Self Tolerance by Tuning T-Cell Activation: An Artificial Immune System for Anomaly Detection

Mário J. Antunes[1,3] and Manuel E. Correia[2,3]

[1] School of Technology and Management, Polytechnic Institute of Leiria, Morro do Lena, Alto do Vieiro, 2411-901 Leiria, Portugal
mario.antunes@ipleiria.pt

[2] Department of Computer Science, Faculty of Science, University of Porto, Rua do Campo Alegre s/n, 4169-007 Porto, Portugal
mcc@dcc.fc.up.pt

[3] Center for Research in Advanced Computing Systems (CRACS), Portugal

Abstract. The Artificial Immune Systems (AIS) constitute an emerging and very promising area of research that historically have been falling within two main theoretical immunological schools of thought: those based on *Negative selection* (NS) or those inspired on *Danger theory* (DT). Despite their inherent strengths and well known promising results, both deployed AIS have documented difficulties on dealing with gradual dynamic changes of self behavior through time.

In this paper we propose and describe the development of an AIS framework for anomaly detection based on a rather different immunological theory, which is the Grossman's Tunable Activation Thresholds (TAT) theory for the behaviour of T-cells. The overall framework has been tested with artificially generated stochastic data sets based on a real world phenomena and the results thus obtained have been compared with a non-evolutionary Support Vector Machine (SVM) classifier, thus demonstrating TAT's performance and competitiveness for anomaly detection.

Keywords: Artificial Immune Systems, Tunable Activation Thresholds, Pattern Recognition, Signal Processing, Support Vector Machine.

1 Introduction

The Vertebrate Immune System (IS) is a complex biological system, conceptually structured into two main functional layers: *innate* and *adaptive*. The anomaly detection embodied by the IS has to cope with a highly dynamic environment where the body is constantly being exposed to external agents (*pathogens*). Thus, this interaction with the environment results in a distinction between what is benign or belong to the organism's own healthy cells and tissues (*self*) from what is harmful and may provoke harm or even motivate a disease (*non-self*). In the majority of cases, pathogens presented to the body correspond to unseen flavours of normal activity that do not represent any serious danger or are benign [1].

J. Suzuki and T. Nakano (Eds.): BIONETICS 2010, LNICST 87, pp. 1–15, 2012.

In the late 19^{th} century, Ehrlich's started by postulating that the IS *"classifies"* pathogens as normal body antigenic components (*self antigens*) or as foreign abnormal chemical structures present in microorganisms (*non-self antigens*). After all these years, immunology continues to be a very vibrant and active open research area where no one has a definitive answer on how the IS is able to accomplish its goals in such an efficient and effective way.

One of the mainstream theories for several years, Negative Selection (NS) [2], assumes that cell activation thresholds have evolved to optimal values and constitute an intrinsic feature of each species. In the early 90's Matzinger described her controversial "Danger Theory" immunological theory, which states that the immune system is activated upon the receipt of molecular signals (danger signals) which indicate damage or stress to the host, rather than by a self-non-self distinction as previously postulated by NS [3, 4].

The IS provides a very appealing metaphor for the development of innovative anomaly detection systems in the form of an Artificial Immune Systems (AIS) [5]. The research in this area is based in principles, mechanisms, models and observed functions of the IS behavior, together with engineering best practices and methodologies. The most relevant AIS developed so far for anomaly detection have been based on the NS approach [6] and on the Danger Theory (DT) [4, 7]. In spite of the results achieved, the deployment of AIS based on both theories to solve real world problems [6, 8] did not yet met the expectations raised by such appealing metaphors [9] [7]. Firstly, the NS approaches proved to be inappropriate for large data sets and have shown to have scaling problems [9, 10]. Secondly, DT approaches have not intrinsic self-tuning mechanisms and require a great deal of expert knowledge beforehand [7, 8].

The well known best of breed AIS are all based on well reasoned immunological metaphors. However, the lack of research into their real biological foundations has led to well known criticism for biologically-inspired engineering approaches [11,12]. As an example, consider an *ideal* anomaly detector. It should be ready to act on a continuous changing environment and it should adapt itself throughout time to tolerate *unseen* and *untrained* forms of normal behavior, thus discriminating ongoing not yet seen anomalies. Such immune-inspired properties that mimic this self-non-self discrimination behavior are essential for the deployment of bio-inspired anomaly detectors within dynamic environments.

In this paper we present a generic AIS framework for anomaly detection, based on a simplified Tunable Activation Threshold (TAT) model, strongly inspired on Grossman's hypothesis [13]. TAT assumes that immune cells (like T-cells) tune their activation thresholds by dynamically updating the levels of two particular enzymes (Kynase and Phosphatase), whose values reflect the recent temporal history of signaling they have been receiving from the environment.

We start by defining a TAT model for T-cells (Section 2) and proceed into Section 3 by presenting the framework and its main building blocks. This is followed by showing and discussing the results obtained with stochastically generated, but real world based, data sets. A comparison is made with the non-evolutionary

classifier Support Vector Machine (SVM) (Sections 4 and 5). Finally, in Section 6 we delineate some conclusions for our work.

2 The TAT Model Adopted

The Grossman's TAT conceptual framework hypothesizes that immune cell activation depends on a dynamically adjusted threshold, which corresponds to the balance between excitation and de-excitation signalling pathways [14]. The activation process is controlled by the activity of two specific enzymes that respond to antigenic signals (S): Kinase (K) phosphorilates molecules that "excite" the cell and Phosphatase (P) that dephosphorilates them, returning the cell to a de-excitation state. The signals are delivered by a particular immune cell named Antigen Presenting Cell (APC).

It is also assumed that T-cell activation is a switch-type response that requires that K supersedes P, at least transiently. At each point in time, T-lymphocytes (T-cells) interact with the peptides presented by APC and receive a stimulus that depends on the *affinity* between its receptor and the peptide ligand, causing the cell to adapt by increasing or decreasing its activation threshold. Also, the de-excitation level is assumed to be intrinsically slow, thus allowing the outcome of a stimulus to depend mainly on the excitation index. Thus, foreign antigens will cause a very fast increase in the cells excitation level, whereas tissue-specific self-ligands will induce a much slower increase excitation level. Accordingly, since the de-excitation levels are kept above the excitation ones, it is possible to maintain tolerance to self for extended periods of time [15]. Within this model, different cells with different antigen-specificity end up having different activation thresholds as they are exposed to different stimuli.

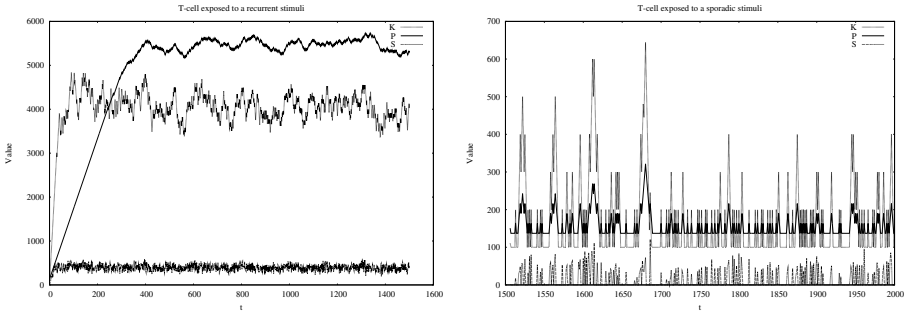
2.1 TAT Dynamics

We have adopted a minimal mathematical model of TAT for T-cells [16], which is also derived from Grossman's hypothesis. Keeping the original Grossman's fundamental thoughts about self regulation and cellular activation, we made the following simplifications:

- both K and P are exposed to the same stimulus S;
- P's basal value (P_0) is higher than K's (K_0);
- S_0 is the initial value for S;
- K's turnover rate (τK) is lower than P's (τP);
- K's slope (ϕK) is higher than ϕP's;
- the IS's speed of response is given by a constant value (t);

We also derived the following values:

- $K_0 = S_0 \cdot \tau\mathrm{K}$ and $P_0 = S_0 \cdot \tau\mathrm{P}$;
- $\tau K = \tau \cdot \tau P$, with $\tau = \frac{\tau K}{\tau P}$;
- $\phi P = \phi \cdot \phi K$, with $\phi = \frac{\phi P}{\phi K}$;



(a) (b)

Fig. 1. TAT dynamics of two individual T-cells in the AIS repertoire. 1(a): the cell is exposed to a recurrent stimuli, like tissue-specific self ligands.1(b): the cell receives an intermittent and strong signal which leads to repeated events of activation.

Figure 1 depicts the TAT activity of two T-cells. S corresponds to a linear increase of both K and P activities until the turnover rate is reached. T-cell activation only occurs if K is higher than P. For a recurrent signal S, K will be transiently higher than P and, if the signal persists P will exceed K and the cell will become inactive. Similarly, on signaling absence, K returns to the initial level at a faster rate than P.

2.2 Signaling Model

In such a model, at each given moment in time, the stimulation history of a T-cell is reflected in the activity of K and P. The signal S sent by the APC to a T-cell is a function of the affinity between the corresponding T-cell Receptor (TCR) and the ligand times the concentration of that peptide in the APC. To give strength to the temporal meaning of the TAT dynamics, S is calculated in a *per APC in the lifespan (LS)* basis, as shown in Algorithm 1.

The values of K and P of each cell repertoire are updated linearly based on the combined signal of all the ligands presented by each APC, as described in Algorithm 2. The use of a linear update of such values gave simplicity to the model, being at the same time in accordance with the Grossman's derived model depicted in [16]. In short, in the presence of a signal ($S > 0$), K and P increase till reach its corresponding maximum values (τK and τP). Otherwise, they decrease gradually till the basal values ($K_0 = S_0 \cdot \tau K$ and $P_0 = S_0 \cdot \tau P$). The growth and decline rates are different for both K and P, which leads to episodes of cell activation when K becomes higher that P.

Grossman also postulated that an immunological response to an APC is usually not initiated by an individual activated cell, being instead initiated by a *population clones* of activated cells. Thus, on an APC processing, an immune response is initiated if the ratio of the population size of activated cells and

Algorithm 1. Signalling model adopted

Input: L_{pep} = List of peptides presented by an APC
Input: L_{tcell} = T-cell Repertoire
Output: S = signal sent for each t-cell in the repertoire

```
forall the tcell in L_tcell do
    forall the peptide in APC's Lifespan (including those in L_pep) do
        a = distance(peptide, tcell)
        c = occurrences of peptide in the APC lifespan
        S+ = Σ(c · a)
    end
end
```

Algorithm 2. Update of K and P for a T-cell, based on a received signal

Input: S = Stimuli received by a T-cell (calculated in Algorithm 1)
Input: t = Real value corresponding to the *speed* of response of the system
Output: Updated values for K and P

```
if (S + S_0) · τK) > K then
    K ← MIN((S + S_0) · τK, K+ = φK · t)
end
else
    K ← MAX((S + S_0) · τK, K− = φK · t)
end

if (S + S_0) * τP) > P then
    P ← MIN((S + S_0) · τP, P+ = φP · t)
end
else
    P ← MAX((S + S_0) · τP, P− = φP · t)
end
```

those that were bound exceeds a threshold. This thus implies that an immune response depends always on the decision of a group of cells (*committee*), instead of an individual one [16].

3 The TAT Based Framework

A generic AIS framework can be divided into three main functional layers [5]: a data representation (Section 3.2), an affinity measure distance between the immune cells receptors and the peptide ligand (Section 3.2) and, finally, an immune-inspired algorithm that maps the system components with the relevant biological IS counterparts (Section 3.4). In what follows we describe in some detail the core building blocks of the TAT-AIS framework, depicted in Figure 2.

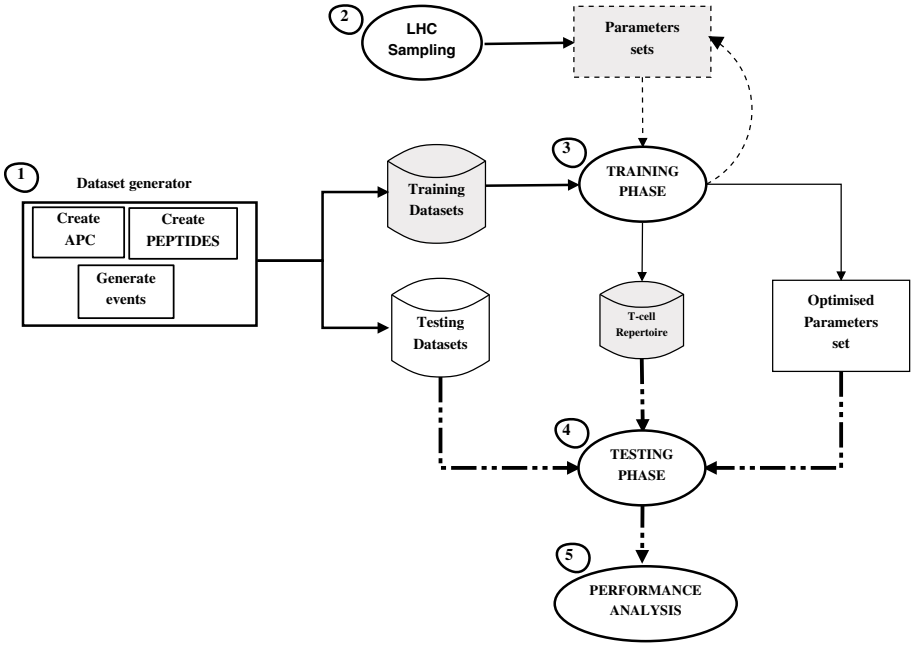


Fig. 2. Building blocks of the TAT-based AIS

Our system is composed by the following three main components: an artificial data set generator, a parameter set sensitivity analysis module and a TAT simulator that implements the TAT model previously described, for processing both training and testing data sets.

The adopted methodology is described as follows. Firstly, using the data set generator we produce the training and testing data sets according to the procedure described in Section 3.2. Then we obtain a set of parameters using an Latin Hypercube (LHC) based sampling method [17] (Section 3.3). Each one of these parameter set is then evaluated against the training data sets in order to obtain an *optimised* parameters set and a list of *trained* detectors (T-cells) that are finally confronted with a group of testing data sets, in order to analyse their behaviour and its performance compared with other competing classifiers.

3.1 The TAT Operation

The TAT simulator requires training, which is comprised by two different steps. Firstly, we have split the training data set into two sub-data sets: one that is used for training TAT with normal examples and the other containing a mixture of randomly interleaved known normal and abnormal examples. We then run the TAT simulator with the normal examples to obtain a list of *self* T-cells that converge to a *"P higher than K"* state. Finally we load the obtained T-cell

repertoire and process the second sub-set. During this phase TAT creates the T-cells responsible for the detection of abnormal patterns. The resulting repertoire is then loaded into the system for further processing the testing data set .

3.2 Data Representation

The most relevant immunological players involved in the TAT model and suitable to be represented in the artificial framework are peptide ligand, TCR and APC. Generally speaking, from a machine learning point of view, both peptide ligand and TCR are *patterns* and can be represented by strings that bound to each other by some affinity measure. APC can thus be seen as a list of characters (peptides) representing a behavior (normal or abnormal). Also, APC are ordered sequentially and tagged into two distinct classes ("Normal" or "Alert") for evaluation purposes.

The reason behind the development and deployment of an artificial data set generator was two-fold. Firstly, we intended to generate simple and easy to understand artificial data sets with which we could test the model and compare the results with other classifiers. Secondly, although being simple, we would like the data sets to be based, as much as possible, on real world phenomena for anomaly detection purposes. We have thus decided to use the spam Enron data sets [18] and identify some of its main characteristics and use them for artificial data set generation. The Enron preprocessed data sets have six personal mailboxes made public after the Enron scandal. The ham mailboxes belong to six employees and combinations of five spam data sets were added to the ham data, coming from different sources [18, 19]. Table 1 describes the distribution of ham and spam words (average values), considering all the spam and ham messages available for each mailbox and an identical size for training and testing data sets.

Table 1. Analysis of Enron spam data sets

Data set	Words	Occurrences	Messages	Words/Msg.	Weight
Ham Training	954	74497	920	78.09	1.81
Ham Testing	140	6906	1838	49.22	1.14
Spam Training	164	7098	954	43.24	**1**
Spam Testing	117	6114	1907	53.34	1.21

From the Table 1 it is possible to identify the average amount of words in each example of the listed data sets (column *Words/Msg*), as well as the proportional relation of each data set with the data set that has the lowest ratio of words per message (column *Weight*). For instance, the ham training data set has 954 different words that appears 74497 times in the 920 messages. Thus, the average of words per message in the ham training data set is about 78. Comparing to the data set with the lowest ratio (spam training), the proportional relation is 1.81. To generate the artificial data sets we defined a symbol set of length 8

for the spam training data set (the one with the lowest ratio) and built the three remaining data sets according to these proportional weights. Thus, the symbols sets lengths are 14 (corresponding approximately to 8 times 1.81), 9 (8 times 1.14) and 10 (8 times 1.21), respectively for ham training, ham testing and spam testing. The Table 2 depicts the symbols sets used to generate the artificial data sets, based on the analysis described previously.

Table 2. Alphabet used to create the peptides

Symbols	**Data sets**	**Tag**
a b c d e f g h i j k l m n	Training,Testing	Normal
A B C D E F G H I	Testing	Normal
1 2 3 4 5 6 7 8	Training,Testing	Alert
+ - . _ : ; ? ! = $	Testing	Alert

In the data sets described above, both peptide ligand and TCR are represented by one character and an APC corresponds to a list of characters. The affinity metric adopted to measure the distance between both peptide ligand and TCR is the character match (one if equal and zero otherwise). This strategy seemed appropriate to calculate the affinity between strings of just one character, which in some way represents *words* of the email messages. Also, an APC tagged as normal has all the peptides produced with symbols picked randomly from the "Normal" alphabets symbols. On the other hand, an APC is labelled as alert if it has at least one peptide belonging to an "Alert" alphabet symbol.

The data sets generation took into count two crucial parameters: the number of peptide ligand (characters) per APC, which varies along the data set and the number and time of occurrence of events for each class ("Normal" and "Alert"). The values for these two parameters were also randomly generated for each data set. The following main features could then be observed in the artificially generated data sets:

- patterns representative of normal behavior appear recurrently, while abnormal ones appear sporadically;
- in the testing phase new unseen patterns representative of normal behavior start appearing recurrently;
- also in the testing phase, new unseen patterns representative of abnormal behavior appears sporadically;
- finally, new unseen anomalies are propagated by the conjunction of both known and unknown patterns, representative of abnormal behavior.

An example of normal and alert APC looks like the following:

```
apc:38:NORMAL: o d n j i m d h p k c k f l b i c l n k i a a k

apc:39:ALERT:  m c 3 3 6 i g n n 6 2 h 6 b e g o j 2 8 f 1 i d 7 p h 2
               7 b m a
```

For the experiments detailed in Section 4 we generated APC with a maximum of 200 peptides. The training data set is composed by 2000 APC sequentially ordered, being the first 75% representative of normal behavior and the remaining related to *known* abnormal examples. The testing data set is twice the size of the training (4000 APC).

3.3 Parameters set Optimisation

The T-cells' TAT dynamics is based on the parameters described on Table 3.

Table 3. Parameters set

Parameter	Range	Description	Type
τ	$[0; 1]$	K and P turnover rate, $\frac{\tau P}{\tau K}$	Optimised
ϕ	$[0; 1]$	K and P slopes rate, $\frac{\phi K}{\phi P}$	Optimised
t	$[10; 100]$	Speed of adaptation and response for the IS	Optimised
LS	$[10; 100]$	Number of APC in the lifespan	Optimised
CS	$[10; 100]$	T-cell maximum clonal size	Optimised
Ct	$[0.00001, 0.8]$	Committee threshold.	Run time
i		update factor to increase/decrease the T-cell CS	Fixed=2
S_0		Initial signal	Fixed=2
CS_0		Initial clonal size	Fixed=2
a		Affinity distance between T-cells and peptides	Fixed=1

We used a hybrid approach to choose a sufficiently good performant parameter set. Firstly, we have listed all the parameters related to the TAT dynamics (Section 3.5) and its corresponding ranges. Then, we generated an LHC sampling for the following parameters: τ, ϕ, t, LS and CS. LHC sampling is a statistical method developed to generate a distribution of collections of parameter values from a multidimensional distribution [17]. For our case we sampled the five parameters into 40 equally probable intervals.

Our model is also evolutionary. So, we defined that Ct will be updated in the testing phase by using a feedback mechanism described in Section 3.6. Ct starts with a fixed value ($Ct = 0.1$) and, as long as the system is having a too high or too low rate of alerts, this parameter increases or decreases accordingly in a gradual way.

Finally, using the combinations calculated by the LHC sampling method and the defined fixed parameters, we run a 10-fold training data set and then calculate the average of the *F1-Measure* thus obtained. This measure is a score of accuracy that considers both the precision p and the recall r, being calculated by $F1 = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}$ [20].

At the end of the 10-fold training process we obtain a good LHC's candidate multidimensional *"square"*, that corresponds to the parameters set with which we obtained the best performance thus far. Depending on the accuracy level, the sampling process can be recursively repeated with an even more strict hypercube, to try to obtain an even more refined parameter set of values that can produce better results.

3.4 General Algorithm

The algorithm 3 describes the TAT processing of each APC. It receives a list of peptides (L_{pep}) and the up to date list of T-cells (L_{tcell}) with its corresponding K and P values.

At each given moment, each T-cell will be stimulated with a signal that corresponds to the sum of signals sent by each affinity-specific peptide. If there is not a T-cell that bounds with a peptide, then a new one is created with the TCR being the string representative of the peptide ligand. After processing all the peptides of the APC, the system calculates the amount of clonal size for both bound and active T-cell. If the ratio between both values is above a threshold (Ct), then the system raise an alarm and increases the clonal size of all the activated cells. Otherwise, the clonal size is decreased (Section 3.5).

3.5 Clonal Size Update Procedure

In our model the clonal size (CS) of each cell corresponds to an integer that varies between an initial value (CS_0) and a maximum (CS_{max}). The clonal size update procedure in each processing phase is depicted in Table 4. We also introduced the meaning of *committee* as being the clones population of cells that bound with a certain specificity the peptides ligand presented by an APC.

In each processing phase, the decision rule is thus based on the ratio between the cells of the committee that are activated and those that simply bond with the peptide ligand but remains quiescent. In the training phase, the learning procedure is supervised and the clonal size of activated T-cells ($K > P$) increases if the *committee* decided in favor to trigger a response. Otherwise, the CS decreases. The testing phase is unsupervised and the clonal size update is based on the APC classification made by the system in each moment. In general, this CS update mechanism allow that sporadically activated T-cells converge to a maximum value (CS_{max}) and the recurrently stimulated (but not activated) ones may have a clonal size near CS_0. For the sake of simplicity the following assumptions was made: the increment factor is fixed, $i = 2$ and CS update is made linearly ($CS \pm = i$)

3.6 Feedback Control Mechanism

The evolutionary tuning of committee threshold is as follow: Ct starts with a predefined value (Table 3). Then, in each APC processing, if the number of

Algorithm 3. General TAT algorithm

Input: L_{tcell} = T-cell Repertoire
Input: L_{pep} = List of peptides presented by an APC
Input: At = Affinity Threshold
Input: Ct= Committee Threshold
Output: Classification of the artificial APC: *Normal* or *Alert*

```
1  T_bind = ()
2  T_active = ()
3  CS_bind = 0
4  CS_active = 0
5  forall the tcell in L_tcell do
6  |  S = 0
7  |  forall the UNIQUE(peptide) in L_pep do
8  |  |  a = distance(peptide, tcell)
9  |  |  c = occurrence of peptide ligand in the APC lifespan
10 |  |  if a >= At then
11 |  |  |  S+ = Σ(c · a)
12 |  |  |  ADD(T_bind, tcell)
13 |  |  end
14 |  end
15 |  UpdateTCell(t, S) (according to Algorithm 2)
16 end
17 forall the tcell in T_bind do
18 |  if K >= P then
19 |  |  ADD(T_active, tcell)
20 |  end
21 end
22 CS_active = ΣClonalSize(T_active)
23 CS_bind = ΣClonalSize(T_bind)
24 Status = Normal
25 if CS_active/(CS_bind+CS_active) >= Ct then
26 |  Status = Anomaly
27 end
28 ReportStatus()
29 UpdateClonalSize(Status) according to the procedure described on Section 3.5
```

triggers observed so far exceeds a preliminary threshold, then the value for Ct is incremented. Otherwise, if there is no trigger during a long period, the Ct decreases. The number of *acceptable* triggers and the values used to increase or decrease the parameter Ct should be domain-dependent. In the experiments we defined 10% as being an acceptable value for the alerts. We imposed that Ct should vary between 0.00001 and 0.8 and its value depends on the observations made to the system in run time.

Table 4. Clonal size update process

	Normal		**Known Alerts**		**Testing**	
	K>P	P>K	K>P	P>K	K>P	P>K
Supervised						
Tag="NORMAL"	↗	↘	↘	↘		
Tag="ALERT"			↗	↘		
Unsupervised						
Trigger an alert					↗	↘
Silent mode					↘	↘

4 Experimental Evaluation and Results

Our working hypothesis is that the deployed TAT based model can be able to recognise new unseen patterns and also to further distinguish between those that are considered self from others included in APC related to abnormal activities and should be identified as non-self. In order to validate our model we compared the results obtained with a non-evolutionary SVM classifier.

For both TAT and SVM algorithms we used the following methodology. Firstly, we trained the system with a 10-fold training data set. Then we manage to process the testing phase with ten different data-sets. Finally we evaluate the performance obtained by each detection algorithm.

4.1 TAT Parameters

We run the parameters optimisation methodology described in Section 3.3 and obtained the parameters set listed on Table 5. The table shows the performance obtained during the training phase (average of the 10-fold training processing) considering the affinity as being the full match of each TCR with each peptide ligand presented in the APC.

Table 5. The parameters set with the best performance during the training phase

τ	ϕ	T	LS	CS	Ct	**Accuracy**	**Precision**	**Recall**	**F1**
0.78758	0.25646	25	99	24	0.1	0.98	1.00	0.92	0.96

4.2 Results

Table 6 depicts the results obtained with both TAT and SVM implementations. The table clearly shows the performance obtained for accuracy, precision, recall and F1 score for both TAT and SVM processing.

Table 6. Results obtained with TAT and SVM processing

	TAT			**SVM**		
Dataset	*Precision*	*Recall*	$F1$	*Precision*	*Recall*	$F1$
1	0.98	0.75	0.85	1.0	0.72	0.84
2	0.98	0.67	0.8	1.0	0.73	0.84
3	0.99	0.75	0.85	1.0	0.75	0.85
4	0.98	0.76	0.86	1.0	0.71	0.83
5	0.99	0.78	0.87	1.0	0.73	0.85
6	0.97	0.76	0.86	1.0	0.74	0.85
7	0.96	0.85	0.9	1.0	0.72	0.84
8	1.00	0.70	0.82	1.0	0.71	0.83
9	0.98	0.72	0.83	1.0	0.67	0.80
10	0.97	0.77	0.86	1.0	0.76	0.86
Mean	0.98	0.75	**0.850**	1.0	0.72	**0.839**

5 Discussion

We have observed that TAT-AIS has interesting properties for anomaly detection, provided the following basic generic requirements are true: *normal behavior is frequent and abnormal behavior is sporadic in time*. By *frequent* we mean a pattern that repeatedly stimulates a set of T-Cells that through time, by the TAT dynamics, stabilizes its enzymatic values ($P > K$). On the other hand, by *sporadic* we mean a pattern that stimulates intermittently a set of T-Cells with such a signal that implies its activation ($K > P$).

Our aim was to validate the appropriateness of using TAT to detect new previously unseen patterns and also to distinguish them between those that correspond to unseen "normal" and "abnormal" behaviors. We have also compared the results obtained with a non-evolutionary SVM classifier. The results show the competitiveness of TAT when comparing it with SVM. The F1-measure varies between 80% and 90% in TAT. In SVM the results are in the range of 80% and 86%. In general SVM has no false positives (Precision=100%) but not all the new alerts are correctly identified. On the other hand, TAT has some false positives (mean of Precision is 98%) but the recall is slightly higher than the SVM.

The performance advantage obtained with TAT-AIS is very tiny, when compared with the non-evolutionary SVM classifier. However, based on these empirical results, we believe that our model can compete with other approaches on the self-non-self distinction for dynamic environments that tend to change gradually throughout time their normality behavior profile. In this paper we were not aware with the performance of the speed of classification. However, in these experiments we observed similar execution times with both models.

6 Conclusions

In this paper we have presented a generic TAT-based AIS framework for anomaly detection and described its main architectural components. We have also presented some results obtained with artificially generated data sets of predefined patterns resulting from normal and abnormal behaviors. Also, we have compared the TAT performance with a non-evolutionary SVM classifier.

The results thus obtained with the AIS are very satisfactory, achieving a high rate of detection and a low level of false positives on the stochastic data sets we have produced.

We are well aware that these stochastic data sets were artificially generated and are most certainly not completely representative of real world phenomenons like the data sets we could obtain with email spam collections. We have however already obtained some preliminary good results with this TAT-based framework, applied both to more complex stochastically generated data sets [21], as well as to network intrusions detection with real network traffic [22]. The research done so far give us confidence on the use of TAT based AIS framework to implement behavior based anomaly detection systems, where the temporal meaning of events is relevant, like spam and intrusion detection. The ongoing research is now on using the model and the framework presented to process the original Enron spam data sets and to compare its performance with the SVM classifier.

Acknowledgments. The authors acknowledge the facilities and research environment gracefully provided by the Center for Research in Advanced Computing Systems research unit. The authors also thank Catarina Silva (Polytechnic Institute of Leiria and CISUC - University of Coimbra) by her insights and help on the SVM data processing.

References

1. Murphy, K., Murphy, K., Travers, P., Walport, M., Janeway, C.: Janeway's immunobiology. Garland Pub. (2008)
2. Burnet, F.: The Clonal Selection Theory of Acquired Immunity. University Press Nashville, Tenn (1959)
3. Matzinger, P.: Tolerance, danger, and the extended family. Annual Review of Immunology 12(1), 991–1045 (1994)
4. Aickelin, U., Greensmith, J.: Sensing danger: Innate immunology for intrusion detection. Information Security Technical Report 12(4), 218–227 (2007)
5. de Castro, L., Timmis, J.: Artificial Immune Systems: A New Computational Intelligence Approach. Springer (2002)
6. Hofmeyr, S., Forrest, S.: Architecture for an artificial immune system. Evolutionary Computation 8(4), 443–473 (2000)
7. Kim, J., Bentley, P., Aickelin, U., Greensmith, J., Tedesco, G., Twycross, J.: Immune system approaches to intrusion detection - a review. Natural Computing 6(4), 413–466 (2007)

8. Greensmith, J., Aickelin, U., Cayzer, S.: Introducing Dendritic Cells as a Novel Immune-Inspired Algorithm for Anomaly Detection. In: Jacob, C., Pilat, M.L., Bentley, P.J., Timmis, J.I. (eds.) ICARIS 2005. LNCS, vol. 3627, pp. 153–167. Springer, Heidelberg (2005)
9. Stibor, T., Mohr, P., Timmis, J., Eckert, C.: Is negative selection appropriate for anomaly detection? In: Proceedings of the 2005 Conference on Genetic and Evolutionary Computation, pp. 321–328 (2005)
10. Kim, J., Bentley, P.: An evaluation of negative selection in an artificial immune system for network intrusion detection. In: Genetic and Evolutionary Computation Conference, pp. 1330–1337 (2001)
11. Andrews, P., Timmis, J.: Tunable Detectors for Artificial Immune Systems: From Model to Algorithm. In: Bioinformatics for Immunomics, pp. 103–127 (2010)
12. Stepney, S., Smith, R., Timmis, J., Tyrrell, A., Neal, M., Hone, A.: Conceptual frameworks for artificial immune systems. International Journal of Unconventional Computing 1(3), 315–338 (2005)
13. Grossman, Z.: Cellular tolerance as a dynamic state of the adaptable lymphocyte. Immunology Reviews 133, 45–73 (1993)
14. Grossman, Z., Paul, W.: Self-tolerance: context dependent tuning of T cell antigen recognition. Seminars in Immunology 12(3), 197–203 (2000)
15. Scherer, A., Noest, A., de Boer, R.: Activation-threshold tuning in an affinity model for the T-cell repertoire. In: Proceedings: Biological Sciences, vol. 271(1539), pp. 609–616 (2004)
16. Carneiro, J., Paixão, T., Milutinovic, D., Sousa, J., Leon, K., Gardner, R., Faro, J.: Immunological self-tolerance: Lessons from mathematical modeling. Journal of Computational and Applied Mathematics 184(1), 77–100 (2005)
17. Helton, J., Davis, F.: Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. Reliability Engineering & System Safety 81(1), 23–69 (2003)
18. Metsis, V., Androutsopoulos, I., Paliouras, G.: Spam filtering with naive bayes-which naive bayes. In: Third Conference on Email and Anti-Spam (CEAS), pp. 125–134 (2006)
19. Abi-Haidar, A., Rocha, L.: Adaptive Spam Detection Inspired by the Immune System. In: Artificial Life XI - 11th Int. Conference on the Simulation Ans Synthesis Os Living Systems, vol. 11, pp. 1–8 (2008)
20. Silva, C., Ribeiro, B.: Inductive Inference for Large Scale Text Classification: Kernel Approaches and Techniques. Springer (2009)
21. Antunes, M., Correia, M.: Temporal Anomaly Detection: an Artificial Immune Approach Based on T-cell Activation, Clonal Size Regulation and Homeostasis. Advances in Computational Biology - Book series vol. 680, pp. 291–298 (2010)
22. Antunes, M., Correia, M.: TAT-NIDS: an immune-based anomaly detection architecture for network intrusion detection. In: Corchado, J.M., et al. (eds.) IWPACBB 2008. ASC, vol. 49, pp. 60–67. Springer, Heidelberg (2008)