

Chapter 3. An Artificial Immune Model for Intrusion Detection

3.1 Introduction

This chapter proposes an artificial immune model for intrusion detection. As seen in the previous chapter, diverse types of artificial immune systems have been applied to various application areas. Among these, intrusion detection is the application area that is the most closely linked to the human immune system. This is because both systems, IDS and the human immune system, aim to protect a host against harmful foreign agents. For this reason, a growing number of proposed artificial immune systems have been applied for intrusion detection.

However, the literature survey of AIS's presented in chapter 2 shows that each AIS is quite different from the others and has its own strengths and problems. The main reason for the absence of a unified artificial immune model is that the human immune system is quite a complicated system to fully implement. Each AIS somewhat simplifies the human immune system by implementing a limited number of human immune features. Hence, in order for an AIS to have selected immune features, it is necessary to comprehend which features of the human immune systems are beneficial to an AIS. This comprehension is possible by assessing whether selected human immune features can allow an AIS to satisfy the requirements of a given application. In addition, in order for an AIS to harness these selected features, it is important to understand how the human immune system is able to obtain these human features. Such understanding can suggest a new artificial immune model equipped with essential immune features that can satisfy all the requirements of a given application.

This chapter focuses on reviewing and assessing the analogy between the human immune system and intrusion detection systems. In order to draw the analogy, a network-based IDS is selected. This is because a network-based IDS can use host-based IDS's as components within a monitoring network, and a network-based IDS monitors multiple hosts in a distributed way as the human immune system does. The chapter starts by identifying a set of general requirements for network-based IDS's and the design goals to satisfy these requirements [Kim and Bentley, 1999a]. Based on these requirements, the salient features of the human immune system that can contribute to the design of competent network-based IDS's are analysed [Kim and Bentley, 1999a]. The analysis leads to a proposal of a new artificial immune model that combines the three evolutionary stages: gene library evolution, negative selection and clonal selection into a single methodology [Kim and Bentley, 1999b]. This chapter then finally presents the scope of the thesis. The thesis scope is defined by clarifying the parts of a proposed artificial immune model that are developed for the thesis and the main research goals.

3.2 Requirements of Network-based IDS's

In chapter 2, network-based IDS's were categorised into three groups according to the overall architecture: *monolithic*, *hierarchical* or *co-operative*. However, each approach shows different problems and no network-based model completely resolves the encountered problems. Before presenting the human immune system features, it is necessary to comprehend which functions are required to design a competent network-based IDS. A careful examination of the literature allows the significant functions to be distilled into seven points:

Robustness: it should *have multiple detection points, which are robust enough against the attack and any system faults on IDS's* [Balasubramaniyan *et al.*, 1998; Forrest *et al.*, 1997]. The critical weak point of an IDS is its failure and subversion by intruders. If intruders already know the existence of an IDS and can subvert it, then the effort to develop the IDS was futile.

Configurability: it should be able to *configure itself easily to the local requirements of each host or each network component* [Balasubramaniyan *et al.*, 1998; Somayaji *et al.*, 1997]. Individual hosts in a network environment are heterogeneous. They may have different security requirements. In addition to hosts, different network components such as routers, filters, DNS, firewalls, or various network services may have various security requirements

Extendibility: it should be *easy to extend the scope of IDS monitoring by and for new hosts easily and simply regardless of operating systems* [Balasubramaniyan *et al.*, 1998; Somayaji *et al.*, 1997]. When a new host is added to an existing network environment and especially when this new host runs a different operating system that has a different format of audit data, it is not simple to monitor it in a consistent manner with existing IDS's.

Scalability: it is necessary to *achieve reliable scalability to gather and analyse the high-volume of audit data correctly from distributed hosts* [Balasubramaniyan *et al.*, 1998]. In the case of the monolithic IDS's, the audit trail collection procedure is distributed and its analysis is centralised [Mykerjee *et al.*, 1994]. However, it is very difficult to forward all audit data to a single IDS for analysis without losing the data. Even if it scales for all audit data correctly, it may cause severe network performance degradation.

Adaptability: it should be *dynamically adjusted in order to detect dynamically changing network intrusions* [Balasubramaniyan *et al.*, 1998; Somayaji *et al.*, 1997]. Computer system environments are not static. Users, vendors and system administrators are constantly changing them. Therefore, the normal activities of networks and intrusions are also continuously changing according to this environment.

Global Analysis: in order to detect network intrusions, it should *collectively monitor multiple events generated on various hosts to integrate sufficient evidence and to identify the correlation between multiple events* [Balasubramanian *et al.*, 1998; Mykerjee *et al.*, 1994]. Many network intrusions often exploit the multiple points of a network. Thus, from a single host, they might appear to be just a normal mistake. However if they are collectively monitored from multiple points, they clearly can be identified as a single attack attempt.

Efficiency: it should be *simple and lightweight enough to impose a low overhead on the monitored host systems and network* [Balasubramanian *et al.*, 1998; Forrest *et al.*, 1997; Somayaji *et al.*, 1997]. A single IDS is expected to perform monitoring, data gathering, data manipulation and decision making. It may impose a large overhead on a system and could place a particularly heavy burden on CPU and I/O, resulting in severe system and network performance degradation.

Even though various approaches have been developed and proposed [Axelsson, 2000] until now, no existing network-based model satisfies these requirements completely [Balasubramanian *et al.*, 1998; Mykerjee *et al.*, 1994].

3.3 The Design Goals of Network-based IDS's

Upon analysis, the requirements identified above can be used to derive three main design goals of an effective network-based IDS. They are being distributed, self-organising and lightweight.

3.3.1 Distributed

The first design goal is being distributed. A distributed network-based IDS delegates its responsibilities to a number of distributed components. A number of independent intrusion detection processes monitor only a small aspect of the overall system. They operate concurrently and co-operate with each other. If a network-based IDS is distributed, it will satisfy the following requirements.

Robustness: for a distributed network-based IDS, the failure of one local intrusion detection process does not cripple an overall IDS even though it causes the minimal degradation of overall detection accuracy.

Configurability: a single intrusion detection process can be simply tailored to local requirements of a specific host without considering the various requirements of other hosts.

Extendibility: even when a new host running a different operating system is added to a network, it is easy to add a new intrusion detection processes on this new host. This is because intrusion

detection processes are independent and thus existing processes do not need to be modified when a new intrusion detection process is added.

Scalability: because audit data collection and its analysis take place in the same place, at a monitored local host, the high volume of audit data is distributed amongst many local hosts. Hence, distributed IDS's are more scalable than IDS's based on a single central server.

3.3.2 Self-organisation

The second goal is being self-organising. Without a central controller having predefined information, a self-organising network-based IDS automatically learns intrusion signatures which are previously unknown and/or distributed. This is achieved through the interaction with changing network environments, various security requirements and other intrusion detection processes. If a network-based IDS is self-organising, it will satisfy the following requirements.

Adaptability: it is highly adaptive because there is no need for manual update of its intrusion signatures as network environments change.

Global analysis: the overall intrusion detection system simply provides the global analysis. This is because it is self-organising from the interactions among a large number of various intrusion detection processes.

3.3.4 Lightweight

The third design goal is being lightweight. A lightweight network-based IDS does not impose a large overhead on a system or place a heavy burden on CPU and I/O. If a network-based IDS is lightweight, it will satisfy the last requirement.

Efficiency: by placing minimal work on each component of the IDS, the main jobs that should be performed by local hosts and networks are not adversely affected by the monitoring.

3.4 Human Immune System features for Network-based IDS's

By performing a careful analysis of the complex capabilities of human immune systems, it is possible to identify several significant features for network-based intrusion detection. Upon investigation, it becomes clear that specific features can act together in order to satisfy each of the three design goals of competent network-based IDS's: being distributed, self-organising and lightweight.

3.4.1 Distributed Model

The human immune system is distributed. The following mechanisms allow the human immune system to detect antigens in a truly distributed way.

Immune Network: the human immune system is implemented through the interactions between a large number of different types of cells. Instead of employing a central co-ordinator, human immune systems sustain the appropriate level of immune responses by maintaining the equilibrium status between antibody suppression and activation using idiotype antibodies [Jerne, 1974; Farmer *et al.*, 1986].

Unique Antibody Sets: the human immune system generates various groups of antibodies to detect different antigens. Its evolution mechanism through natural selection of gene libraries and clonal selection maintains a number of different sets of antibodies. Therefore, each antibody set is unique and independent. These properties do not require any central co-ordinator and they allow the human immune system to detect antigens in a local antibody level [Somayaji *et al.*, 1997].

3.4.2 Self-organisation

The overall immune response is composed of three evolutionary stages: gene library evolution generating effective antibody, negative selection eliminating inappropriate antibodies and clonal selection cloning well-performing antibodies. These three stages are self-organising rather than being directed by a central organ or predefined information.

Gene Library Evolution: antibodies recognise antigens by the complementary properties that only antigens, not self-cells, show. Thus, some knowledge of antigen properties is required to generate competent antibodies. The human immune system learns this knowledge by its evolution over time and hence provides us with efficient and 'knowledge-rich' DNA. Because of this evolutionary self-organisation process, our gene libraries act as archives of information on how to detect commonly observed antigens [Tizard, 1995].

Negative Selection: as the second stage, this eliminates inappropriate and immature antibodies, which bind to self. The important constraint that the immune system has to satisfy is not to attack self cells. Instead of having any global information about self cells, this constraint satisfaction is performed in the thymus and bone marrow by presenting self cells, and removing any antibodies which attack these cells [Forrest *et al.*, 1997; Paul, 1993].

Clonal Selection: as the third stage, this process clones antibodies performing well. In contrast, antibodies performing badly die off after a given life time. Thus, according to currently existing

antigens, only the fittest antibodies survive. Similarly, instead of having the predefined information about specific antigens, it self-organises the fittest antibodies by interacting with the currently existing antigens [Paul, 1993; Tizard, 1995].

3.4.3 Lightweight

The human immune system is lightweight. The following mechanisms allow it to be lightweight and are focused on three ideas: i) how a vast number of antigens can be detected with a smaller number of antibodies, ii) how the known antigen information can be reused efficiently and iii) how numerous antibodies can be generated with a limited number of genes. Approximate binding, memory cells, gene expression and somatic mutation provide the answers to these questions respectively.

Approximate Binding: The immune response activates when the affinity of antibody and antigen binding is above a certain threshold. In other words, a single antibody can detect any number of antigens as long as their affinity is above the threshold. This approximate binding contributes to increase the generality of immune systems [Forrest *et al.*, 1997].

Memory Cells: memory cells store the genetic information of previously detected antigen epitopes and respond efficiently and quickly when they meet the same antigens in the future [Somayaji *et al.*, 1997; Tizard, 1995]. Because memory cells have a longer life span than ordinary antibodies, they retain immunity without the need to create the same antibodies again.

Gene Expression: the immune system maintains antibody diversity in order to ensure the effective detection of a wide range of antigens. In an antibody development process, known as gene expression, several genetic mechanisms are employed to generate diverse antibodies from the gene libraries. The main idea of these mechanisms is that a vast number of new antibodies can be generated from new combinations of gene segments in the gene libraries [Paul, 1993; Tizard, 1995].

Somatic Hypermutation: the immune system learns dynamically changing antigens via clonal selection. During a clonal selection process, cloned antibody secreting cells trigger a *somatic hypermutation* process. Somatic hypermutation mutates the portion of genes that are randomly selected from antibody clones. Mutated offspring of activating antibodies are expected to have wider variations of their antigen matching genes. The key property of these mechanisms is that new antibodies are generated from mutants of useful detectors that have detected recently appearing antigens [Paul, 1993; Sompayrac, 1999].

In summary, this analysis shows that the human immune system is distributed through its immune network and unique antibody sets. It is self-organising because of the three evolutionary processes of gene library evolution, negative selection and clonal selection. It is lightweight because of the generality of approximate binding, gene expression, somatic hypermutation and the efficiency of memory cells.

Since the human immune system is distributed, self-organising and lightweight, it clearly fulfils the design goals for network-based intrusion detection systems. Perhaps most importantly, the mechanisms used by human immune systems satisfy the three goals in an elegant and highly optimised way and this motivates future research harnessing such processes. Because of this study, it is thought that the application of computer immune systems to network-based intrusion detection is likely to provide significant benefits over other approaches.

3.5 Artificial Immune Model for Network Intrusion Detection

3.5.1 Overview

The human immune system has been successful at protecting a human body against a vast variety of foreign pathogens or organisms [Tizard, 1995]. This remarkable property is attractive to computer security researchers and artificial intelligence researchers. Based on the studies by immunologists, a growing number of computer scientists have proposed several different computer immune models [Dasgupta, 1998a]. The main idea of these models is distinguishing self, which is normal, from non-self, which is abnormal. With respect to network intrusion detection, it is viewed that the normal activities of monitored networks as self and their abnormal activities as non-self. Many sophisticated network intrusions such as sweeps, co-ordinated attacks and Internet worms are detected by monitoring the anomalies of network traffic patterns [Porras and Valdes, 98]. Most network-based IDS's monitor network packets and their identified anomalies show critical signatures of these network intrusions [Mykerjee *et al.*, 1994; Porras and Valdes, 1998]. Thus, the artificial immune model is designed for distinguishing normal network activities from abnormal network activities and expected to detect various network intrusions¹.

The overall architecture of the novel artificial immune model is presented in Figure 3.1. The artificial immune model for network intrusion detection consists of a primary IDS and secondary IDS's. For a human body, at the bone marrow and the thymus, various detector cells are continuously generated and distributed to secondary lymph nodes, where the detector cells

¹ Most network-based IDS's operating in real network environments monitor the audit trails generated by a local host together with the network activities. This kind of approach is more reliable at detecting various intrusions. Even though the artificial immune model proposed in this chapter restricts its monitoring scope to network activities, it

reside to monitor living cells. The distributed detector cells monitor all living cells and detect non-self cells invading into secondary lymph nodes. For the artificial immune model, the primary IDS, which we view as the bone marrow and thymus, generates numerous detector sets. The architecture shown in Figure 3.1 is assumed to monitor a single network domain. Therefore, all the input network packets transferred to a monitored single network domain firstly arrive at the first router². Each individual detector set describes abnormal patterns of these network traffic packets. It is unique and transferred to each local host. It is viewed that local hosts act as secondary lymph nodes, detectors as detector cells and network intrusions as non-self cells. At the secondary IDS's, which are local hosts, detectors are background processes which monitor whether non-self network traffic patterns are observed from network traffic patterns profiled at the monitored local host. The primary IDS and each secondary IDS have communicators to allow the transfer of information between each other.

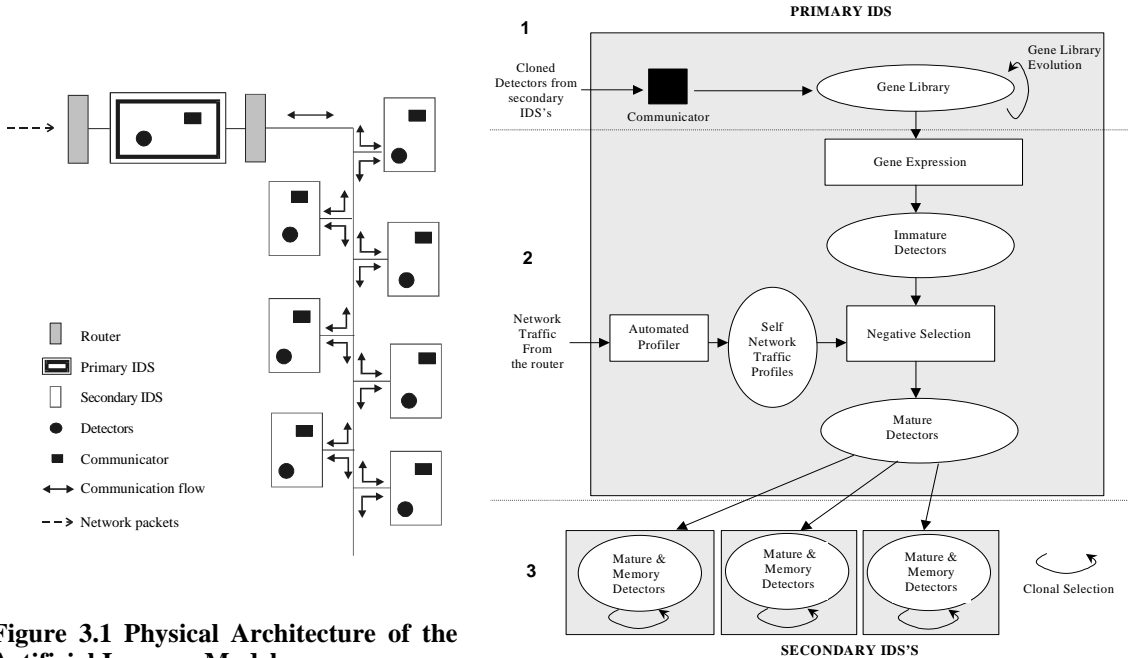


Figure 3.1 Physical Architecture of the Artificial Immune Model

Figure 3.2. Conceptual Architecture of the Artificial Immune Model

Three main goals were identified for designing an effective network-based IDS's: being distributed, self-organising and lightweight. Furthermore, it is shown that several sophisticated mechanisms of the human immune system allow it to satisfy these three goals. For the proposed artificial immune system, these mechanisms are embedded in three evolutionary stages: gene library evolution, negative selection and clonal selection. While the currently existing computer immune models focus on the use of a single significant stage according to their perceived

could be extended by monitoring local audit trails and this extension might be possible by employing one suggestion, a host-based computer immune system, introduced in [Somayaji *et al.*, 1997].

purpose [Dasgupta, 1998a; Forrest *et al.*, 1997; De Castro and Von Zuben, 2000b], the new artificial immune model proposed in this chapter combines these three significant evolutionary stages into a single methodology. The overall conceptual architecture of the proposed artificial immune model is shown in Figure 3.2. In Figure 3.2, stage one indicates gene library evolution, stage two presents negative selection and stage three shows clonal selection. The functions in each stage and how these three stages operate together for performing network intrusion detection are described in the following two sub-sections: Primary IDS and Secondary IDS's.

3.5.2 Primary IDS

The primary IDS performs the first two evolutionary processes: gene library evolution and negative selection. At the gene library evolution stage, it aims to gain general knowledge on detectors that have been effective. At the negative selection stage, it aims to generate a number of diverse detectors, which do not match self, and transfer a number of unique detector sets to distributed local hosts. In order to achieve these tasks, it contains the following components (shown in Figure 3.2).

At the first stage, a gene library is generated and maintained by an evolution process³. The *gene library* of the artificial immune model stores the potentially effective detectors. The potentially effective detectors are the selected ones after they detect anomalous network traffic activity at secondary IDS's. These detectors are transferred from the secondary IDS's to the gene library at the primary IDS. Hence, the gene library collects detectors having useful genes to detect anomalous network traffic patterns and these genes can be rearranged or mutated to generate new detectors. On the other hand, when an offspring detector of detectors in the gene library makes a mistake and detects self-network patterns as an anomaly, the parent detectors in the gene library are also deleted with the offspring detectors. This mechanism drives the artificial immune model to perform *gene library evolution*. This process allows the artificial immune model to learn knowledge of currently existing intrusions regardless of whether they were detected previously or not, making it self-organising by re-using information about previously detected anomalies. Furthermore, its self-organising feature allows it to be lightweight. This is because it does not have to contain all the information of intrusions that have been detected so far. Instead, it holds only the smaller and limited number of detectors that currently survive⁴.

² This assumption can be extended for monitoring large-scale networks which include a number of different domains. It is achieved simply by installing a single primary IDS on each domain and monitoring each domain independently.

³ It should be noted that this evolutionary process is a simulation of the natural evolutionary process for gene libraries. In nature, the DNA (gene libraries) of an organism cannot change within the lifetime of that organism. Evolution operates on populations of organisms, evolving gene libraries based on which organisms survive (i.e., how effective their immune systems are, throughout their lives). This is clearly computationally expensive, so in this model we treat the gene library as a population in itself and evolve it with a single artificial immune system. However, unlike gene library evolution, the other two evolutionary processes within the model operate in a conceptually similar manner to natural immune systems.

⁴ Note that the gene library evolution process described here is different from what was proposed in an earlier model [Kim and Bentley, 1999b]. The earlier model had more emphasis on updating the fitness of individual genes than detectors. For the main goal of this research, that is investigating how the gene library evolution process provides the

At the second stage, the *gene expression* process generates various *immature detectors* by applying various genetic mechanisms to detectors, which are randomly selected from the gene library. These mechanisms can lead to the generation of a vast number of possible immature detectors from combinations of genes [Tizard, 1995]. This process permits the artificial immune model to detect numerous intrusions using a smaller number of detectors, making it lightweight. The *automated profiler* produces a self network traffic profile of raw network traffic packets transferred from the first router. However, the raw network traffic volume is huge and the normal activity patterns are hidden. The automated profiling component reduces the huge volume of raw network packets into a self profile. The fields of the *self network traffic profile* are identical to those of the generated immature detectors. In other words, specific values of these fields can determine whether the observed network activities are normal (the self-profiles), or anomalous (the immature detectors). However, some immature detectors can be false detectors because they have novelty generated via gene expression. These false immature detectors are removed by the *negative selection* process, which matches them to a self network profile produced by an automated profiler. If the match strength between an immature detector and the self profile is over a pre-defined threshold, this new immature detector is considered as a false detector which wrongly identifies self as anomalies, and thus it is eliminated [Forrest *et al.*, 1997]. This process removes false immature detectors by presenting self without any global information about self and hence it shows the property of self-organisation.

Finally, the surviving detectors from negative selection become *mature detectors*. Unique sets of detectors and self network traffic profiles are selected from these mature detectors based on each network connections in order to transfer them to local hosts. This selection guarantees the uniqueness of individual detector sets. These unique detector sets detect network intrusions independently at a local host level [Somayaji *et al.*, 1997] and permit the artificial immune model to be distributed. The selected detector sets and self network traffic profiles are transferred to the second router and it distributes them to their corresponding secondary IDS's.

In order to perform above processes, the primary IDS needs to communicate with the secondary IDS's. For example, the former needs to transfer mature detectors to the latter and the latter needs to send newly found useful detectors to the former. The *communicator* controls any type of communication between the primary IDS and the secondary IDS's.

self-organising feature of the system, the model presented here generalises the gene library evolution process proposed in the earlier model. The major idea of the gene library process for providing the self organising feature is achieved through maintaining information on how to detect commonly observed intrusions (See Section 3.2 Self-organisation). Therefore, the generalised gene library evolution process that keeps useful detectors rather than detector genes still follows the equivalent basic mechanism that was proposed in the earlier model of the gene library process.

3.5.3 Secondary IDS

The secondary IDS's perform the last evolutionary process: *clonal selection*. Its main tasks are detecting various intrusions with a limited number of detector sets and cloning the identical detectors that are performing well, producing memory detectors and driving the gene library evolution in the primary IDS. These tasks are achieved by the operations of several components: self network profiles, unique detector sets, network traffic anomaly detection, clonal selection of detectors, memory detectors and a communicator.

In order to perform *network traffic anomaly detection*, the detectors of *unique detector sets* and *self network profiles* transferred from the primary IDS are compared. First of all, the match strength between a detector and the self profile is measured. When this strength is over a pre-defined threshold, this process informs the communicator. This approximate binding helps make the artificial immune model lightweight. This is because one detector can bind to a number of different intrusions if only their match strength is over the threshold [Somayaji *et al.*, 1997].

After detecting anomalies, the secondary IDS's perform *clonal selection*. When a new detector detects an abnormal network traffic activity, this detector remains as a *memory detector* in a secondary IDS and clones itself. The cloned detectors can be transferred to other hosts. They act as misuse detectors. They quickly detect the same intrusions in the future that have been previously detected. Furthermore, these detectors will be added to the gene library in the primary IDS if they do not exist in the gene library. This drives the gene library evolution in the primary IDS. As the anomaly detection of detectors in local hosts continues, each local host will have more memory detectors and the number of detectors that need to be transferred to each local host will decrease. This process allows the model to be self-organised and lightweight. Instead of having predefined information about specific intrusions, it self-organises the fittest detectors by detecting the currently existing intrusions. In addition, the evolved gene library and memory cells decrease the efforts to create various new detectors, helping to make the model lightweight.

The final decision of whether a network intrusion has occurred is made according to the collective decisions from several local hosts. The artificial immune model employs the agent communication mechanism suggested by Balasubramaniyan *et al.* [1998]. When suspicious activity is detected by anomaly detection process at any secondary IDS, it sends a signal to a *communicator*. The communicator increases the risk level and sends a signal to the communicators in other hosts and the primary IDS. Other communicators, which receive the signal, increase the risk level. If suspicious activities are found from several hosts within a short time, the risk level in each host and the primary IDS will be rapidly increased. When this risk

level becomes above a certain threshold, a communicator can inform the breach of network intrusion to a security officer through a user-interface.

3.5.4 Summary of Artificial Immune Model

The artificial immune model described above consists of the primary IDS and the secondary IDS's. It combines three evolutionary stages. *Gene library evolution* simulates the first stage of evolution, which learns knowledge of currently existing antigens. This process allows the model to be lightweight and self-organising. *Gene expression* and *negative selection* form the second stage of evolution, generating diverse immature detectors and selecting mature detector sets by eliminating false immature detectors in a self-organising way. The transfer of unique detector sets to the secondary IDS's also occurs at this stage, making the model distributed. *Clonal selection* is the third stage of evolution, detecting various intrusions with a limited number of detector sets using approximate binding, and generating memory detectors. This generality and efficiency results in the model being lightweight. In addition, this process drives the gene library evolution in the primary IDS⁵ [Kim and Bentley, 1999a].

3.6 Thesis Scope

In the previous sections, the analogy between the human immune system and intrusion detection systems in a network environment is reviewed. From this review, it is understood that salient features of the human immune system can contribute to the design of a competent IDS and these features are obtained through the coordinated actions of several sophisticated immune mechanisms. This new understanding led to the proposal of a new artificial immune model that can be more beneficial for intrusion detection.

As the next step of this research, the thesis thoroughly examines whether each sub-component of the proposed artificial immune model indeed provides the useful features within an intrusion detection context. In order to verify the effectiveness of the new model, it is necessary to assess whether the proposed system satisfies the requirements of IDS's that are presented in this chapter. However, the research scope of this thesis is focused on investigating whether the integration of separate immune algorithms into one system is effective to detect intrusions. For this research purpose, this study emphasises the analysis of the advantageous features and problems of three evolutionary stages: negative selection, clonal selection and gene library evolution.

These three evolutionary stages are identified as the main processes that allow an AIS to be self-organising. Hence, the research performed in this thesis is concentrated on whether the

⁵ Discussion of the proposed artificial immune model is presented in Appendix B. From this discussion, the new artificial immune model is analysed with respect to the requirements of a network-based anomaly detector identified in section 3.2 Requirements of Network-based IDS's.

developed system shows adaptability to continuously changing self and non-self environments. The AIS developed for the thesis has immune components that can result in the other two important features, being distributed and lightweight. However, the extensive tests that study whether the developed system provides these two features are not performed within the scope of this thesis. Instead, the research reported in the remaining chapters emphasises how the three integrated evolutionary stages can permit an AIS to harness self-organising features. For this research goal, the AIS developed here is tested with real network traffic data and other machine learning benchmarking data sets. In addition, all the tests reported in this thesis are performed in a single host under a simulated dynamic environment. New understanding through the research performed for this thesis corroborates the promising advantages of a newly proposed artificial immune model for intrusion detection. This corroboration is discussed in Chapter 8: Conclusion.

3.7 Summary

This chapter reviews and assesses the analogy between the human immune system and intrusion detection systems. In order to draw the analogy, a network-based IDS is selected. A set of general requirements for network-based IDS's and the design goals to satisfy these requirements are identified. Based on these requirements, the salient features of the human immune system that can contribute to the design of competent network-based IDS's are analysed. The analysis shows that the coordinated actions of several sophisticated mechanisms of the human immune system satisfy all the identified design goals. These results lead to a proposal of a new artificial immune model that combines the three evolutionary stages: gene library evolution, negative selection and clonal selection into a single methodology. This chapter then finally presents the scope of this thesis by clarifying the parts of a proposed artificial immune model that are developed for this thesis and the main research goals.