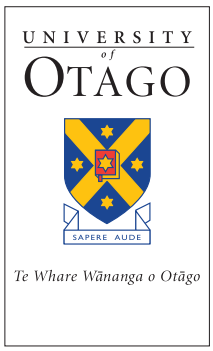

Positive and Negative Selection in a Multilayer Artificial Immune System

Melanie Middlemiss

The Information Science Discussion Paper Series

Number 2006/03
January 2006
ISSN 1177-455X

University of Otago

Department of Information Science

The Department of Information Science is one of seven departments that make up the School of Business at the University of Otago. The department offers courses of study leading to a major in Information Science within the BCom, BA and BSc degrees. In addition to undergraduate teaching, the department is also strongly involved in postgraduate research programmes leading to MCom, MA, MSc and PhD degrees. Research projects in spatial information processing, connectionist-based information systems, software engineering and software development, information engineering and database, software metrics, distributed information systems, multimedia information systems and information systems security are particularly well supported.

The views expressed in this paper are not necessarily those of the department as a whole. The accuracy of the information presented in this paper is the sole responsibility of the authors.

Correspondence

This paper represents work to date and may not necessarily form the basis for the authors' final conclusions relating to this topic. It is likely, however, that the paper will appear in some form in a journal or in conference proceedings in the near future. The authors would be pleased to receive correspondence in connection with any of the issues raised in this paper, or for subsequent publication details. Please write directly to the authors at the address provided below. (Details of final journal/conference publication venues for these papers are also provided on the Department's publications web pages: http://www.otago.ac.nz/informationscience/pubs/). Any other correspondence concerning the Series should be sent to the DPS Coordinator.

Department of Information Science
University of Otago
P O Box 56
Dunedin
NEW ZEALAND

Fax: +64 3 479 8311
email: dps@infoscience.otago.ac.nz
www: http://www.otago.ac.nz/informationscience/

Positive and negative selection in a Multilayer Artificial Immune System

Melanie Middlemiss, *Student Member, IEEE*

***Abstract*— The immune system is a complex and distributed system. It provides a multilayered form of defence, capable of identifying and responding to harmful pathogens that it does not recognise as "self". The framework proposed in this paper incorporates a number of immunological concepts and principles, including the multilayered defence and the cooperation between cells in the adaptive immune system. An alternative model of positive selection is also presented. It is suggested that the framework discussed here could lead to reduced false positive responses in anomaly detection tasks, such as intrusion detection, as well being extended to a population of computational immune systems that are able to maintain population diversity of recognition and response.**

I. Introduction

Parallels have been drawn between the human immune system and anomaly detection problem domains, particularly with regards to intrusion detection systems (IDS) [1], [2], [3], [4]. The human immune system, for the most part, successfully protects the body from harmful pathogens that come in many varied forms. Each type of pathogen has a different cellular structure, method of replication and mechanism for entering the body. The immune system has evolved complex structures and methods for identifying these pathogens and removing or responding to the threat that they possess. The widely held view in immunology is that the main function of the immune system is to distinguish between "self" (cells belonging to the individual) and "non-self" (pathogens) [5]. An alternative view is that of the "Danger theory" proposed by Matzinger [6], [7] and discussed in relation to artificial immune systems by Aickelin [8], [9]. However the self/non-self theory still stands as the foundation view of immune system function.

These parallels have lead to much development and research in the area of artificial immune systems (AIS). While the human immune system is complex in its methods and systems, there are only a select few immunological principles that are modelled in AIS. A recent report by Dasgupta and Yu [10] cites 456 journal articles, conference papers and technical reports in the field of AIS. Despite this volume of work there are only a small number of immunological principles that have been used in these models [11], [12]. If we take a broader look at the immune system there are several features that could provide a novel approach to developing an artificial immune system for anomaly detection. In particular the multilayered approach to detection and response afforded by the human immune system is one in which has been previously modelled by Dasgupta [13], but could be explored further. The negative selection algorithm initially presented by Forrest [1] has been well researched, but the additional process of positive selection which immature lymphocytes are subject to in the adaptive immune system is not as well researched.

In this paper a framework is presented that incorporates principles from both innate immunity (the "built-in" defence of the immune system) and adaptive immunity (the memory and specific defence abilities). A multilayered approach is taken and a method for positive selection is presented which is modelled on an alternative immunological theory [14]. The focus is on the development of a framework that is able to adapt and evolve a diverse population of detectors which can recognise and respond to known anomalous behaviour (known as misuse in IDS), as well as previously unknown anomalous behaviour. The framework provided in this paper is expected to be applied in the area of host-based computer intrusion detection, where each computer on a network would have an individual computational immune system. It is the intention that this framework will be extended to allow the population (network) of individuals (computers) to work together to maintain diversity of recognition and response within the population. Methods of communication and "vaccination" techniques to implement this are left for future research and not discussed here. For the remainder of this paper, the discussion refers to the development of a computational immune system for a single individual within the population.

The remainder of the paper is structured as follows. Section 2 contains an overview of some immune system concepts involved with the framework presented in this paper. Section 3 discusses two theories of positive selection in the immune system and their implication for artificial immune systems. Section 4 then introduces the framework and the components are discussed in Section 5. Section 6 gives a description of the three main phases involved in the framework. The paper concludes in Section 7 with a discussion of future work.

II. Immune Systems Background

Humans and other vertebrates have a complex and sophisticated immune system that involves several layers of defence. The first and most obvious level is that of the body's external defences. This includes the tissues that cover and line the body, and the various physiological conditions that are present in the body. For example, unbroken skin provides a barrier from invading pathogens. Likewise, the mucous membranes lining the respiratory tract stop pathogens from

Melanie Middlemiss is a PhD student with the Department of Information Science, University of Otago, Dunedin, New Zealand, (phone: +64 3 479 8315; email: mmiddlemiss@infoscience.otago.ac.nz).

entering the respiratory system [5]. The remainder of this section gives an overview of some of the concepts and principles involved in the second and third layers of this defence system: namely innate and adaptive immunity.

A. Innate immune system

The innate immune system consists of specialised cells and molecules that are responsible for the initial response to any pathogen which has entered the body. This is achieved through pattern recognition receptors (PRRs) on the surface of the innate immune system cells. PRRs recognise pathogen associated molecular patterns (PAMPs) which are found on the cell surface of invading pathogens and never on the host [15]. These molecular patterns have been conserved through evolution and are common to many pathogens. A single cell can have multiple different PRRs present on its cell surface. Each receptor is specialised towards the recognition of a specific type of pathogen. For example some of the types of PRRs known to be found on macrophages (a type of innate immune system cell) include Mannose receptors, Scavenger receptors, Tol-like receptors and CD14 receptors. The CD14 receptor for example, binds to a particular molecule (*lipopolysaccharide–LPS*) which is found only in the cell wall of Gram-negative bacteria, e.g. *E.coli, Neisseria, Salmonella* [16].

The pattern recognition receptors found on innate immune system cells are germline encoded. Consequently the responses of the innate immune system are "built-in" from birth and do not change with age or experience of infection [5]. The ability of the PRRs to recognise a broad range of pathogens provides a non-specific defence for the immune system and through signalling and communication mechanisms, contributes to the induction of appropriate adaptive immune system responses.

B. Adaptive immune system

As opposed to the innate immune system which is non-specific in its defence against harmful pathogens, the adaptive immune system initiates a response specific to the pathogen that has entered the body. The adaptive immune system also provides memory capabilities to the immune system. This functionality is provided through a particular type of cell found only in the adaptive immune system - the lymphocyte. Table I describes the four main properties of lympohocytes that make them distinct from other innate immune system cells [5, pg 118]. These differences combine to maintain the flexibility of the immune system.

Lymphocytes circulate through the blood and lymphatic system waiting to encounter antigens (the foreign molecules belonging to pathogens that invade the body). Each antigen has a particular shape that is recognised by the receptor present on the lymphocyte cell surface. The ability of the immune system to recognise and respond to the millions of different antigens that it encounters comes from the large lymphocyte receptor repertoire. Millions of lymphocyte cells are developed and selected in a way that the receptors on each cell surface recognise different antigens, ensuring diversity within the immune system.

TABLE I
FOUR MAIN DIFFERENCES BETWEEN LYMPHOCYTES AND INNATE IMMUNE CELLS, SUCH AS MACROPHAGES [5, PG 118].

1) They *recirculate* through the blood, tissues, and lymphoid organs, waiting to encounter foreign molecules (or antigens).
2) They are individually *specific* for the antigens they recognise. This is due to the possession of antigen-specific surface receptors.
3) When they recognise 'their' antigen, they *respond* by proliferating and switching on a particular function (e.g. cytotoxicity; secretion of antibody or cytokines).
4) Once they have functioned, some of them remain for years as *memory* cells, with the capacity for faster and larger future responses.

The ability of the adaptive immune system to identify new and previously unseen pathogens through recognition of antigen by the lymphocytes is one of the underlying principles behind the development of artificial immune systems [17]. In the adaptive immune system there are two types of lymphocytes: B cells and T cells. These lymphocytes begin as a stem cell found in the bone marrow and go through a maturation process. B cells complete their maturation in the bone marrow, while T cells move to the thymus to complete maturation. Much of the research in AIS that models lymphocyte development and behaviour incorporates the main properties of B and T cells into the concept of a generic detector.

One of the important processes that lymphocytes undergo is that of central tolerance. This involves the elimination of lymphocytes that recognise self antigens and would otherwise initiate an immune response to self. This provides immunological tolerance towards self and has been widely modelled in artificial immune systems– implemented as the negative selection algorithm [1]. It is suggested here that, by ignoring the differences between B and T cells, several potentially useful properties for developing an AIS are ignored. For this reason we investigate the differences and in particular a way to model the additional process of positive selection that T cells undergo. The immunological processes underlying positive selection of T cells is discussed in the following section.

III. POSITIVE SELECTION

The T cell receptor (TCR) is similar to that of the B cell, except that it is not secreted as antibody and does not bind directly to antigens. Rather it binds to peptides (small fragments of protein broken down from the pathogen) presented in a complex with a specialised self molecule called the Major Histocompatibility Complex (MHC). Recognition of the self-MHC:peptide complex means that T cells undergo a slightly different process to B cells. While T cells are selected

for their lack of recognition of self (negative selection as in B cells), they must also be able to recognise the self MHC molecule.

MHC is a specialised complex involved in regulating the T cell response in the immune system [16]. MHC is highly polymorphic, with an estimated 10^{13} possible combinations [15]. Humans inherit genes for about six different MHC molecules, promoting diversity within the population. The widely held view of the reason behind this difference between the MHC of individuals is that this ensures the immune systems of all individuals do not react in the same way to pathogens. It could be envisaged that if all members of a population respond in the same way, a new pathogen would be able to destroy the entire population [5].

Positive selection is an area in immunology where there are contrasting views. This arises around the ability of positive and negative selection to work together to both *retain* cells that recognise the self-MHC:peptide complex, while also *removing* cells that recognise any self peptides [18]. There are two theories regarding this issue: first that the interactions required for positive and negative selection are similar and receptor affinities drive the two processes (avidity hypothesis); and second that the interactions required are different, with different signals inducing different responses for positive and negative selection (differential signalling hypothesis) [15], [14].

A. Avidity hypothesis

According to the avidity hypothesis (AvH) the avidity[1] with which the TCR binds to the self-MHC:peptide complex will determine whether the cell is positively or negatively selected. T cells are transported from the bone marrow to the thymus where, during positive selection, those cells that demonstrate a relatively weak avidity for the self-MHC:peptide complex presented by epithelial cells are induced to die. This results in the removal of cells that are unable to recognise the self-MHC molecule. During negative selection, T cells with a strong avidity for the self-MHC:peptide complex presented by dendritic cells are also induced to die. The result of this positive and negative selection is a repertoire of T cells with receptors that can be considered to have a “medium” avidity for the self-MHC:peptide complexes presented to them.

Several studies in the area of AIS have used the concept of positive selection to model the generation of detectors [2], [22]. The positive selection algorithm used in these papers is essentially the inverse of the negative selection algorithm described by Forrest [1]. Instead of removing detectors that match self strings, detectors that do not match self strings above a given affinity threshold are removed. This results in a set of detectors that describe the self space rather than the nonself space as with the negative selection algorithm [19]. When combined with the negative selection algorithm in a manner that models the avidity hypothesis, affinity thresholds are used to restrict which detectors should and should not be removed. The result from such a process is a set of detectors whose affinity is not too high when matched to self (negative selection) and also not too low (positive selection).

Aicklen and Cayzer [8] compare positive and negative selection and show that the effectiveness of their use depends on the nature of the problem domain. Using positive selection, in which the detector set represents the self space, a new string that was non-self would have to be compared with all detectors to confirm it was in fact non-self. However, if the new string was self it would only have to match one detector in order to confirm it was a self string. In the case of negative selection this is reversed and a non-self string would only have to match one of the detectors that covers the non-self space. It follows that with a new self string, using negative selection a comparison would have to be made with all detectors to confirm the string was self. This implies that the choice of positive or negative selection using these methods, is dependent on the self/non-self distribution of the problem domain.

B. Differential signalling hypothesis

The alternative view of the differential signalling hypothesis (DSH) assumes that the interactions required for positive and negative selection are different. Cohn discusses this theory and asserts that “restrictive recognition of peptide (*P*) by the TCR requires both allele-specific recognition of the MHC-encoded restricting element (*R*) and specific recognition of the peptide bound to it” [14, pg. 375]. This suggests that there must be two different structures which combine to make up the T cell receptor, as opposed to the single structure implied by the avidity hypothesis. Figure 1 gives a simplified view of the components of such a T cell receptor according to the differential signalling hypothesis. The two structures, or paratopes, of the receptor are anti-*R*, which is germline encoded and anti-*P*, which is somatically encoded.

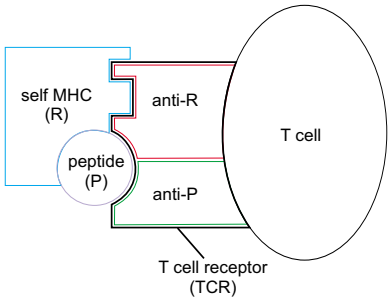

Fig. 1. Simplified view of TCR according to the differential signalling hypothesis.

The differences between these two hypotheses for positive selection and their implications for TCR structure is described by Cohn [21] as a ‘singly recognitive multiple receptor’ in the avidity hypothesis and as a ‘multiple recognitive single receptor’ in the differential signalling hypothesis. This means that the TCR in the AvH recognises a single

[1]the sum total of the strength of binding of two molecules or cells to one another at multiple sites - distinct from affinity [15, pg. 685]

paratope– the MHC:peptide complex (*singly recognitive*) and the avidity of this match determines what type of signal is sent to the T cell, resulting in different outcomes– positive or negative selection (*multiple receptor*). Alternatively in the DSH the TCR is thought to be comprised of two different paratopes which recognise *R* and *P* respectively (*multiple recognitive*). Where the TCR could send different signals in the AvH, here the TCR can only deliver a single signal and the stage of development that the T cell is in will determine how that signal is interpreted (*single receptor*).

As far as the author is aware, the differential signalling process has not been used previously in artificial immune systems.

IV. Communication Between Cells in the immune System

One of the key functions of the immune system is communication and there are many aspects to this communication. The majority of cells in the immune system act under the influence of signals that they receive from the cells [5]. These signals direct the function of the cells and can be delivered either: 1) through direct contact between cells, or 2) through specialised molecules called cytokines.

While the innate and adaptive components of the human immune system have quite distinct functions, there is a lot of communication and interaction that occurs between the cells of each component. Cytokines are responsible for most of this type of communication. Communication, in the form of costimulatory signals and cell-cell interaction also occurs within the components of the immune system. For example, in the adaptive immune system these communication mechanisms determine how the cells will respond.

In the adaptive immune system there are two ways in which a response can be initiated: 1) antibody-mediated or humoral response, or 2) cell-mediated response. The humoral response is driven by direct B cell interaction with antigen and defends against extracellular pathogens. The cell-mediated response is driven by indirect activation of T cells through recognition of the MHC:peptide complex and defends against intracellular pathogens. There are many points in these response pathways at which communication and signalling occurs.

This communication and signalling helps in achieving peripheral tolerance in the immune system. It is possible that some cells which are able to survive the central tolerance processes could then incorrectly respond to self. A mechanism to stop this harmful process is required, and this is achieved by the requirement of a signal from another cell [16]. For example T cells require costimulatory cells from antigen presenting cells before a response is initiated. “By requiring that helper T cells only recognise presented antigen, Mother Nature guarantees that the decision to deploy the deadly adaptive immune system is not made by a single cell” [18, pg. 51].

In anomaly detection problem domains a common issue is that normal behaviour is often classified incorrectly and the system continuously raises alarms. The cooperation and costimulation between cells in the immune system ensures that an immune response is not initiated unnecessarily, thus providing some regulation to that response. Implementing an error-checking process provided by cooperation between layers or types of detectors could reduce the level of false positive alerts in an anomaly detection system.

The relevance of internal communication mechanisms, such as costimulation has been acknowledged in the area of artificial immune systems. One method that has been described to achieve this is through the use of human interaction to provide the costimulation signal [20]. If a detector raised an alarm, the user has a certain length of time to decide if the alarm was a true alert. If the detector does not receive this costimulation it is destroyed. Otherwise it becomes a memory detector with an indefinite lifetime. While this models costimulation, it still requires user intervention and could interfere with normal user activities if the user is prompted too often to stimulate the system.

In [24] costimulation is achieved in an AIS for intrusion detection within a network of computers by a check to determine whether other systems in the network have also raised an alarm. This costimulation is intended to reduce the number of false alarms, but it may lead to an increase in the number of missed attacks, as all detected packets must be costimulated.

The following section describes how the immune concepts presented in the previous sections could be modelled in an AIS.

V. Framework for Multilayer Artificial Immune SYstem

The aim of the framework presented in this section is to develop a multilayered detection system that can be used to create a population of individuals which are able to identify a variety of known and unknown anomalous behaviour. This multilayered approach is designed to provide a form of internal costimulation, similar to peripheral tolerance, and reduce the number of false positive responses from the system.

The framework consists of several components which comprise a multilayered system of detection. These components (MHC, Antigen Presenting Detectors, T and B detectors) are described in the following section. Sections VII–VIII then outline the three phases of the system: initialisation, incremental learning and adaptation, and identification.

VI. Representation

In this framework ‘self’ is defined as background knowledge of normal data. In an IDS application, for example, this would be normal system behaviour modelled by a number of features or attributes relating to the type of behaviour of interest. ‘Nonself’ is therefore any data that is not self, or abnormal. In the IDS application, nonself would be anomalous behaviour. Knowledge of existing anomalous behaviour (misuse or attack information) is also used in the framework, incorporating the concept of a ‘built-in” knowledge of nonself that the innate immune system provides.

A. MHC

As discussed in Section III, the Major Histocompatibility Complex (MHC) promotes diversity of recognition and response within a population of individuals. One method of modelling MHC is presented by Hofmeyr [23]. A permutation mask is created for each detector set, which when applied causes identical detectors to match data examples in different ways. This creates diverse recognition abilities within a population that contain similar detector sets.

An alternative method of modelling MHC is presented by Kim [25] whereby MHC is modelled as a set of rules extracted from a 'self' knowledge base using the Apriori rule mining algorithm. An immature T detector is then generated in the form of an "If-Then" rule that contradicts the rules in the MHC set. These rules are filtered using a rule confidence threshold in a process analogous to the avidity hypothesis for positive selection.

In this work as we are attempting to model the multilayered detection of the immune system the MHC is represented as a set of binary string feature masks which are highly specific to known types of anomalous behaviour. MHC is involved in the positive selection process of T cell development. When antigen presenting cells break down pathogens the MHC presents the peptide fragments on the antigen-presenting cell surface. This is analogous to a feature selection process where the MHC presents feature subsets of the pathogen that it knows T cells will be able to recognise. As there are a large but finite population of MHC in the human immune system and each individual has a small subset of these, a global MHC set is generated for the population and a local MHC set is extracted from this for each individual.

The global set of MHC are formed through feature extraction from a base of background knowledge. A method that could be used to perform this feature extraction is information gain (IG). Features would be ranked and those features with an IG ratio above a threshold are used to create a set of feature masks. The threshold is selected experimentally and is used instead of a feature count cut-off to allow the creation of different length feature masks.

B. Antigen Presenting Detectors

Foreign pathogens are presented to T cells by Antigen Presenting Cells (APC). There are different types of cells that can act as APCs, including cells from both the innate immune system (for example macrophages) and adaptive immune system (for example B cells). These APCs identify pathogens through their receptors and break down the pathogen into peptides which are presented in a complex with MHC. This is analogous to a set of detectors that are developed to recognise conserved patterns in anomalous behaviour.

In this framework antigen presenting detectors (APD) are modelled by rules mined from existing background knowledge of 'normal' and 'anomalous' behaviour. The rules relating to anomalous behaviour are extracted and added to a local APD rule list for an individual. In a population of systems, each individual would have a localised domain of background knowledge and would develop a personalised list of APD rules specific to behaviour they have encountered in the past.

From this local APD rule set a population of APDs is constructed for the local system. An APD consists of the random pairing of an APD rule and a MHC mask. This is analogous to the APC, which recognises a particular pathogen and the MHC, which presents a particular peptide fragment to T cells.

As with the APCs in the innate immune system which these detectors are modelled on, this knowledge is "built-in" from initial generation. However, unlike in the innate immune system, where this base of knowledge does not change with age or experience of infection, in the computational immune system it is desirable for the set of APDs evolve over time. This evolution would involve the expansion of the APD set to incorporate new knowledge of anomalous behaviour– either through the addition by the user of preformed rules, or through communication with other computational immune systems in the population.

C. T detectors

The T cells of the adaptive immune system are used in the framework to develop a set of detectors that can regulate the systems recognition and response. In the immune system central tolerance is achieved through negative and positive selection. Peripheral tolerance is also achieved, ensuring the immune system components do not respond against self, by the requirement of T cells for a costimulatory signal from an APC.

In this framework the differential signalling hypothesis is modelled and as such we take the view that the T detector is comprised of two parts: a rule (the T cell recognition receptor) and an MHC feature mask (recognising the MHC that is presented by the antigen presenting cell). The MHC mask is randomly generated and compared with the masks in the local MHC set. If the average affinity between the T detector MHC mask and the MHC set masks is above a given threshold (Thr_{PS}) the T detector is positively selected and the development of the detector continues. If the T detector MHC mask does not match above the Thr_{PS}, the T detector is removed and a new T detector is generated. The value for Thr_{PS} is selected experimentally.

After the positive selection process, a random rule is generated for the new T detector. This rule is masked by the T detectors MHC mask, and is subjected to a negative selection process. Local background knowledge of normal behaviour is used to match a randomly generated rule. If the affinity, or strength, of the match between the rule and any data example in this 'normal' data is above a given affinity threshold (Thr_{NS}) it is removed and another T detector is generated. The value for Thr_{NS} is selected experimentally. A modified Euclidean distance measure is used to determine affinity, where the affinity is high if the Euclidean distance is low and conversely the affinity is low if the Euclidean distance is high.

When a T detector passes both the positive and negative selection processes it is added to the local T detector set.

D. B detectors

The final component in the framework is a set of B detectors analogous to B cells in the adaptive immune system. These are similar to generic detectors modelled in other AIS. They undergo the same process of negative selection as T detectors, but are not subjected to positive selection against the local MHC set. This leads to a set of detectors that are selected for their difference to known normal behaviour, but they are not restricted by the feature masks of the MHC. Consequently the B detector set adds a more general form of detectors for anomaly detection to the framework.

VII. Initialisation Phase

There are three main phases described in the framework. The first of these is the initialisation phase. This involves the generation of each of the four sets described above for a local system. The initial number in each set is arbitrary. The number in the local MHC set will be proportionally small according to the number of MHC masks in the global MHC mask list. This global list is generated for each subtype in the anomalous behaviour space. For this reason this framework is more suited to complex problem domains, such as intrusion detection. In intrusion detection there is normal behaviour and behaviour that constitutes anomalous behaviour. However anomalous behaviour can be subdivided into different types of attacks. For example SYN flood, Ping of Death, Perl, Imap, or Ipsweep attacks [26]. Therefore the global MHC list will contain a feature mask for each of these subtypes and the number depends on how many are present in the global background knowledge. The local MHC set will have a small random subset of these global MHC.

The size of the APD set depends on the number of rules in the APD rule list. The number of APDs should be selected so that the probability of duplication amongst the APD set is minimised. As for T and B detectors, the size is set according to the computational cost involved with generating the detectors. All sets are continuously adapted during the lifecycle of the system and therefore the detector set sizes could be expanded to improve detection if necessary.

VIII. Identification Phase

The objective of the framework is to develop a system that is able to detect anomalous behaviour. This is achieved through the identification phase. A data example presented to the system is subjected to the multilayered identification process of the system. This data example is first matched to the rules of the local APD set. This is analogous to the innate APCs in the human immune system recognising conserved patterns in pathogens.

For each APD that matches the data example, the MHC mask associated with the APD is matched to the MHC masks in the local T detector set. This step models the way in which antigen presenting cells present the MHC:peptide complex and T cells recognise first the MHC part of the complex. It also provides some regulation to the system, in that confirmation of the anomalous nature of the data example is required from the T detectors. This is similar to the way in which T cells require the costimulatory signal from APCs to achieve peripheral tolerance.

If there is a match between the APD MHC mask and a T detector MHC mask, the data example and T detector rules are filtered using the APD MHC mask. The affinity of the match between the data example and the detector rule is then measured and if it is above a given threshold Thr_{AFF} the data example is identified as anomalous.

At each point if there is no match, the data example is matched to the rules in the B detector set. This allows a general detection by cells that have developed according to their dissimilarity to normal behaviour.

IX. Incremental Learning and Adaptation Phase

As in the immune system, the initial detector sets are developed so that they should be able to identify anomalous behaviour. However the immune system is adaptive and allows the cells to die, new cells to develop, and cells that are good at detecting harmful pathogens proliferate and are mutated to become even better detectors. This is modelled in the incremental learning and adaptation phase of this framework.

After initialisation, a further traning set of labelled data is passed through the framework for identification (as described above). For each T detector involved in a true positive detection (anomalous data example correctly classified as anomalous), the associated APDs and data examples are found. Each T detector rule and mask are cloned several times. Each clone is mutated and the fittest clone is found and replaces the original T detector. A fitter clone is one that can perform better across all data examples. The objective of this process is to produce a clone that is able to identify more similar anomalous data examples in the future.

During this phase underperforming detectors are removed and replaced with new ones. Any detectors that have not been involved in a detection in time period t are removed and replaced with a new detector. This is an ongoing adaptation process to ensure the system does not get overwhelmed by detectors that are unable to assist in the identification of anaomalous behaviour.

X. Conclusions

This paper has presented an initial overview of a framework for a multilevel artificial immune system. This framework has combined serval immune principles, including using positive and negative selection in the generation of detector sets. The motivating idea behind this framework is to reduce the false positive rates in anomaly detection problem domains, for example in an intrusion detection system. The focus now is to implement and test this framework using a benchmark intrusion detection dataset.

REFERENCES

[1] Forrest, S., Perelson, A. S., Allen, L. and Cherukuri, R., "Self-Nonself Discrimination in a Computer," *Proc. IEEE Symposium on Research in Security and Privacy*, IEEE Computer Society Press, 1994, pp. 202–212.

[2] Forrest, S., Hofmeyr, S. and Somayaji, A., "Computer Immunology," *Communications of the ACM*, vol. 40, no. 10, 1997, pp. 88–96.

[3] Kim, J. and Bentley, P., "Negative selection and niching by an artificial immune system for network intrusion detection," *Proc. Late Breaking Papers at the 1999 Genetic and Evolutionary Computation Conference*, Edited by Brave, S. and Wu, A. S., Orlando, Florida, USA, 1999, pp. 149–158.

[4] Kephart, J. O., "A Biologically Inspired Immune System for Computers," *Artificial Life IV: Proceedings of the Fourth International Workshop on the Synthesis and Simulation of Living Systems*, Cambridge, MA, US: MIT Press, 1994, pp. 130–139.

[5] Playfair, J. H. L. and Bancroft G. J., *Infection and immunity*, Oxford; New York: Oxford University Press, 2004.

[6] Matzinger, P., "Tolerance, danger, and the extended family," *Annual Reviews of Immunology*, vol. 12, 1994, pp. 991–1045.

[7] Matzinger, P., "The danger model: A renewed sense of self," *Science*, vol. 296, no. 5566, 2002, pp. 301.

[8] Aickelin, U. and Cayzer, S., "The Danger Theory and Its Application to Artificial Immune Systems," *Proc. 1st International Conference on ARtificial Immune Systems (ICARIS-2002)*, Edited by Timmis, J. and Bentley, Peter, University of Kent at Canterbury, UK, 2002, pp. 141–148.

[9] Aickelin, U., Bentley, P., Cayzer, S., Kim, J. and McLeod, J., "Danger Theory: The Link between AIS and IDS?," *Hewlett Packard Labs: HPL-2003-138*, 16 July, 2003.

[10] Dasgupta, D. and Yu, S., "Artificial Immune Systems: A Bibliography," *CS Technical Report: CS-03-002*, Computer Science Division, The University of Memphis, USA, Dec, 2003.

[11] de Castro, L. N. and Timmis, J., *Artificial Immune Systems: A New Combutational Intelligence Approach*, Springer-Verlag, 2002.

[12] Dasgupta, D. and Ji, Z. and Gonzalez, F., *Artificial Immune System (AIS) Research in the Last Five Years*, *Proc. Congress on Evolutionary Computation*, Edited by Sarker, R. et. al., Canberra, Australia: IEEE Press, 8–12 Dec 2003, pp. 123–130.

[13] Dasgupta, D., Yu, S. and Majumdar, N. S., "MILA - Multilevel Immune Learning Algorithm," *Proc. Genetic and Evolutionary Computation Conference*, Edited by. Cantu-Paz, E. et. al., Chicago, USA: Springer-Verlag, 2003, pp. 183–194.

[14] Cohn, M., "An alternative to current thinking about positive selection, negative selection and activation of T cells," *Immunology*, Blackwell Publishing, vol. 111, 2004, pp. 375–380.

[15] Janeway, C. and Travers, P., *Immunobiology : the immune system in health and disease*, Current Biology; London, San Francisco New York: Garland Publishers, 5th Edition, 2001.

[16] Lydyard, P. M., Whelan, A. and Fanger, M. W., *Immunology*, London: Bios Scientific, 2nd Edition, Instant Notes Series, 2004.

[17] Somayaji, A., Hofmeyr, S. and Forrest, S., "Principles of a Computer Immune System," *Proc. New Security Paradigms Workshop*, Great Langdale, Cumbria, UK, 1997, pp. 75–82.

[18] Sompayrac, L., *How the immune system works*, Malden, Mass.: Blackwell Science, 2nd Edition, 2003.

[19] Esponda, F., Forrest, S. and Helman, P., "A Formal Framework for Positive and Negative Detection Schemes," *IEEE Transactions on Systems, Man, and Cybernetics- Part B: Cybernetics*, vol. 34, no. 1, 2004, pp. 357–373.

[20] Hofmeyr, S. and Forrest, S., "Architecture for an Artificial Immune System," *Evolutionary Computation*, vol. 7, no. 1, 2000, pp. 1289–1296.

[21] Cohn, M., "Tritope model of restrictive recognition by the TCR," *Trends in Immunology*, vol. 24, no. 3, 2003, pp. 127–131.

[22] Somayaji, A. and Forrest, S., "Automated Response Using System-Call Delays," *Proc. Usenix*, San Diego, California, 2000.

[23] Hofmeyr, S. and Forrest, S., "Immunity by Design: An Artificial Immune System," *Proc. Genetic and Evolutionary Computation Conference (GECCO)*, Edited by. Banzhaf, W. et. al., 1999, pp. 1289–1296.

[24] Williams, P., Anchor, K., Bebo, J., Lamont, G. and Gunsch, G., "CDIS: Towards a Computer Immune System for Detecting Network Intrusions," *Lecture Notes in Computer Science*, vol. 2212, 2001, pp. 117–133.

[25] Kim, J., Ong, A. and Overill, R. E., "Design of an Artificial Immune System as a Novel Anomaly Detector for Combating Financial Fraud in the Retail Sector," *Proc. Congress on Evolutionary Computation*, Edited by. Sarker, R. et. al., Canberra, Australia: IEEE Press, 2003, 405–412.

[26] Kendall, K., "A Database of Computer Attacks for the Evaluation of Intrusion Detection Systems," *Masters Thesis*, Massachusetts Institute of Technology, 1999.