International Journal of Multimedia and Ubiquitous Engineering
Vol. 9, No. 11 (2014), pp. 85-98
http://dx.doi.org/10.14257/ijmue.2014.9.11.09

A Fast Immune Recognition Model based on Immune Response

Yuan Tao[1] and Min Hu[2]

[1]*Computation Center, Shanghai University, Shanghai 200444, China*
[2]*SILC Business School, Shanghai University, Shanghai 201800, China*
[1]*taoyuan103@shu.edu.cn*, [2]*minahu@163.com*

Abstract

This paper proposes a fast immune recognition model based on immune response. Inspired by biological immune system, AIS for anomaly detection has been adopted widely because of its analogy with body resistance in the immune system provided against agents which causes diseases. This paper mainly studies correspondence between immune response and anomaly detection. Compared with traditional AIS, the change of antigen type is judged by statistical techniques, which ignores the differences of the different systems. Antibody recognition is initiated by the proposed model only when the type of antigen has changed, which improves the efficiency of the algorithm. Complexity analysis shows the proposed algorithm is a linear algorithm. The usefulness of the proposed model is demonstrated through experiments. The experiments illustrate the availability and feasibility of the model.

Keywords: *Antigen Evaluation, Antibody Recognition, Immune Response*

1. Introduction

The immune system is a remarkable information processing and self-learning system that offers inspiration to build Artificial Immune Systems (AIS). The field of AIS has obtained a significant degree of success as a branch of Computational Intelligence since it emerged in the 1990s [1]. The powerful information processing capability, pattern recognition, learning, memory and immune distributive nature provide rich metaphors for its artificial (computational) counterpart [2]. Immune-based techniques have been applied in a wide area of applications and successful in anomaly detection [3].

Biological immune system has successfully solved the problem of unknown virus detection [4]. The immune system has a natural advantage in anomaly detection and has the ability to learn the unknown risks. Therefore, immune techniques applied to anomaly detection are currently gaining significant attention from researchers in various fields, and have been successfully developed due to their effectiveness and robustness in anomaly detection [5-7].

People have been continuously trying to establish appropriate theories that can best explain the recognition mechanisms of the immune system. There are many models in biology that attempt to explain the immune system behavior. In the last decade, the immune system has drawn significant attention as a potential source of inspiration for novel approaches to solving complex computational problems [8]. Among those theories, Self-Nonself (SNS) model, Infectious-Nonself (INS) Model and Danger Model are most popular.

Anomalies are patterns in data that do not conform to a well-defined notion of normal behavior. Anomaly detection refers to the problem of finding patterns in data that do not conform to expected behavior [9]. The anomaly detection problem can be stated as a two-class problem: given an element of the space, classify it as normal or abnormal. The

ISSN: 1975-0080 IJMUE

techniques include statistical, machine learning, data mining and immunological inspired techniques [10].

Anomaly detection techniques can operate in three modes: supervised anomaly detection, semi-supervised anomaly detection, and unsupervised anomaly detection. As it is hard to get anomaly class covering every possible anomalous behavior, semi-supervised and unsupervised techniques are more widely applied than supervised techniques. And, there is another problem that immune inspired systems have features whose the application environment must be considered on the implementation [11]. Besides, the complexity of the algorithm is also a very important problem.

In terms of the difficulties of anomaly detection, a fast immune recognition model based on immune response (FIRM) is proposed. FIRM combines the statistical technique and the sliding window technique to evaluate the change of the antigen's type, which overcomes the weakness of it hard to get the anomaly patterns. Antibody recognition is initiated only when the type of antigen has changed, which improves the efficiency of the algorithm. The rest of the paper is organized as follows: Section 2 briefly gives related work; Section 3 describes FIRM; the key issues of FIRM are given in Section 4; Section 5 shows the experimental results; finally, the conclusion is given in Section 6.

2. Related Work

2.1. Immunological Recognition Models and Theories

With extensive knowledge of the mammalian immune system, people have been continually trying to establish appropriate theories that can best explain the recognition mechanisms of the immune system in different angles. There are mainly three models in immunology, from Self–Nonself (SNS) model to Danger Model, as shown in Figure 1. These have important influences on AIS.

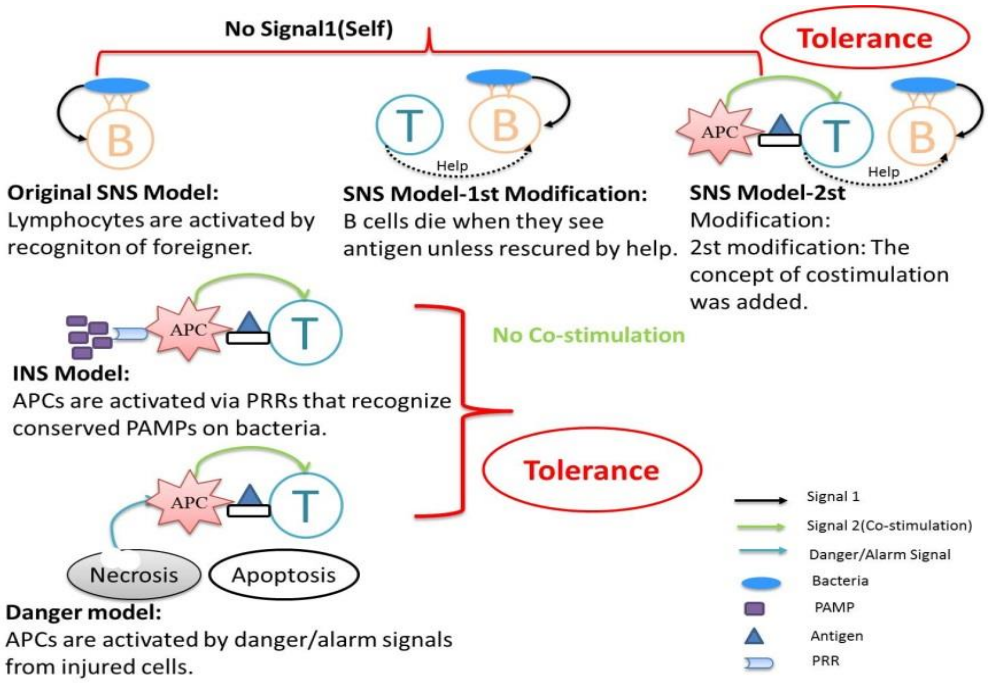


Figure 1. Three Important Immunological Models

Proposed by Burnet in 1959, the Self-Nonself (SNS) model is the first immunologic model. SNS model supposes that the immune system distinguish between self and nonself, tolerating self and attacking nonself. As new evidence continuously accumulated SNS was modified twice in 1969 and in 1975. In1969, Bretscher and Cohn added a new cell (T cell) and a new signal (help), proposing that the B cell would die if it recognized antigen in the absence of help [12]. In 1975, Lafferty and Cunningham proposed that T cells also need a second signal (named "costimulation"), which they receive from "stimulator" cells (APCs) [13]. Costimulation principle suggested that the immune response is initiated by APCs. However, APCs are not antigen specific but they capture all sorts of self and foreign substances, then the immunity cannot be directed only against non-self. The concept of costimulation was therefore essentially ignored.

In 1994, Forrest et al. proposed negative selection algorithm (NSA), which is based on the principles of self/nonself discrimination in the natural immune system [14]. Although a diverse family of negative selection algorithms has been developed, the essential characteristics of the original negative selection algorithm still remain. The core principle of negative selection algorithms is to distinguish self from nonself. NSA is a two-class algorithm. The main limit is that the self and non-self is too clearly.

To solve the problems emerging from SNS model and costimulation, in 1989, Charles Janeway proposed the Infectious-Nonself (INS) Model. According to his theory, APCs are quiescent until they are activated via a set of germ line-encoded pattern recognition receptors(PRRs) that recognize evolutionary distant conserved pathogen-associated molecular patterns(PAMPs) on bacteria. It proposes that the innate immune system predominantly discriminates infectious nonself from noninfectious self by means of a system of pattern recognition receptors(PAMPs) [15]. If PAMPs of antigen match pattern in immune system, APC is activated to produce co-stimulation which induces helper T cells to initiate immune response. This model embodies the importance of local matching on immune recognition. Self/non-self discrimination can be accomplished by pattern recognition with PRRs. However, INS model has inspired few researches [16].

In 2002, Polly Matzinger proposed the Danger Model in the journal Science. Polly Matzinger suggested that the immune system is more concerned with damage than with foreignness, and is called into action by alarm signals from injured tissues, rather than by the recognition of nonself [17]. The theory is not complete, and there are some doubts about how much it actually changes behaviour and / or structure. Nevertheless, the theory contains enough potentially interesting ideas to make it worth assessing its relevance to Artificial Immune Systems [18]. There have been many successful applications of AISs, such as computer security, optimization, data mining, and anomaly detection, *etc*. [19-21].

Besides, new immune theories are constantly proposed to accommodate incompatible new findings, such as the cell death recognition model.

Cell death recognition model has four principles. First, only antigens shedding from apoptotic or necrotic cells, rather than from healthy cells, can be presented to naı¨ve T cells. Second, either apoptotic cells or necrotic cells, but not healthy cells, can attract phagocytes, namely DC (dendritic cells) or macrophages that are also APC(antigen presenting cells), to scavenge dead cells. Third, only macrophages or DC residing in non-lymphoid tissues phagocytose dying/dead cells, migrate to lymphoid tissues and then present antigens to naive T cells in lymphoid tissues or liver. Fourth, tolerance or adaptive response is not dependent on whether the antigens are self or nonself, but on the ways of cell death during antigen presentation [22]. Therefore, cell death recognition model supposes that apoptotic cells actively induce tolerance, and that necrotic cells initiate immune response.

2.2. Immune Response

As mentioned in Section 2.1, the immunological recognition models and theories focus on to what immune system responses. A comparison of the immune response in different immunological recognition models and theories is given in Table 1.

Table 1. A Comparison of the Immune Response in Different Immunological Recognition Models and Theories

Immunological Recognition Models and Theories	Response to
SNS model	Nonself
INS model	Infectious nonself
Danger Theory	The dangerous
Cell death recognition model	Necrotic cells

The immune response is the core event in the development of immunity. The immune response refers to how your body recognizes and defends itself against bacteria, viruses, and substances that appear extracorporeal and harmful. The immune system protects the body from possibly harmful substances by recognizing and responding to antigens [23].

Therefore, we take the immune response as our immune inspiration. The mapping between anomaly detection and immune system based on immune response is shown in Figure 2.

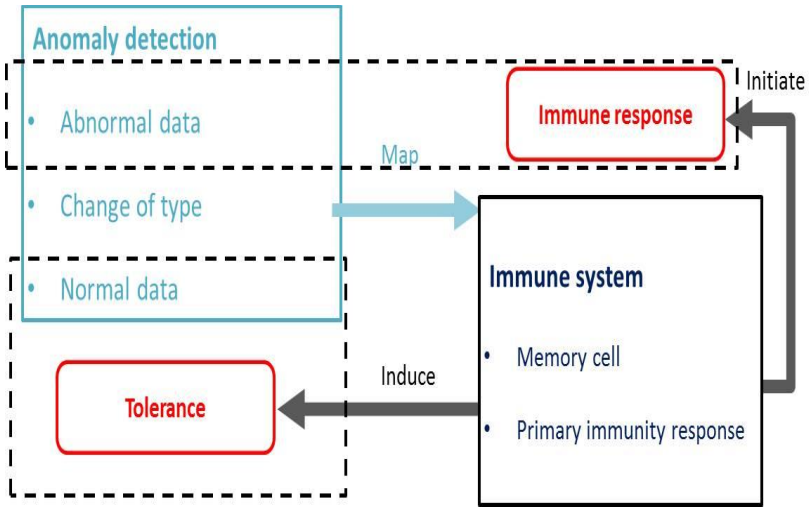


Figure 2. Mapping between Anomaly Detection and Immune System based on Immune Response

3. Fast Immune Recognition Model based on Immune Response

The problem can be formulated as follows: define antigen set $\{AG \in R^d\}$ and antibody set $\{AB \in R^d\}$; take input data as antigen$\{Ag \in AB\}$ and reference samples as antibody$\{Ab \in AG\}$; Ag should be recognized by Ab; the proposed method should be able to calculate the recognition results. As shown in Figure 3, a fast immune recognition model based on immune response (FIRM) is proposed.

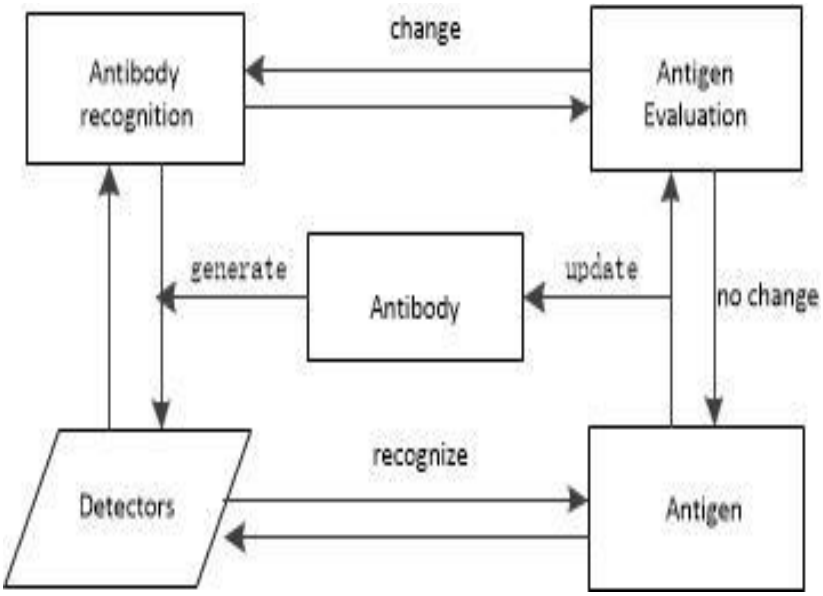


Figure 3. Fast Immune Recognition Model based on Immune Response

Inspired by immune response, FIRM mainly studies when to initiate immune response. Regarding memory of the immune system, FIRM first evaluates antigen. The antigen is treated as the prior antigen if the type of the antigen does not change; otherwise, antibody recognition is activated to determine whether or not it initiates immune response. In FIMR, abnormal data corresponds to the antigen which initiates adaptive immune; normal data corresponds to the antigen which induces response tolerance.

The key issues of FIRM include antigen evaluation, antibody recognition and antibody evolution.

4. Key Issues of FIRM

4.1. Antigen Evaluation

As for traditional AIS, similarity measurement is most popular to recognize antigen. Anomaly detection is becoming increasingly challenging because of the size and complexity of input data. That is why feature extraction and dimension reduction are always needed to preprocess data. At this point, statistical techniques are adopted to improve the efficiency of FIRM.

Apart from calculating the mean, estimating the variance of a random vector is the most basic problem in statistics. It has numerous applications in sciences, social sciences and humanities [24]. In many systems, inputting data is generally not independent, and to directly deal with independent samples is cumbersome. Moreover, there is no need to initiate antibody recognition if the state of antigen has not been changed. Aiming at this purpose, antigen evaluation is used to calculate the state of antigen.

Variance measures how far each number in the set is from the mean. Variance is calculated by taking the differences between each number in the set and the mean, squaring the differences (to make them positive) and dividing the sum of the squares by the number of values in the set. Variance is the best choice to calculate the state of antigen. And, variance is defined as:

$$\sigma^2 = \frac{\sum(X-\bar{x})^2}{N} \quad (1)$$

However, Ag may be a single sample in some systems, $N = 1$. So the key problem is to calculate variance. The sliding window approach not only creates poorest linear approximations but runs the quickest, and it solves the single sample to calculate variance of Ag, as shown in Figure 4.

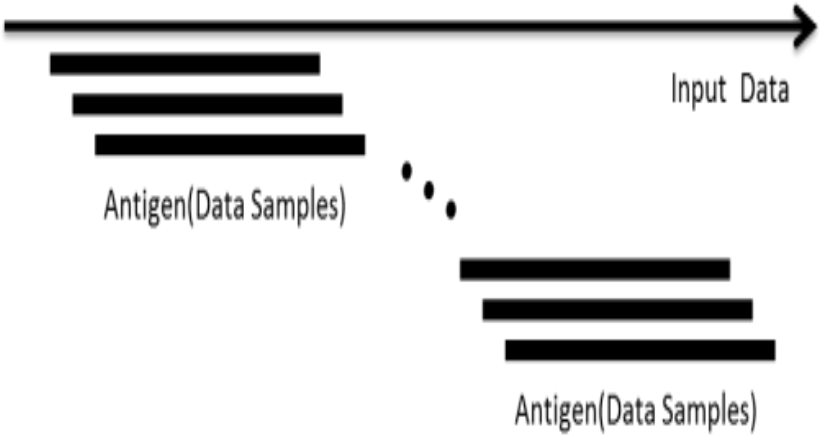


Figure 4. The Sliding Window

Antigen evaluation is defined as formula (2):

$$\mathrm{AE}(\mathrm{Ag(i)}) = \frac{\sum(\mathrm{Ag(i)}-\overline{\mathrm{Ag}})^2}{\mathrm{N}} \quad (2)$$

Where N is the window size of the sliding window, $\mathrm{Ag(i)} = \{Ag_i | Ag_i \in AG \;\; i = k, \cdots, \mathrm{k} + \mathrm{N} - 1\}$, $\overline{\mathrm{Ag}}$ is mean of Ag(i). To improve the running efficiency of the proposed algorithm. $\mathrm{AE}(Ag(i))$ could be calculated incrementally by maintaining the sum and square:

$$\mathrm{AE}(\mathrm{Ag(i)}) = \frac{\sum \mathrm{Ag(i)}^2 - 2\sum \mathrm{Ag(i)}\overline{\mathrm{Ag}} + \sum \overline{\mathrm{Ag}}^2}{\mathrm{N}} \quad (3)$$

Therefore, $\mathrm{AE}(\mathrm{Ag(i)})$ can be written as:

$$\mathrm{AE}(\mathrm{Ag(i)}) = \frac{\mathrm{N}\sum \mathrm{Ag(i)}^2 - (\sum \mathrm{Ag(i)})^2}{\mathrm{N}^2} \quad (4)$$

The pseudo code of antigen evaluation algorithm is described in Algorithm 1:

Algorithm 1. Antigen evaluation algorithm
for k=1 to n do
Create data samples from input data $\mathrm{Ag(i)} = \{Ag_i | Ag_i \in AG \;\; i = k, \cdots, \mathrm{k} + \mathrm{N} - 1\}$
Evaluate the change of antigen' type:
$\mathrm{AE}(Ag(i))$
if antigen' type has changed

Initiate antibody recognition
end if
end for

4.2. Antibody Recognition

As for antibody recognition, distance (or similarity) measure is preferred. Antibody recognition is defined as formula

$$AR(Ab, Ag) = \{f_m(Ab, Ag) | Ab \in AB, Ag \in AG\} \quad (5)$$

Where f_m could be Euclidean Distance, Manhattan Distance, Hamming Distance, *etc*.

The pseudo code of the antibody recognition algorithm is described in Algorithm 2.

Algorithm 2. Antibody recognition algorithm

if antigen' type has changed (AE)
Initiate antibody recognition
Measure the similarity between Ag and Ab
$AR(Ab, Ag)$
end if

As shown in Algorithm 2, antibody recognition is not necessary for every antigen. It is activated to determine whether or not antigen initiates immune response when the type of antigen has changed.

4.3. Antibody Evolution

Inspired by biological immune mechanism, antibody evolution can be described as:

$$R = \{R_c, R_m, R_o\} \quad (6)$$

Where R_c is clonal operator, R_m is mutation operator, R_o is optimization selection operator.

In the immune operators, clone operator is the most commonly used and the number of the clone is also an important problem. In order to make the number of clone within a controllable range, R_c is difined as formula (7):

$$R_c^i = \left\lceil \frac{P_i(NC-(N-1))+(N-1)}{N} \right\rceil \quad (7)$$

$$P_i = \frac{f_{mi}}{\sum_{k=1}^{N} f_{mk}} \quad (8)$$

Where N is the size of AB, C is the best antibody's maximum clone number. According to formula (7): R_c^i is proportional to P_i, and $1 \le R_c^i \le C$.

Mutation operator is another important member in the immune operator. As for R_m, mutation probability is defined as:

$$g = \beta \exp(-\frac{1}{f_m(Ab,Ag)}) \quad (9)$$

Where β has the effect of adjusting the mutation probability, g decreases with the increase of $f_m()$.

According to the application and purpose of the model, R_o can choose different selection criterias, for example, mutated affinity cannot reduce, and antibody concentration is under control, *etc*.

4.4. Fast Immune Recognition Algorithm based on Immune Response

Fast immune recognition algorithm based on immune response (FIRA) is the most important issue of FIRM. The outline of FIRA is illustrated in Figure 5.

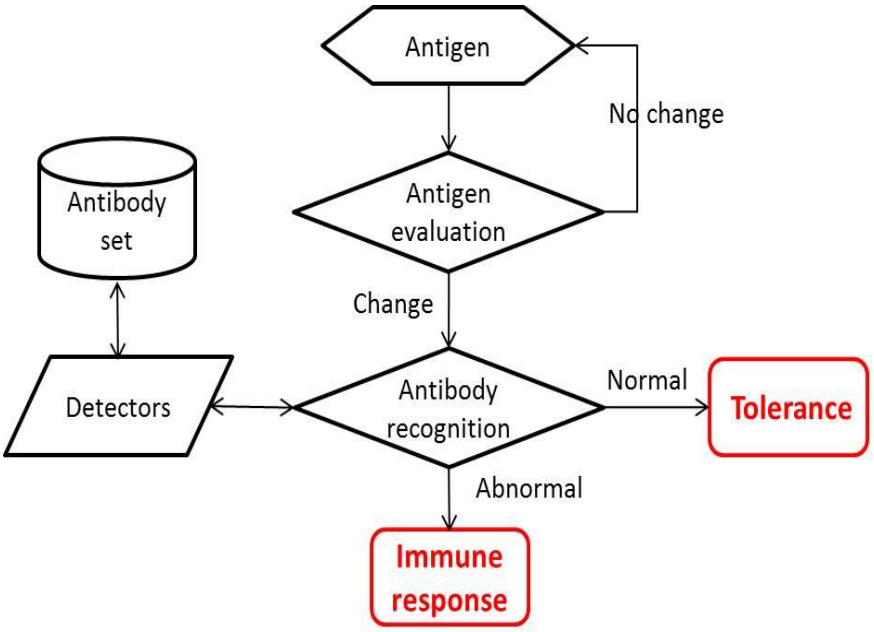


Figure 5. Flowchart of FIRA

Now let's analyze the algorithm complexity of FIRA. Suppose type of Ag has changed at every input data sample, complexity of FIRA is $O(n(N + \text{detector\_number})) = O(n)$ because N and detector_number (the number of the detectors) are constants; suppose Ag has not changed, FIRA is $O(n(N)) = O(n)$ because N is a constant. Therefore FIRA is a linear algorithm.

5. Experiments

In this section, we perform experiments to verify the effectiveness of the proposed approach for anomaly detection.

In order to test the performance, accuracy is taken as measure of effectiveness:

$$\text{accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \tag{10}$$

Where TP represents true positives, TN is true negatives, FP is false positives and FN is false negatives

5.1. Breast Cancer Dataset

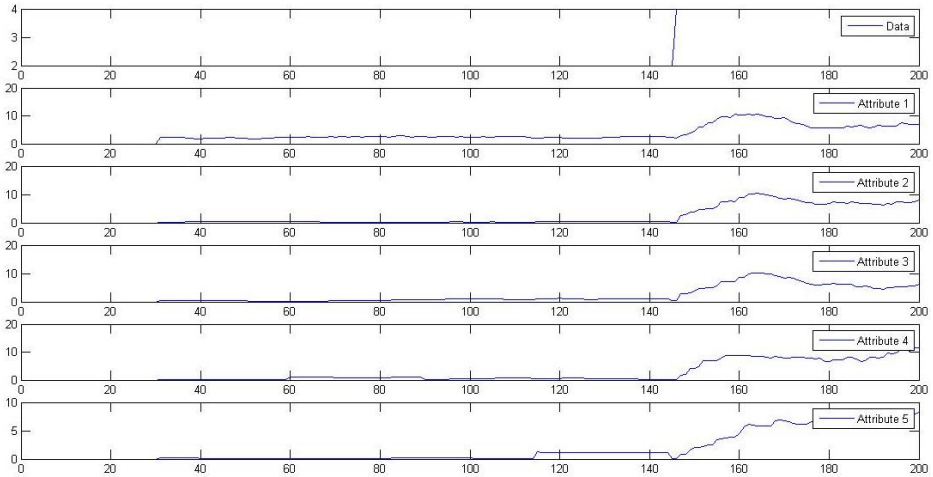
Breast cancer dataset of UCI has been used as a test set for different anomaly detection approaches. Its properties are shown in Table 2.

Table 2. The Properties of Breast Cancer Dataset

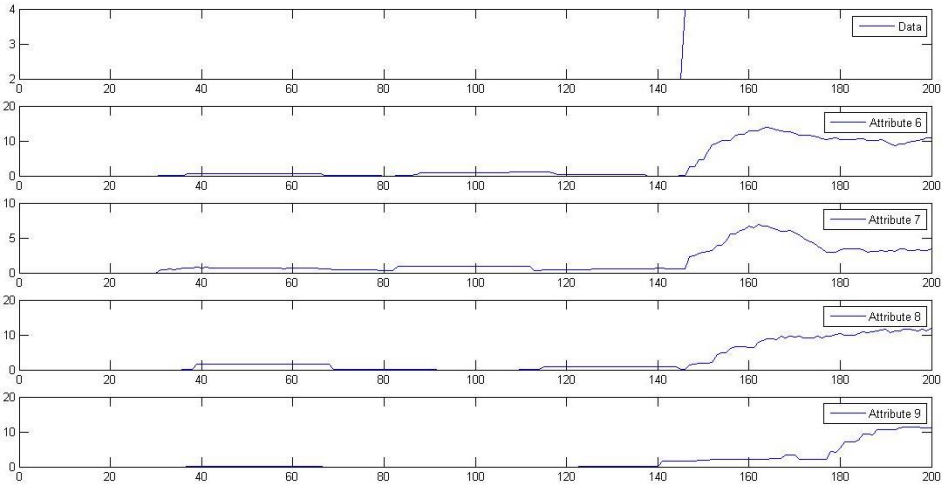
Instances	Attributes	Classification
699	10	2

Each record of breast cancer dataset contains 10 attributes. The first nine dimensions represent the feature of data, which describe the test results of the potential breast cancer patients, and the last dimension describes diagnosis (benign or malignant) of them.

200 recorders data are randomly selected in the following experiments. According to the classification property, "2" represents normal, "4" represents abnormal.



(a) Evaluation of the First 5 Attributes



(b) Evaluation of the Last 4 Attributes

Figure 6. Antigen Evaluation

An advantage of FIRM's antigen evaluation is that there are no parameters. As shown in Figure 6, the value of evaluation from attribute 1 to attribute 8 changes obviously in different classifications. The change of the attribute 9's value cannot represent classification result. Therefore, the attribute 9 can be abandoned. The rest 8 attributes are used to conduct detection. The result is shown in Figure 7.

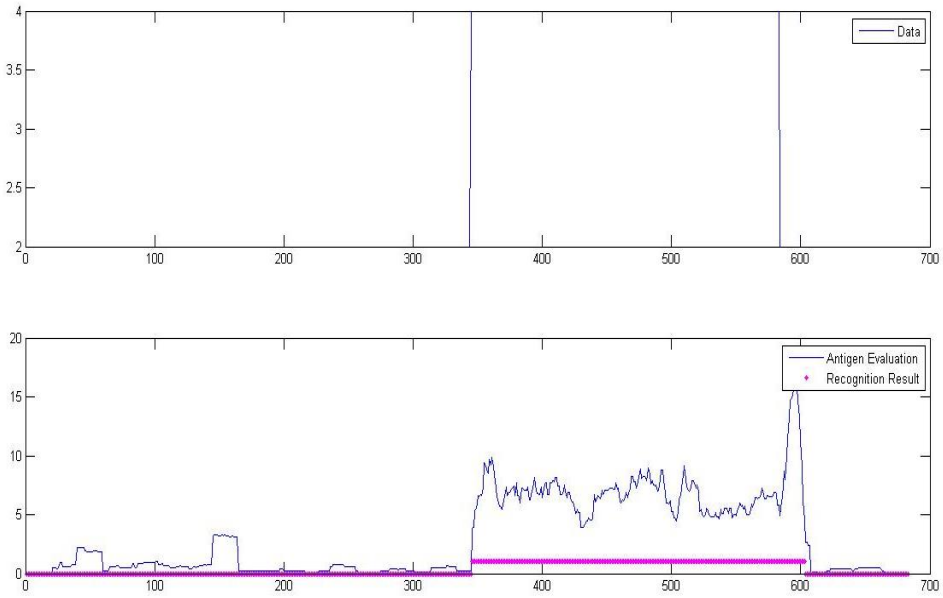


Figure 7. Recognition rResult

"0" represents normal, "1" represents abnormal. As shown in Figure 7, the type of the antigen is evaluated by antigen evaluation. When the type of the antigen changes, antibody recognition is activated to determine whether or not the antigen is normal. Accuracy is 98.6%. Compared with [11], FIRA has a good work at Breast Cancer Dataset.

5.2. Iris Dataset

The properties of iris dataset of UCI are shown in Table 3

Table 3. The Properties of Iris Dataset

Instances	Attributes	Classification
150	5	3

Iris is three classifications dataset including iris-setosa, iris-versicolor and iris-virginica. Take iris-setosa as normal, iris-versicolor and iris-virginica as abnormal.

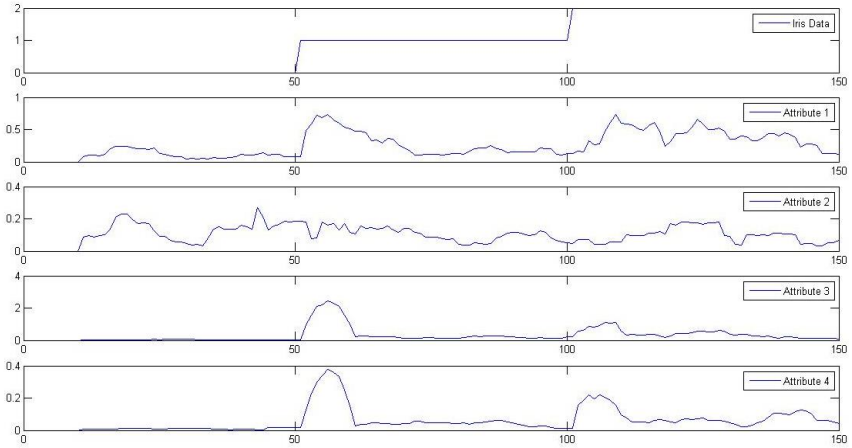


Figure 8. Antigen Evaluation

As shown in Figure 8, the change of the attributes' value cannot reveal classification result clearly. As the dimensions of data are not much, iris is not taken to reduce dimension.

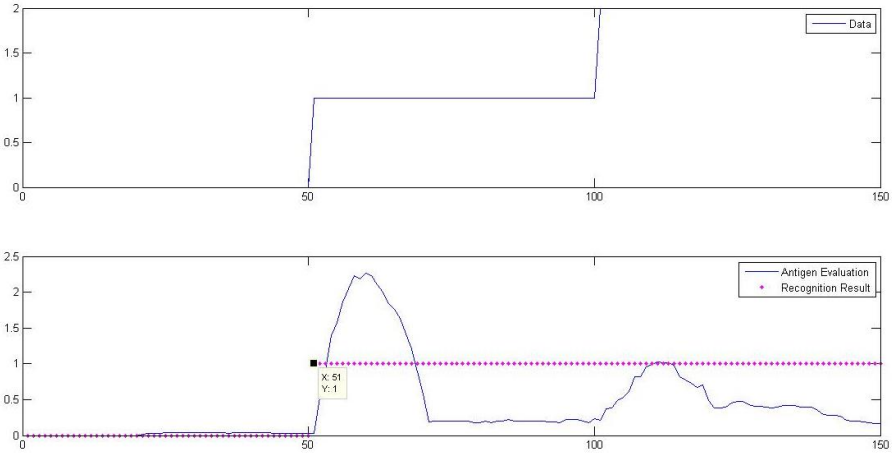


Figure 9. Recognition Result

"0" represents normal, "1" represents abnormal. As shown in Figure 9, the value of antigen evaluation changes violently in the type conversion. Accuracy is 100%. Experiments show the proposed method has high accuracy.

6. Conclusion

In immunology, new models and theories are constantly being proposed. Although immunological recognition models and theories may have different explanations of the immune system, they focus on to what immune system responses. This paper studies the correspondence between immune response and anomaly detection. A fast immune recognition model based on immune response (FIRM) is proposed. The major contributions of FIRM are as follows:

Statistical techniques and sliding window are used to evaluate the type of antigen, which has no parameters. In this way, the most advantage is that it can ignore the differences of different systems.

Another advantage of antigen evaluation is that it can be used for data dimension reduction. The experiments show it works well in dimension reduction.

Compared with traditional AIS, antibody recognition is initiated by the proposed model only when the type of antigen has changed, which improves the efficiency of algorithm. Complexity analysis shows FIRA is a linear algorithm.

For antibody evolution, the number of the clone is always an important problem. FIRM calculates the number of the clone by defining formula, which reduces the randomness of the clone's number.

Acknowledgements

This paper is supported by "Science and Technology Innovation Plan" Major Project of Shanghai Committee of Science and Technology under Grant No.13511504803;Yong University Teachers Training Plan of Shanghai Municipality under Grant No. ZZSD12060; Innovation Fund of Shanghai University under Grant No. SDCX2012017.

References

[1] D. Dasgupta, S. Yu and F. Nino, "Recent advances in artificial immune systems: models and applications", Applied Soft Computing, vol. 11, no. 2, **(2011)**, pp. 1574-1587.
[2] C. A. Laurentys, R. M. Palhares and W. M. Caminhas, "A novel Artificial Immune System for fault behavior detection", Expert Systems with Applications, vol. 38, no. 6, **(2011)**, pp. 6957-6966.
[3] D. Dasgupta and F. Nino, "Immunological computation: theory and applications", CRC Press, **(2008)**.
[4] A. S. Perelson and G. Weisbuch, "Immunology for physicists", Reviews of modern physics, vol. 69, no. 4, **(1997)**, pp. 1219-1263.
[5] M. Vella, M. Roper and S. Terzis, "Danger theory and intrusion detection: possibilities and limitations of the analogy", ICARIS: 9th International Conference on Artificial Immune Systems, Springer: Lecture Notes in Computer Science, **(2010)**, pp. 276-289.
[6] U. Aickelin and S. Cayzer, "The danger theory and its application to artificial immune systems", arXiv preprint arXiv:0801, no. 35-49, **(2008)**.
[7] J. Greensmith, U. Aickelin and S. Cayzer, "Introducing dendritic cells as a novel immune-inspired algorithm for anomaly detection//Artificial Immune Systems", Springer Berlin Heidelberg, **(2005)**, pp. 153-167.
[8] D. Dasgupta, S. Yu and F. Nino, "Recent advances in artificial immune systems: models and applications", Applied Soft Computing, vol. 11, no. 2, **(2011)**, pp. 1574-1587.
[9] V. Chandola, A. Banerjee and V. Kumar, "Anomaly detection: A survey", ACM Computing Surveys (CSUR), vol. 41, no. 3, **(2009)**, pp. 15.
[10] F. A. González and D. Dasgupta, "Anomaly detection using real-valued negative selection", Genetic Programming and Evolvable Machines, vol. 4, no. 4, **(2003)**, pp. 383-403.
[11] G. C. Silva, R. M. Palhares and W. M. Caminhas, "A transitional view of immune inspired techniques for anomaly detection//Intelligent Data Engineering and Automated Learning-IDEAL", **(2012)**, pp. 568-577.
[12] I. Aydin, M. Karakose and E. Akin, "Chaotic-based hybrid negative selection algorithm and its applications in fault and anomaly detection", Expert Systems with Applications, vol. 37, **(2010)**, pp. 5285-5294.
[13] K. J. Lafferty and A. Cunningham, "Aust. J. Exp. Biol. Med.Sci.", vol. 53, no. 27, **(1975)**.
[14] S. Forrest, A. S. Perelson, L. Allen, *et al.*, "Self-nonself discrimination in a computer//Research in Security and Privacy", Proceedings, 1994 IEEE Computer Society Symposium on IEEE, **(1994)**, pp. 202-212.
[15] C. A. Janeway, "Approaching the asymptote? Evolution and revolution in immunology//Cold Spring Harbor symposia on quantitative biology", Cold Spring Harbor Laboratory Press, **(1989)**, pp. 54: 1-13.
[16] S. Yu and D. Dasgupta, "Conserved self pattern recognition algorithm//Artificial Immune Systems", Springer Berlin Heidelberg, **(2008)**, pp. 279-290.
[17] P. Matzinger, "The danger model: a renewed sense of self. Science", vol. 296, no. 5566, **(2002)**, pp. 301-305.
[18] U. Aickelin and S. Cayzer, "The Danger Theory and its Application to Artificial Immune Systems", Research Report HPL-2002-244, **(2002)**; HP Labs, Bristol.
[19] J. Greensmith, U. Aickelin and S. Cayzer, "Detecting danger: The dendritic cell algorithm//Robust Intelligent Systems", **(2008)**, pp. 89-112.

[20] A. Iqbal, J. Greensmith, U. Aickelin and J. Twycross, "Detecting Danger: Applying a Novel Immunological Concept to Intrusion Detection Systems", presented at 6th International Conference in Adaptive Computing in Design and Manufacture, Bristol UK, **(2004)**.
[21] M. A. Maarof, "Danger Theory and Intelligent Data Processing", presented at World Academy of Science, Engineering and Technology, **(2005)**.
[22] E. W. Sun, "Cell death recognition model for the immune system", Medical Hypotheses, vol. 70, **(2008)**, pp. 585–596.
[23] P. A. Miescher, L. Zavota and A. Ossandon, "Autoimmune disorders: a concept of treatment based on mechanisms of disease//Springer Seminars in Immunopathology", Springer-Verlag, vol. 25, no. 1, **(2003)**, pp. S5-S60.
[24] K. M. Abadir, W. Distaso and F. Zikes, "Design-free estimation of large variance matrices", Working paper, Imperial College London, **(2012)**.

Authors

Yuan Tao, received her PhD in Computer Application Technology from Shanghai University in 2011. She is a lecturer in Computation Center at the Shanghai University, China. Her main research interests are in the areas of intelligent information processing, artificial intelligent and pattern recognition.

Min Hu, Received her PhD in Control Theory and Control Engineering from Shanghai University in 2006. She is an associate professor in Information Management at the Shanghai University, China. Her main research interests are in the areas of data mining, risk identification method, artificial intelligent and computer application.