578	www	0.5703
pobox	1.0000	guarante	0.8448	prize	0.5181
ringtone	1.0000	ppm	0.8381	text	0.4745
entri	1.0000	co	0.8242	uk	0.4272
polif	1.0000	cs	0.8206	stop	0.4111

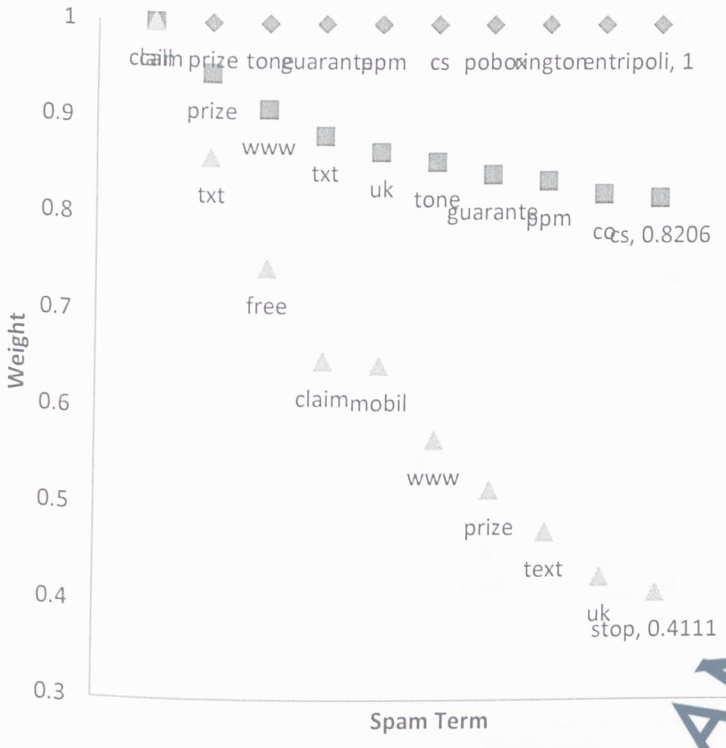


Figure 3.11: The Scatter Plot For Top 10 Spam Terms, With The Pre-processing

Table 3.11: The Top 10 Spam Terms, Without The Pre-processing

TF		IG Ratio		CHI^2	
Term (antigen)	**Weight (input signal)**	**Term (antigen)**	**Weight (input signal)**	**Term (antigen)**	**Weight (input signal)**
p	1.0000	Â	1.0000	Â	1.0000
claim	1.0000	p	0.9568	p	0.6837
prize	1.0000	FREE	0.8086	call	0.5111
co	1.0000	claim	0.7911	Call	0.5063
ppm	1.0000	www	0.7798	FREE	0.4034
Cs	1.0000	prize	0.7721	to	0.3920
URGENT	1.0000	STOP	0.7440	www	0.3899
tone	1.0000	co	0.7157	or	0.3647
awarded	1.0000	ppm	0.7127	mobile	0.3390
Box	1.0000	Cs	0.7065	claim	0.3157

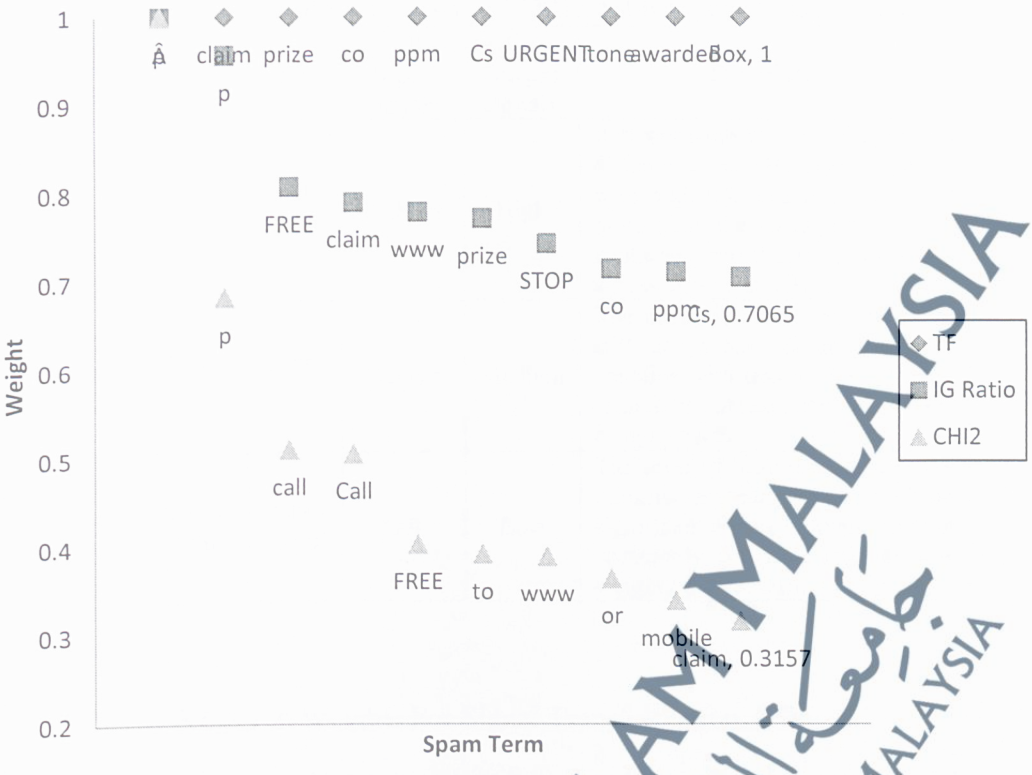


Figure 3.12: The Scatter Plot For Top 10 Spam Terms, Without The Pre-processing

3.6.3.2 Weight For Input And Output Signals

Greensmith (2007) reported in her thesis that DCA is sensitive to a weight value and it is reacted differently for various weights. For sensitivity analysis of weights, this testing also executed a various range of risk scale and multi weights for signal transformation from input to output signals. This objectively to identify which value is producing an optimized result of the proposed algorithm and model.

3.6.3.2.1 Risk Scale Value Range, S

This testing is applied in two (2) different range value for risk scale as tabulated in Table 3.12. This is employed during the malicious measurement. These risk scales are used for translating the numerical value into three (3) different levels for input and output signals.

Table 3.12: Two (2) Different Ranges Of Risk Scale

S1	S2	Risk level		Description
		Input signals	Output signals	
1.00 - 0.70	1.00 - 0.80	PAMPs	High	Reflects catastrophe effect that is difficult to handle. Spam message with high risk usually consists of more than one spam term. It also containing text that request users to access any given URL.
0.69 - 0.40	0.79 - 0.50	Danger	Medium	The effect is lesser than high but still dangerous. Spam messages contains with less weight of spam term in its context such as call back or reply SMS.
0.39 - 0.00	0.49 - 0.00	Safe	Low	The level of hazard could be very minimal or nearly secure with no significant damage. Spam message commonly contains with the low weight of spam term (negligible).

3.6.3.2.2 Weights For Signal Transformation, WM

Weights to transform input signals into output signals are one of the mandatory prerequisites for signal processing intentionally for the algorithm. The higher the transforming weights are depicted as the thicker line, as illustrated in Figure 2.5 of Chapter 2. These weights values are applied in Ding, Yu, & Yang (2013); Pereira (2011), and Kim, Bentley, Wallenta, Ahmed, & Hailes (2006). The transforming weight applied for this testing is tabulated in Table 3.13. This weight is only applied for risk assessment using DCA since dDCA mechanism does not require weight for signal transformation.

Table 3.13: Two (2) Different Matrices For Transforming Weights

Signals	WM1			WM2		
	PAMPs	Danger	Safe	PAMPs	Danger	Safe
CSM	1	0.5	1.5	2	0	2
smDC	0	0	1	1	0	1
mDC	1	0.5	1.5	2	3	3

3.6.3.2.3 Anomaly Threshold, t_m And T_k

For the simulation explained in this Chapter 3, the given value for anomaly threshold (both for DCA and dDCA) is set to 0.1340. This value is derived from the dataset proportion of the initial population, and the number of spam messages is divided by number of total messages (747 divided by 5,574). There are a few more options to set and test the value for anomaly threshold, which is further explained and tested in Chapter 5.

3.7 The Classifier: Empirical Differences Between DCA And dDCA

3.7.1 Dendritic Cell Algorithm (DCA) And Deterministic Dendritic Cell Algorithm (dDCA)

As elaborated in Gu, Greensmith, & Aicklein (2011), there are six (6) versions of DCA extension in its development pathway which is Prototype DCA (pDCA), Libtissue DCA (ltDCA), Deterministic DCA (dDCA), Extended DCA (xDCA), Robotic DCA (rDCA) and Integrated DCA (iDCA). Every version has been applied specifically in certain fields and it is recorded that certain versions of the DCA is successfully applied, in particular to some domains such as classification and robotic.

This research only employed the original version of DCA and dDCA with the following justifications:

i. The developed version other than DCA and dDCA has loads of stochastic decisions (randomly determined) and they are complex algorithms (Greensmith & Aickelin, 2009). For instance, in Gu et al. (2011), ltDCA contained too many arbitrary and random components and made it behave as an infeasible algorithm;

ii. DCA and dDCA is suitable for anomaly classification compared to others which

- pDCA applied to binary classification problem and capable of performing 2-class discrimination on an ordered dataset;

- ltDCA applied to problems in computer and network security including port scan detection and sensor network security. With some modification ltDCA is also applied as a robotic classifier for physical security;
- rDCA is applied to solved problems in the robotics field (Oates, Kendall, & Garibaldi, 2008a).

iii. dDCA that is a simpler version than another version could potentially minimize the computational cost since a large amount of randomness has been removed. DCA and dDCA are much simpler than other variants and dDCA has reduced numbers of parameters compared to DCA. Greensmith & Aickelin (2008) and Brownlee (2011) identified that complexity has been reduced in dDCA to make it more feasible and amenable to analysis. And even though numerous parameters has been removed, there is no reduction in algorithm performance as clarified in Greensmith & Aickelin (2009).

3.7.2 Antigen Multiplication

The antigen multiplication is applied with the objective to verify its effect in optimizing the accuracy rate. As elaborated in F Gu, Greensmith, & Aickelin (2008), the antigen multiplication is implemented to overcome the problem of antigen deficiency or insufficient antigens that are supplied to the initial DC population. This antigen multiplication makes several copies of each individual antigen which can be fed to multiple dendritic cells. However, a different findings in Gu et al. (2008) shown that DCA could not be optimized using antigen multiplication.

In this testing, the same 747 spam messages (antigen) are multiplied or copied for 10, 40 and 100 times prior to creating the initial population. The process for severity assessment is repeated with multiplied antigens and the effect of it is identified.

3.8 Experimental Setup For Testing The DCA And dDCA

Prior to the prototype development, the design of the RiCCA is required to be evaluated and verified. This can be done via conducting a series of simulations. This simulation executed with the following arrangements with all of three (3) pre-selected term weighting schemes are applied at all rounds using DCA.

Table 3.14: Four (4) Series Of Simulated Experiments

Experiment No.	Range of Risk Value Scale (S)	Weight Matrix (WM)
1	S1	WM1
2	S1	WM2
3	S2	WM1
4	S2	WM2

The testing as tabulated in Table 3.14 then is repeated, but with the dDCA classifier. This is conducted purposely to execute some comparable analysis between DCA and dDCA. Finally, the test is conducted with multiplied antigens employing the best-identified risk value scale and weight matrix. A detailed of experiment process flow is depicted in the following Figure 3.13.

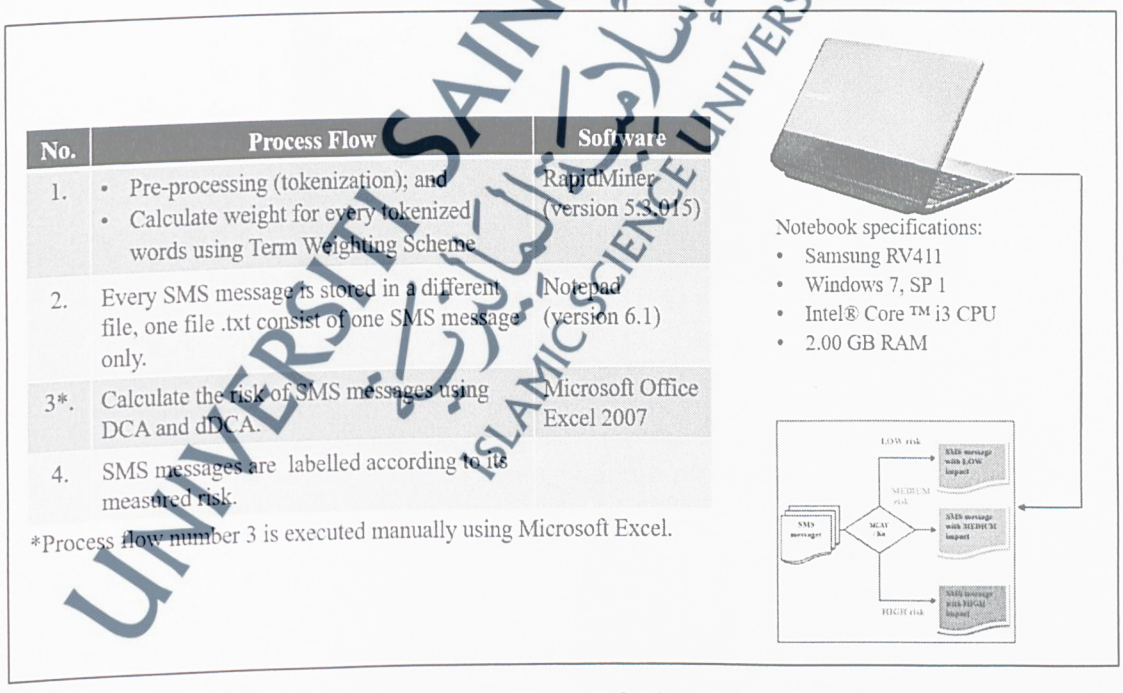


Figure 3.13: The Process Flow Of Simulated Experiment

3.9 Results And Findings Of Initial Simulation

The designed model is to test for initial simulation and the experimental setup is arranged as described in Section 3.6 – 3.8 and simplified as tabulated in Table 3.14.

In this simulation, a metric of True Positive (TP) is used to measure the performance; which total number of spam messages is identified precisely as malicious. Malicious is defined with these two (2) conditions:

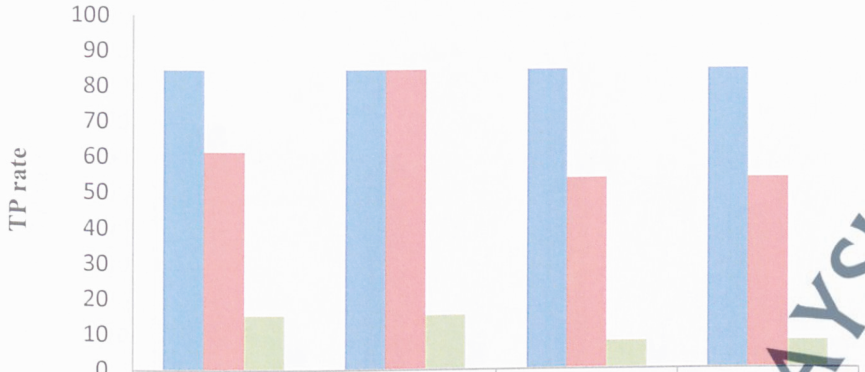
i. matured cell consists of high and medium tokens; and

ii. matured cell consists of high tokens only.

Table 3.15: TP Rate For Matured Cell Consists Of High And Medium Tokens

Scale, S and Signals Weight Matrix, WM	TP, with pre-processing (%)			TP, without pre-processing (%)		
	TF	IG Ratio	CHI^2	TF	IG Ratio	CHI^2
S1 and WM1	**100**	69.2	30.7	**84.6**	61.5	15.4
S1 and WM2	**100**	76.9	30.8	**84.6**	**84.6**	15.4
S2 and WM1	**84.6**	69.2	23.1	**84.6**	53.8	7.7
S2 and WM2	**92.3**	46.2	23.1	**84.6**	53.8	7.7

(a) TP rate, with pre-processing



	S1 & WM1	S1 & WM2	S2 & WM1	S2 & WM2
TF	84.6	84.6	84.6	84.6
IG Ratio	61.5	84.6	53.8	53.8
CHI2	15.4	15.4	7.7	7.7

(b) TP rate, without pre-processing

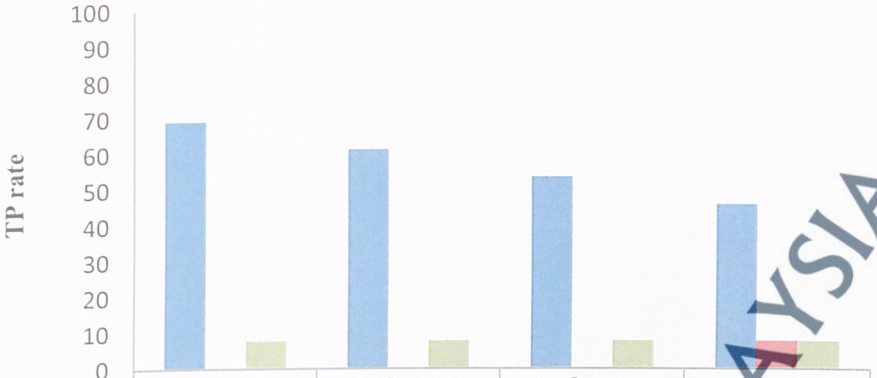
Figure 3.14: TP Rate With Different Weighting Schemes, Multiple Values For Risk Scale And Transforming Weights. These Figures (a) And (b) Corresponds To Table 3.15

Table 3.16: TP Rate For Matured Cell Consists Of High Tokens

Scale, S and Signals Weight Matrix, WM	TP, with pre-processing (%)			TP, without pre-processing (%)		
	TF	IG Ratio	CHI^2	TF	IG Ratio	CHI^2
S1 and WM1	76.9	30.7	15.4	69.2	0	7.7
S1 and WM2	76.9	30.8	15.4	61.5	0	7.7
S2 and WM1	46.2	30.8	15.4	53.8	0	7.7
S2 and WM2	46.2	23.1	15.4	46.2	7.7	7.7

	S1 & WM1	S1 & WM2	S2 & WM1	S2 & WM2
TF	76.9	76.9	46.2	46.2
IG Ratio	30.7	30.8	30.8	23.1
CHI2	15.4	15.4	15.4	15.4

(a) TP rate, with pre-processing



	S1 & WM1	S1 & WM2	S2 & WM1	S2 & WM2
TF	69.2	61.5	53.8	46.2
IG Ratio	0	0	0	7.7
CHI2	7.7	7.7	7.7	7.7

(b) TP rate, without pre-processing

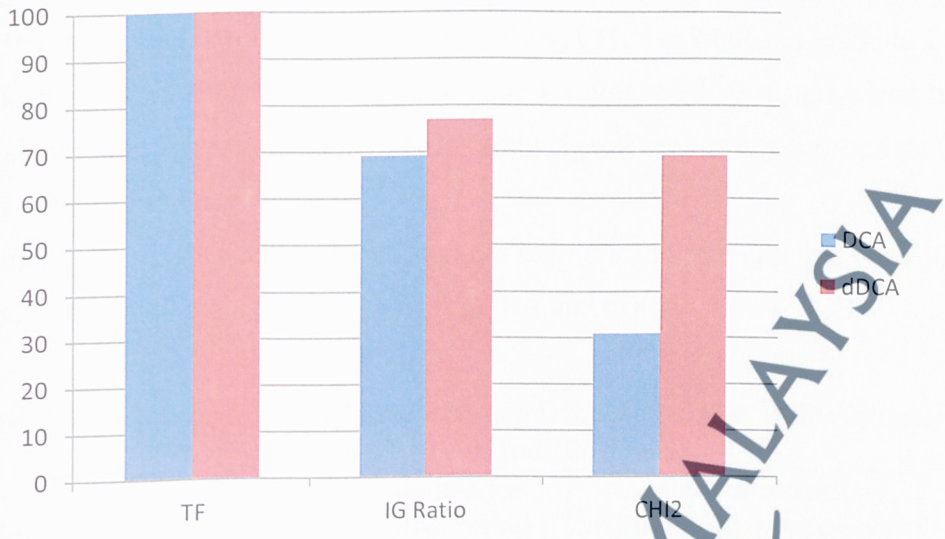
Figure 3.15: TP Rate With Different Weighting Schemes, Multiple Values For Risk Scale And Transforming Weights. These Figures (a) And (b) Corresponds To Table 3.16

Referring to Table 3.15 and 3.16, it is shown that TF is the best suited and resulted for the highest rate of TP. The rate even optimized in its risk classification when the process is considered the pre-processing phase and counted both high and medium tokens as the mature DCs. Other than that, the risk scale value range and transforming weights also giving significant effect to the high rate of precise classification.

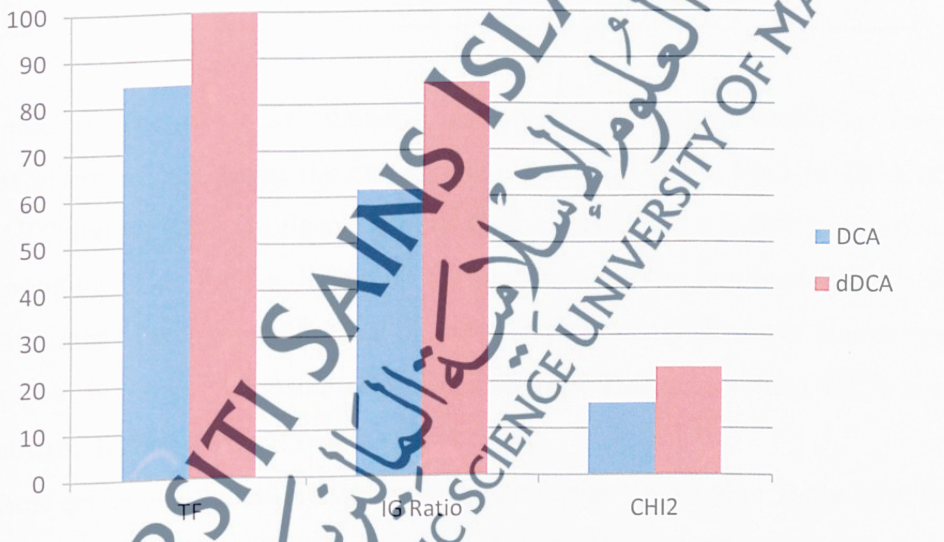
Then, the test is continued with all three (3) weighting schemes and best-identified risk scale (S1) from previous testing is employed with dDCA as the classifier.

Table 3.17: Comparison Of TP Rate Between DCA And dDCA

Weighting Schemes	DCA (High & Medium for MCAV)		dDCA (K_α)	
	With Pre-Processing	Without Pre-Processing	With Pre-Processing	Without Pre-Processing
TF	**100**	84.6	**100**	**100**
IG Ratio	69.2	61.5	76.9	84.6
CHI^2	30.8	15.4	69.2	23.1



(a) TP rate, with pre-processing



(b) TP rate, without pre-processing

Figure 3.16: The TP Rate Value For DCA And dDCA With S1 And WM1. These Figures (a) And (b) Corresponds To Table 3.17

According to the simulation result, as tabulated in Table 3.17 and graph depicted in Figure 3.16, the classifier dDCA shows a higher and optimized TP classification rate compared to DCA, especially with TF as weighting scheme, both with pre-processing and vice versa. From these simulations, it is obvious that CHI^2 is not a suitable

weighting scheme to be applied in this field of the domain (spam risk assessment). With referring to the top ten (10) of calculated weights using CHI^2 (as tabulated in Table 3.10 and 3.11), this scheme demonstrated a weak value for almost all spam terms and this caused a signal decay (less availability of activating signals) that is not appropriate for risk assessment in detecting danger.

Finally, the test is continued with antigens multiplication for creating the initial population with 10, 40 and 100 times, both for DCA and dDCA.

Table 3.18: Comparison Of TP Rate Between DCA And dDCA With Antigen Multiplication Of 10, 40 And 100 Times

Classifier	Weighting Schemes	Antigen multiplier with pre-processing			Antigen multiplier without pre-processing		
		10x	40x	100x	10x	40x	100x
DCA	TF	-na-	-na-	-na-	50	50	50
	IG Ratio	0	100	100	0	80	60
	CHI^2	0	0	0	0	0	0
dDCA	IG Ratio	0	100	100	50	50	50
	CHI^2	0	0	0	0	0	0

Based on the result as tabulated in Table 3.18, antigen multiplier has a significant effect of increasing the classification accuracy value, both in DCA and dDCA for TF and IG Ratio. Antigen multiplication revealed that it is able to contribute its corresponding value for the derivation of input signals. The saturated value is the factor of 40 times of antigen multiplication, with pre-processing. However, this antigen multiplier does not affect the result of the CHI^2 scheme. This shows that CHI^2 is not a suitable scheme for the derivation of the input signal.

These series of simulation have its final outcome as tabulated in Table 3.19. For a sample of comparative analysis between DCA and dDCA, some messages of risk concentration for context assessment is shown in this table. The characteristics for the experimental setup are stated as follows:

- Term weighting scheme: Term Frequency (TF)
- Matured content referred to both high and medium tokens
- Text pre-processing: Applied
- Using S1 as risk scale and WM1 as the transforming weights
- Antigen multiplication: Not applied.

Table 3.19: *Outcome Of Risk Concentration Calculation Using DCA And dDCA*

Text ID No.	Content of the spam messages	Measured risk level using DCA			Measured risk level using dDCA		
		Output signal	Risk value	Risk level	Output signal	Risk value	Risk level
S614.txt	Wanna have a laugh?Try CHIT-CHAT on your mobile now! Logon by txting the word: CHAT and send it to No: 8883 CM PO Box 4217 London W1A 6ZF 16+ 118p/msg rcvd	mDC>smDC	0.83	High	K_a>T_k	0.63	Medium
S712.txt	If you don't, your prize will go to another customer. T&C at www.t-c.biz 18+ 150p/min Polo Ltd Suite 373 London W1J 6HL Please call back if busy	mDC>smDC	1.00	High	K_a>T_k	0.81	High
S509.txt	Congratulations U can claim 2 VIP row A Tickets 2 C Blu in concert in November or Blu gift guaranteed Call 09061104276 to claim TS&Cs www.smsco.net costÂ£3.75max	mDC>smDC	0.89	High	K_a>T_k	0.72	High
S75.txt	Your credits have been topped up for http://www.bubbletext.com Your renewal Pin is tgxxrz	mDC>smDC	1.00	High	K_a>T_k	0.86	High
S181.txt	You have 1 new voicemail. Please call 08719181503	mDC>smDC	1.00	High	K_a>T_k	0.78	High
S230.txt	500 free text msgs. Just text ok to 80488 and we'll credit your account	mDC>smDC	1.00	High	K_a>T_k	0.61	Medium
S36.txt	Text & meet someone sexy today. U can find a date or even flirt its up to U. Join 4 just 10p. REPLY with NAME & AGE eg Sam 25. 18 -msg recd@thirtyeight pence	mDC>smDC	1.00	High	K_a>T_k	0.57	Medium
S533.txt	Bored housewives! Chat n date now! 0871750.77.11! BT-national rate 10p/min only from landlines!	mDC>smDC	0.75	High	K_a>T_k	0.52	Medium
S15.txt	Did you hear about the new "Divorce Barbie"? It comes with all of Ken's stuff!	mDC<smDC	0.00	Low	K_a<T_k	-0.11	Low
S111.txt	Romantic Paris. 2 nights, 2 flights from A£79 Book now 4 next year. Call 08704439680Ts&Cs apply.	mDC>smDC	0.67	Medium	K_a>T_k	0.40	Medium
ES292.txt	RM0.00. FREE RM15k travel voucher! Own a spacious Townvilla with ZERO entry cost & EARN up to RM20k for home owner programme. Call 0123568311 for details. *T&C apply	mDC>smDC	0.83	High	K_a>T_k	0.54	Medium
ES453.txt	ZI6036S Bonus code RM10 ncity888.com http://goo.gl/eR0CI5 12 win Newtown SCR 888 Sun city 100% Fast Withdraw Whatapps 0182546092 Wechat: csncity8	mDC>smDC	1.00	High	K_a>T_k	0.94	High
ES435.txt	Congratulations, you have successfully activated your TM Rewards membership account. To log into TM Rewards, go to www.tm.com.my/tmrewards	mDC>smDC	1.00	High	K_a>T_k	0.59	Medium

Referring to the risk concentration value as tabulated in Table 3.19, it is clearly proven that dDCA produced a better risk level classification in terms of numerical granularity. The combination of more than one risky term may result in high risk such as *free, call, text back* or messages that contain *URL* requesting users to access it. Considering these messages that contain information requested users to response (*call, text,* and *accessing URL link*) are in the precise risk, high and medium. While messages that contain no need to response (for instance a message with ID 15.txt) produced a low-risk concentration level or perhaps can be considered as a legitimate message.

3.10 Summary

The creation of AIS involves the translation of basic immunological models into feasible algorithms. Chapter 3 defined the original DCA into an enhanced version for risk assessment of text spam message and namely as RiCCA or Risk Concentration for Context Assessment. This is executed with conducting a series of simulation with the deployment of the small size of the dataset for testing. This is to design the model and manually test it prior to developing the prototype using computer programming for automatic simulation that afterward will be employed using a larger set of data.

Once the prototype is developed (Chapter 4), a large scale of testing is conducted with performance measurement which is articulated in Chapter 5.

From the simulation presented in this Chapter, there are a few discriminatory factors that control and influence the TP value of risk classification and ensure the success functionality of the classifier. These factors are considered in the design and development of the RiCCA prototype, which are:

i. the pre-processing of text obviously returns a higher TP rate of risk classification;

ii. the precise term weighting schemes give the effective weight for input signals as required in the Danger Theory. For instance, findings from the testing remarked CHI^2 as an unreliable scheme for this task since it showed signal decay even for important spam words. Application of antigen multiplication for CHI^2 also showed that this scheme is unsuitable for input signal derivation;

iii. different risk value range and weight for signal transformation (from input to output signals) does have its significant effects on the accuracy of risk classification. This is showed that the DCA classifier is sensitive to weight value changes;

iv. matured cell that consists of both high and medium tokens resulted in higher accuracy rate compared with consideration of high tokens only as matured cell;

v. the variants of the Danger Theory (DCA and dDCA) has a slightly different in terms of calculation for risk concentration and complexity;

vi. dDCA produced a finer grained risk concentration value (numerical information granularity) compared to DCA; and

vii. antigen multiplication is applied to overcome the problem of antigen deficiency and also contributed in accuracy classification.

The taxonomy of research that becomes the focus points of this study is represented in Figure 3.17. It shows the interconnection between phases of SMS spam management and its relation to the dataset that later subsequently fed as input signals for the DCA classifier. Considering the literature presented in Chapter 2, the taxonomy of direction of this research is constructed as follows, which the filled box with bold font is the focus of the study.

The RiCCA prototype is developed in JAVA for simplifying the testing with the larger size of the dataset. The development of this prototype is elaborated in Chapter 4 and further tested (in Chapter 5) to verify the findings as claimed and discussed in this chapter.

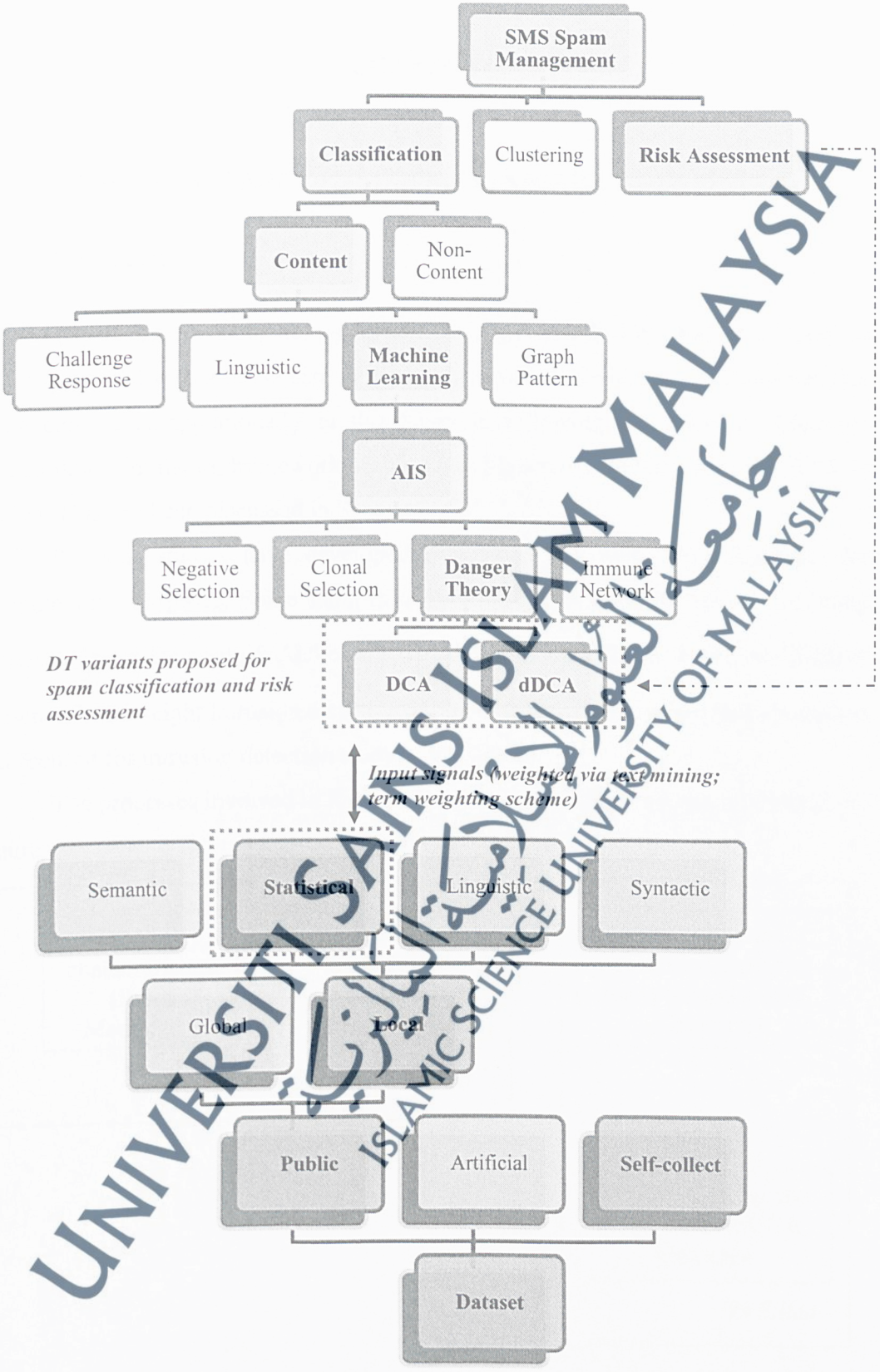


Figure 3.17: The Taxonomy Of Research Direction