CHAPTER 3
SPAM MANAGEMENT MODEL

3.1 BACKGROUND

This chapter presents the proposed model for managing text spam. The discussion on how the AIS algorithms will be used in spam management model and the relationship between BIS is also included in this chapter. Detail discussion on the Clonal Selection and Immune Network Theory of BIS are also presented in this chapter.

3.2 INTEGRATED MOBILE SPAM MODEL

A spam management model known as the Integrated Mobile Spam Model (IMSM) is to manage 'spam' text messages. This model consists of three phases with different types of AIS algorithms used, as shown in Figure 3.1.

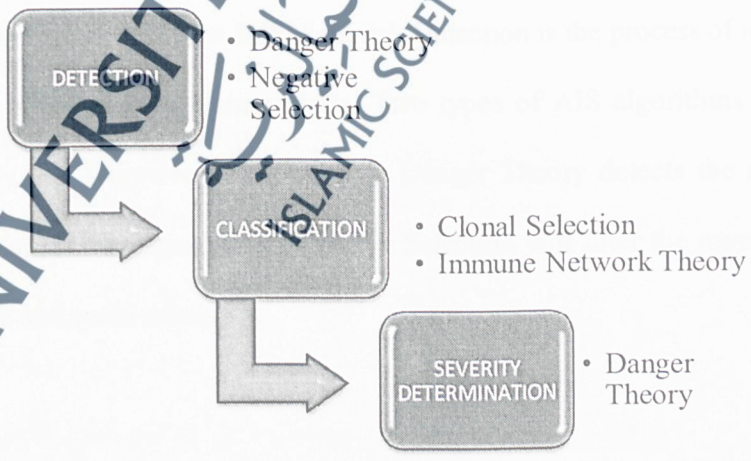


Figure 3. 1: Integrated Mobile Spam Model (IMSM).

In spam management process, most of the previous researches investigated only on how to detect and filter spam messages into ham and spam and the process ends after finishing identifying the messages without further study on the spam messages. Based on this situation, a new spam management model named as IMSM is introduced to explore more about managing spam messages. Figure 3.1 presented IMSM, and this model consists of three phases; detection, classification (i.e. clustering) and severity determination. This model is created because it uses the same algorithms in every phase (i.e. AIS algorithms) and this motivates us to explore the capability of the AIS algorithms in managing spam for each phase. AIS algorithms are chosen because they are inspired from human immune system to defend body from any virus or pathogens that can cause diseases or sickness. Thus, this proposed model wants to apply the same concept for identifying 'spam' text messages in mobile phone. This model is different from the existing models such as Risk Management Model and Intrusion Detection System in terms of the algorithms used in every phase and stage of the model.

Detection is the first phase in the IMSM model. Detection is the process of identifying the SMS message whether it is spam or ham. Two types of AIS algorithms are used; the Danger Theory and Negative Selection. The Danger Theory detects the messages that show they are spam messages while Negative Selection will filter the messages in order to identify ham and spam messages.

The second phase in IMSM model is classification or also known as clustering. Classification is the process to categorize and cluster the spam messages into several types of spam according to set rules. Clonal Selection and Immune Network Theory are used for this phase. Clonal Selection will clone the spam messages into several groups meanwhile the Immune Network Theory identifies which messages has close relationship and meaning with groups of spam depending on content and meaning of the messages.

The last phase is severity determination. This phase is to determine the level of danger for identified spam messages according to types of spam messages. The proposed level of danger can be divided into low, medium and high and the Danger Theory is used for this process. Danger Theory is used in both detection and severity determination phase. As the function of this theory is to detect and identify the danger signal, it help in detection phase to identify the danger messages which is spam messages while for severity determination, it help in determine the level of danger for the spam messages (i.e. identify details about the type of danger for spam messages).

The proposed IMSM model involves three phases and each phase is related to each other. For example, before doing classification of spam messages into several groups, detection process is needed to identify spam messages and determine the level of danger for each messages; it needs to identify the types and groups of spam. From this model, a spam model to integrate spam mobile is shown in Figure 3.2.

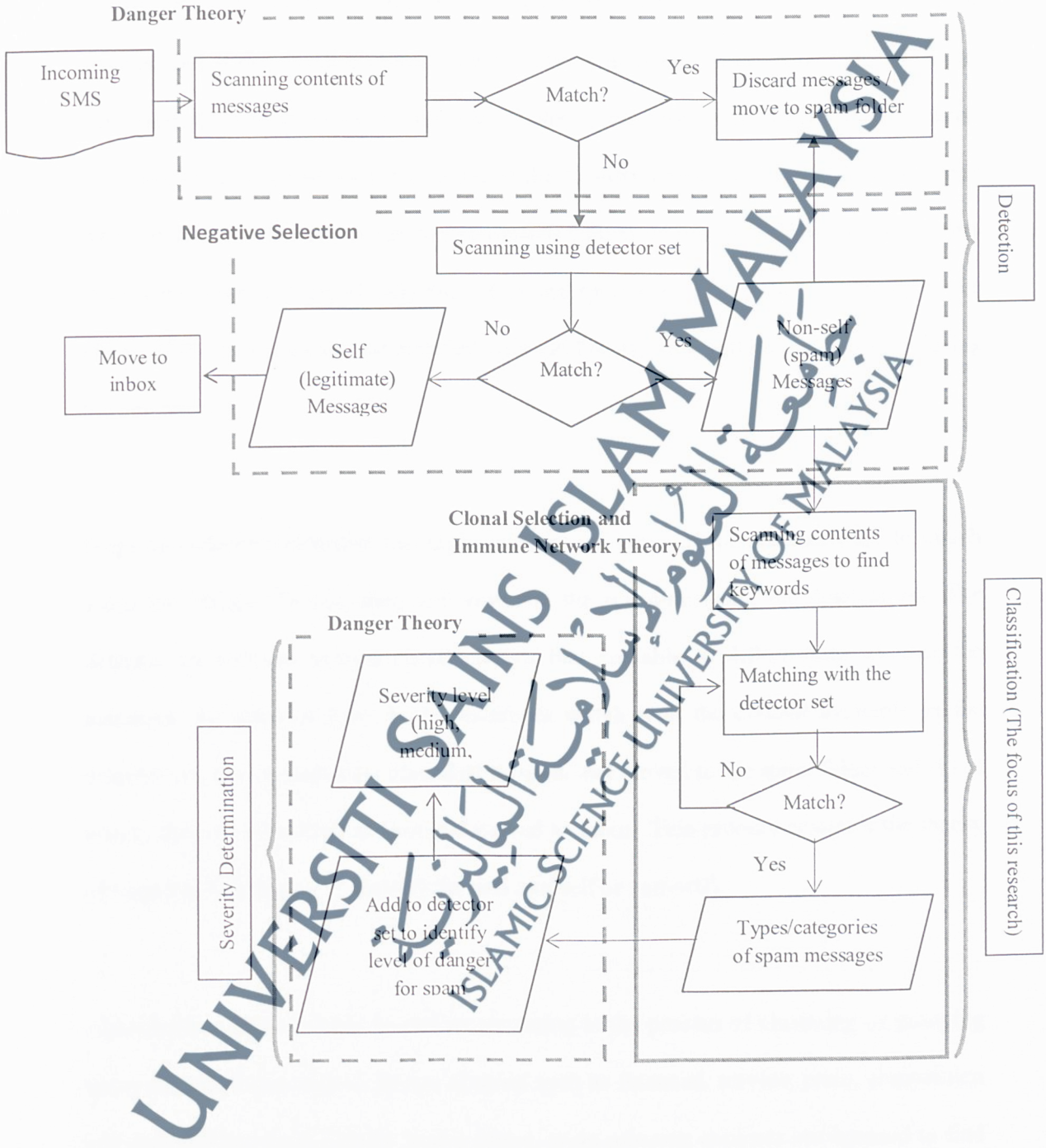


Figure 3. 2: Block diagram for spam model

Figure 3.2 shows the block diagram for the spam engine. Detection phase is the process to identify the SMS messages either ham or spam. The contents of incoming SMS will be scanned and if the contents match with the library available in the server, the messages automatically will be moved to the spam folder or the server will discard the messages. This process is inspired by the concept of the Danger Theory to detect spam messages where it detects cells that are already showed as danger. In this model, the Danger Theory detects the messages that already show they are spam (i.e. danger) before coming to the phone. If the messages are not matched with the library, they will be scanned again using the detector set.

Negative Selection identifies messages into ham or spam. If the messages fail to match using the Danger Theory, they will continue the scanning using the detector set. The detector set contains several characteristics that are able to differentiate whether the messages are spam or ham. If the messages match with the content available in the detector set, the messages are classified as spam and moved to the spam folder and if not match, they are classified as ham and moved to inbox. This process emulates the theory of Negative Selection (i.e. identify the cell into self or non-self).

Classification phase or also known as clustering is the process of clustering or grouping spam messages into several groups of spam such as financial, service, prize, competition and chat (Delany et al., 2012). In this phase, spam message contents are scanned to find the keywords for each group in the detector set. If the messages match with the detector

set, the messages will go to their defined categories and if not match, it will be scanned again in the detector set due to updated contents of messages in line with the transformation of technology. Clonal Selection occurs when the messages are divided into their own categories after the scanning process. Clonal Selection is the process of cloning cells into several cells and in this spam model, cloning process is the process of grouping spam messages into several groups.

Immune Network Theory is the interaction between antigens with antibody or among themselves. In this model, it happens when the interaction between the contents of the messages with keywords is available in the detector set. Further discussion is presented in sub-section 3.3.

Severity determination is the process to identify the level of danger for each spam message according to the group and cluster of spam. Each cluster of spam will calculate its level of danger in the detector set and results will show whether the level of danger is low, medium or high. Danger Theory is the process of identifying cells that are already in danger so the same concept is used for this model to identify the level of danger for each spam message. Table 3.1 summarizes the AIS algorithms that are being used in the spam management model and their relationship with BIS.

Table 3. 1: The implementation of AIS in the proposed model

Phase	AIS Algorithm	Process	Implementation from BIS
Detection	Negative Selection	- This theory detects self and non-self antigens. - Self is defined as legitimate messages while non-self as spam messages. - To detect the messages that receive self or non-self based on some criteria such as pattern and size of messages.	Process occurs in thymus, an organ of lymphatic system. In the negative selection, T-cells with high affinity (i.e. the strength) interaction with the thymic dendritic cells are eliminated and those with intermediate affinity survive.
	Danger Theory	- This theory only detects danger signal. - Assumes that danger signal is the messages that are already show it is spam based on the behaviour and characteristic of messages.	Process occurs in cell-mediated immunity and it detects a cell that is already in danger and affected by virus.
Classification	Immune Network Theory	- This theory is about the connection network between one component to other same component or to different component (example: the connection between an antibody with an antigen). - If spam messages (i.e. antigen) bind with correct detector (i.e. antigen) the messages will go to correct cluster while if they incorrectly bind with different detector, the process of binding will be repeated again using another detector. - Example: Detector for fraud binds with spam messages is categorized as pornography, so it will destroy the messages, but if it binds with spam messages categorized as gift, no response will happen.	The interaction between antibody and antigen or antibody with antibody, depends on the concentration of the affinity.
	Clonal Selection	- Proliferate into types of spam messages. - Example 	The process of cloning and proliferating of B-cells and producing two types of cells (i.e. plasma cells and memory cells).
Severity Determination	Danger Theory	- Determines the level of danger for spam messages and three levels of danger are proposed (i.e. low, medium and high).	Process occurs in cell-mediated immunity and it detects a cell that is already in danger and affected by virus.

3.3 CLUSTERING/ CLASSIFICATION USING AIS ALGORITHMS

Since the focus of this thesis is on the classification/clustering, a detail explanation related to AIS algorithms (i.e. Clonal Selection and Immune Network Theory) are presented in order to give further understanding on the theoretical perspective of the used BIS. The defined category of the spam text messages will be determined based on the keywords of the spam. Examples of categories are financial, prize, service, offer, claim, ringtone, chat, date, sex, game/competition, advertisement, entertainment, job, mail messages and voicemail. These categories were obtained from the previous study related to spam messages (Delaney et al., 2012), with further addition of cluster categories resulted from our conduct of research.

a. Immune Network Theory

Niels Jerne (1974) proposed that cells and molecules of the immune system do not only recognise foreign substances but also recognise, respond to and are regulated by each other. It is important to know the concepts of the epitope, paratope and idiotope in order to understand the principle of this theory. Antigen is anything that is able to induce an immune response and it has epitope (or antigenic determinant), a part of an antigen that has complementary with the V region of an antibody specific for that antigen. Paratope is the recognized part in antibody and it is important in the binding process between antigen and antibody. Figure 3.3 shows the binding process of paratope in the antibody with the epitope in the antigen.

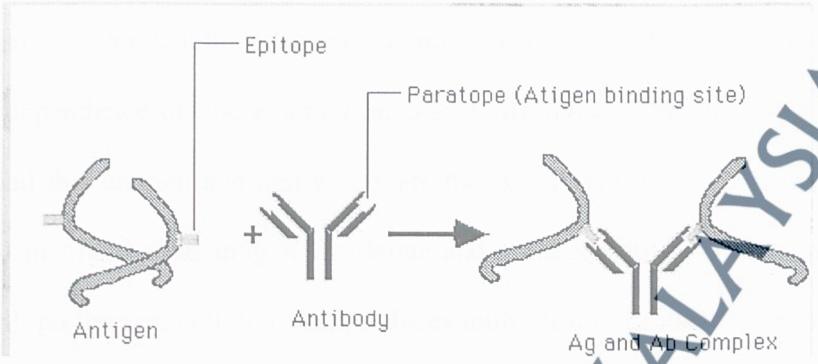


Figure 3. 3: Epitope and paratope

(Immune, 20 March 2014)

The binding of antibody paratope with antigen epitope causes other cells to assist with elimination. Hence, the antibody paratope and antigen epitopes are complementary and analogous to key and locks. Nevertheless, the theory also suggests that antibodies also own a set of the epitope and capable being recognized by other antibodies. Epitopes unique to an antibody type are termed idiotope and the group of antibodies sharing the same idiotope belong to the same idiotype. In this network, two cells are connected if the affinities (i.e. the strength with which an antibody molecule binds an epitope) they shared exceed a certain threshold and the strength of the connection is directly proportional to the affinity they shared (Nanda, 2009).

b. Clonal Selection Theory

One of the factors computational scientists inspired BIS to develop AIS is because of the ability of immune system to remember the previous pathogen and also remember the structure of pathogen. Frank Macfarlane Burnet (1957) introduced the Clonal Selection

Theory that describes the mechanism by which immune cells can generate memory to a specific antigen. He said that the antibody repertoire was produced spontaneously by cells without any dependence on the external antigen. Only those antibodies with sufficient specificity bind that antigen and hence activate the B cell. He then concluded that only those B cells specific for the antigen proliferate and produce antigen-specific antibodies. The B cell will go through cell division; produces multiple clones itself. Some of those B cell clones become memory cells (to remember the types of pathogen when second attacks occur) and also plasma cells (to produce antibodies to fight pathogen). Figure 3.4 shows the basic of the Clonal Selection algorithm.

Figure 3. 4: The basic of process cloning in Clonal Selection.

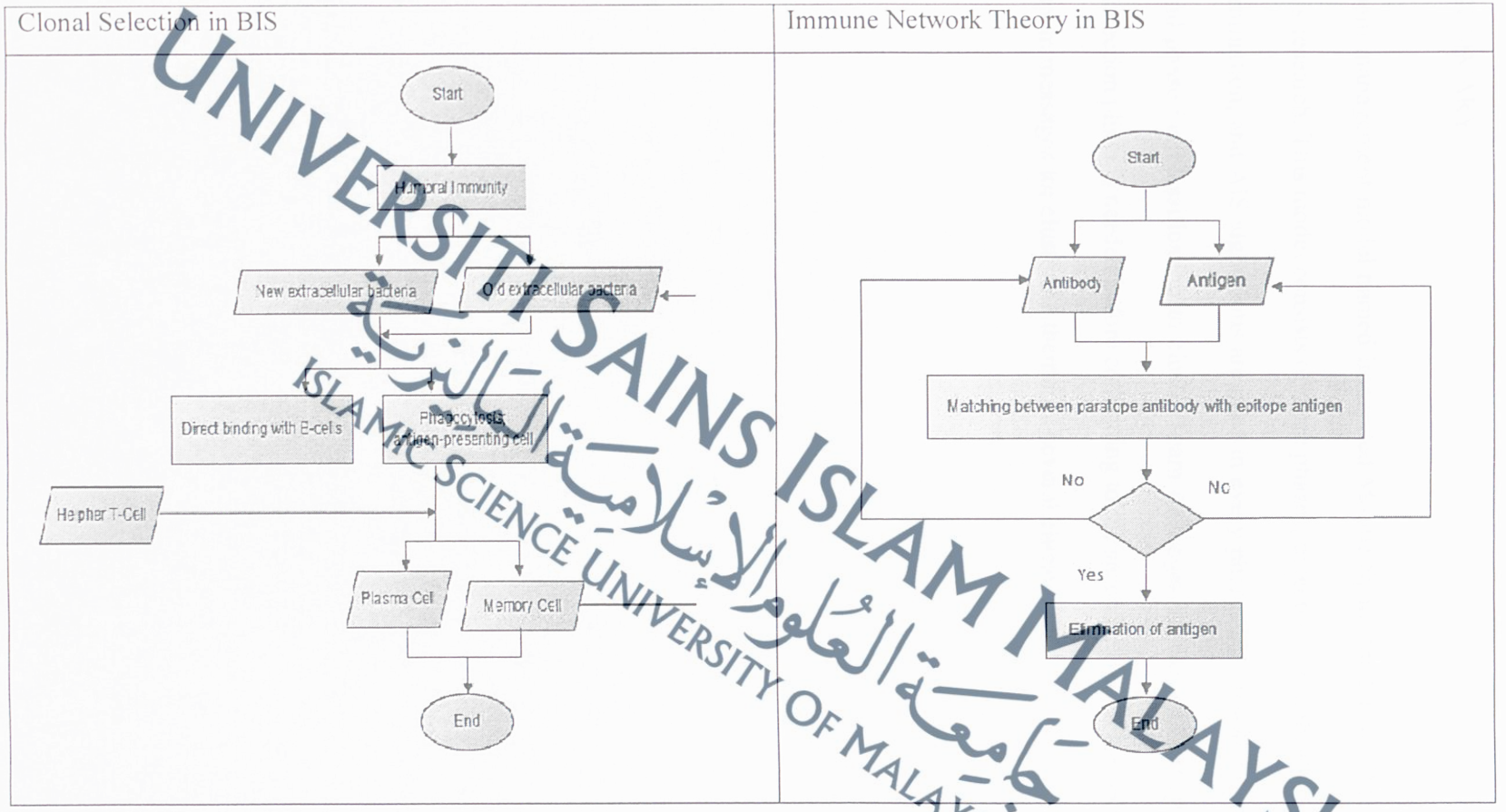
(Nanda, 2009)

In the process of Clonal Selection, the antibody will recognize the antigen which structure fits the pathogen. After it finds suitable antigen, the process of cloning occurs and produces two types of cells which are plasma cell and a memory cell.

The role of Clonal Selection and Immune Network Theory is suitable in clustering the 'spam' text messages into several groups. Ideally in the Clonal Selection, spam messages will be divided into categories based on the contents and keywords of the messages. It is similar with the process that occurs in the immune system where B-cells will clone into two types of cells (i.e. plasma cell and memory cell). We focus in memory cell because it can remember the structure of the pathogen when another attack happens for the second time.

In the Immune Network Theory, spam messages will find the similarity contents with the groups of spam. For example, if the message received contains "Congratulations! You win", it means the keywords "win" match with the category of the prize. In BIS, an antibody will find suitable antigen that matches between paratope and epitope to eliminate the pathogen. Figure 3.5 shows the process that occurs in the Clonal Selection and Immune Network Theory for BIS.

Figure 3. 5: Flow chart of Clonal Selection and Immune Network Theory



3.4 SUMMARY

A spam management model named Integrated Mobile Spam Model (IMSM) is proposed in this research. This model consists of three phases; detection, classification and severity determination, and AIS algorithms are used in every phase. This research focuses on the second phase i.e. classification to cluster spam messages into several groups. However, the detection phase is needed before continuing to the second phase as we need to identify the spam messages for clustering them into several clusters.