Pattern Recognition Approaches inspired by Artificial Immune System

Aanchal Malhotra
Amity School of Engineering & Technology, Amity University, Noida-201303, India

Abhishek Baheti
Amity School of Engineering & Technology, Amity University, Noida-201303, India

Shilpi Gupta
Amity School of Engineering & Technology, Amity University Noida-201303, India

ABSTRACT

In this paper, we have presented a survey on Pattern Recognition technique using a new computational paradigm of Artificial Immune System. Inspired by the biological immune system, it aims to provide solutions for problems in a vast range of domains using novel computational tools. The main use of AIS in pattern recognition is in the field of data mining. The basic immune theories used to explain how the immune system performs pattern recognition are described and their corresponding computational models are presented. We also present a survey on the applications of AIS in various fields related to pattern recognition.

General Terms

Pattern Recognition, Artificial Immune System, Data Mining, Algorithms.

Keywords

Pattern Recognition; Adaptive Immunity; Artificial Immune System; Negative Selection; Clonal Selection; Immune Networks.

1. INTRODUCTION

The vertebrate immune system is believed to run on one of the most complex biological structures within an organism that protects the body against foreign molecules known as antigens. Its incredible pattern recognition capability distinguishes between foreign cells (non-self or antigen) entering the body and the body cells (self or antibody). Similar to the way the nervous system inspired the development of Artificial Neural Networks (ANN), the immune system has now led to the emergence of Artificial Immune Systems (AIS) as a novel computational intelligence paradigm [14]. "*Artificial Immune Systems (AIS) are adaptive systems, inspired by theoretical immunology and observed immune functions, principles and models, which are applied to problem solving*." [15] They are a class of computationally intelligent systems inspired by the principles and processes of the vertebrate immune system.

Most AIS systems aim at solving complex computational problems from mathematics, engineering, and information technology.

***Correspondence to:** Aanchal Malhotra*
Aanchal Malhotra is at Amity School of Engineering & Technology, Amity University. She is also a Research Intern at TCS Lab, IIT-Madras, India.

2. BIOLOGICAL AND ARTIFICIAL IMMUNE SYSTEMS

As we know, the biological Immune system consists of two components: Innate and Adaptive Immunity. Innate immunity comprises of an immediate, but non-selective protective response. On the other hand, adaptive immunity is the one that enables the immune system to recognize and memorize specific pathogens (previously encountered), and to command stronger attacks each time the registered pathogens are encountered.

In a nutshell, B-cells (origin: Bursa) and T-cells (origin: Thymus) are vital participants in human immune system. The B-cells are the precursors of antibody forming lymphocytes. Upon encountering an antigen, Major Histocompatibilty Complex (MHC) stimulates B-cells by a series of complex chemical pathways. The T-cells undergo cloning and somatic hyper-mutation on sufficient stimulation of B-cells. The antigens are then attacked by killer T Cells and removed from the system [16]. The immune system maintains a memory of the infection by the antigen, as a tool to elicit a quicker and more-efficient immune response in case of subsequent infestations. The biological immune system is a highly parallel, distributed, and adaptive system. It uses learning, memory, and associative retrieval to solve recognition and cataloging tasks [17]. According to De Castro et al [14], the most appealing characteristic of the IS is the presence of *receptor sites* on the immune cell surfaces enabling recognition of limitless range of antigenic patterns. The biological immune system possess the capability to identify relevant patterns, remember past encounters, and to use data mining approaches to construct efficient pattern detectors. Thus the extraordinarily efficient information processing capabilities of the immune system provide important aspects in the field of computation and pave the way for the emergence of Artificial Immune System as a novel computational paradigm.

3. AIS ALGORITHMS' CLASSIFICATION

3.1 Negative Selection

3.1.1 The Biology

The process of deleting auto-immune lymphocytes (anti-self) is termed clonal deletion. It is carried out via a mechanism called negative selection that operates on lymphocytes during their maturation. For T-cells this mainly occurs in the thymus,

which provides an environment rich in self-antigens. Immature T-cells that strongly bind these self-antigens undergo a controlled death (apoptosis). Thus, the T-cells which survive this filtration process should be non-reactive to self-antigens. The property of lymphocytes not to react to the self is called immunological tolerance [15].

3.1.2 The Algorithm

The Negative Selection is one of the mechanisms of the natural immune system that has inspired the developments of most of the existing Artificial Immune systems. In the T-cell maturation process of the immune system, if a T-cell in thymus recognizes any self cell, it is eliminated before deploying for immune functionality. Similarly, the negative selection algorithm generates detector set by eliminating any detector candidate that match elements from a group of self samples [18].

As described in [14] from the perspective of pattern recognition, negative selection provides an alternative approach by storing information about the complement (non self) set of the patterns to be recognised (self). The first negative selection algorithm was proposed by Forrest *et al* (1994) [1] to detect data manipulation caused by a virus in a computer system. In Ayara et al. (2002) [2] the NSMutation algorithm is presented. It introduces somatic hyper mutation, eliminates redundancy and possesses tunable parameters. It consists of three phases: define self-data, generate candidate detector and compare the generated detector with self-data based on affinity threshold. Gonzalez and Cannady (2004) [4] presented a self-adaptive negative selection approach for anomaly detection. It uses self-adaptive techniques for parameter tuning. The main two phases of the algorithm are: generate the initial population and the evolution of the population.

Igawa and Ohashi (2008) [3] proposed a new negative selection algorithm named Artificial Negative Selection Classifier (ANSC) for multi-class classification. It introduces a cutting method to reduce the effect of noise. It combines the negative selection algorithm with clonal selection (section 3.2) mechanism to solve issues that prevent negative selection algorithms from being applied to classification problems. These issues include random searching, overfitting, and incomplete information. Some other researchers proposed negative selection algorithms can be found in Zeng et al. (2007) [5], Xia et al. (2007) [6] and Zhengbing et al. (2008) [7].

The most popular negative selection algorithm is the one developed by Forrest *et al* (1994) [1] and is described in the following section. Given an appropriate problem representation, define the set of patterns to be protected and call it the *self- set* (**P**). Based upon the negative selection algorithm, generate a set of *detectors* (**M**) that will be responsible to identify all elements that do not belong to the self-set, i.e., the non self elements. The negative selection algorithm runs as follows:

1. Generate random candidate elements (**C**) using the same representation adopted;
2. Compare (match) the elements in **C** with the elements in **P**. If a match occurs, i.e., if an element of **P** is recognized by an element of **C**, then discard this element of **C**; else store this element of **C** in the detector set **M**.

After generating the set of detectors (**M**), the next stage of the algorithm consists in monitoring the system for the presence of non self patterns. In this case, assume a set **P*** of patterns to be protected. This set might be composed of the set **P** plus other new patterns, or it can be a completely novel set. For all elements of the detector set, that corresponds to the non self patterns, check if it recognizes (matches) an element of **P*** and, if yes, then a non self pattern was recognized and an action has to be taken. The resulting action of detecting non self varies according to the problem under evaluation.

3.2 Clonal Selection

3.2.1 The Biology

The Clonal Selection Principle is the fundamental features of an immune response to an antigenic stimulus. It establishes the idea; only the cells capable of recognizing the antigen will proliferate. Thus, when a B-cell receptor recognizes a non-self antigen with a certain affinity, it is selected to proliferate and produce antibodies in high volumes. The antibodies are released from the B-cell surface to retaliate the invading non self-antigen. Antigen-antibody complexes operate to eventually eliminate the antigens by a complex biochemical cascade. During reproduction, the B-cell progenies (clones) undergo a hyper mutation process that, together with a strong selective pressure, results in B-cells with antigenic receptors presenting higher affinities for the target antigens. The process of mutation and selection is known as the maturation of the immune response [19]. In addition to differentiation (into antibody producing cells), a section of activated B cells population with high antigenic specificity are selected to become memory cells for long life spans. These memory cells are pre-eminent in future responses to the explicit antigenic pattern, or a similar one [15].

3.2.2 The Algorithm

Clonal Selection algorithms are a class of algorithms inspired by the clonal selection theory of acquired immunity. Several artificial immune algorithms have been developed imitating the clonal selection theory. As illustrated in [15] the two important features of clonal selection relevant from the perspective of computation are:

1. The proliferation rate of each immune cell is proportional to its affinity with the selective antigen: the higher the affinity, more clones are produced.
2. The mutations suffered by the antibody of a B-cell are inversely proportional to the affinity of the antigen it binds.

Utilizing these two features, de Castro and Von Zuben (2000) [8] developed one of the most popular and widely used clonal selection inspired AIS called CLONALG, which has been used to perform the tasks of pattern matching and multi-modal function optimization. When applied to pattern matching, a set of patterns S, to be matched are considered to be antigens. Given a set of patterns to be recognized (**P**), the basic steps of the CLONALG algorithm are as follows:

1. Randomly initialize a population of individuals (**M**);
2. For each pattern of **P**, present it to the population **M** and determine its affinity with each element of the population **M**;
3. Select n_1 of the best highest affinity elements of **M** and generate copies of these individuals proportionally to their affinity with the antigen. The higher the affinity, the higher the number of copies, and vice-versa;
4. Mutate all these copies with a rate proportional to their affinity with the input pattern: the higher the affinity, the smaller the mutation rate, and vice-versa.

5. Add these mutated individuals to the population **M** and re-select n_2 of these maturated (optimized) individuals to be kept as memories of the system;
6. Repeat Steps 2 to 5 until a certain criterion is met, such as a minimum pattern recognition or classification error.

Later in the following years, many improved and advanced versions of CLONALG algorithm to overcome its inherent drawbacks were implemented.

3.3 Immune Networks

3.3.1 The Biology

The defining concept of immune network theory is that any lymphocyte receptor within an organism can be recognized by a subset of the total receptor repertoire. According to the immune network theory, the receptor sites (antigenic determinants) on the surface of one immune cell, named *idiotopes*, can be recognized by receptors on other immune cells. These idiotopes are displayed in and/or around the same portions of the receptors that recognize non-self antigens. It is suggested that the immune cells are capable of recognizing each other, what endows the system with an Eigen behavior that is not dependent on foreign stimulation. Several immunologists have refuted [20] this theory, however its computational aspects are relevant and it has proved itself to be a powerful model for computational systems [14].

3.3.2 The Algorithm

In 1974, Jerne [21] proposed an immune network theory to help explain some of the observed emergent properties of the immune system, such as learning and memory. Recently, the most influential artificial immune network (AIN) models found in the literature are [9] and [10]. An updated version, called AINE [11] uses artificial recognition ball (ARB) to represent a number of similar B-cells (not a single B cell). The network of antibodies is generated by matching the data item by Euclidean distance to antigen or other ARBs. A link between two B-cells is created if the affinity (distance) between two ARBs is below a network affinity threshold (NAT). The results show that the combination of normalizing the stimulation levels of ARBs in the network and the resource allocation mechanism leads to the biasing of AINE towards the strongest pattern in the data set to emerge (Knight2001) [12]. The work presented in [13] makes use of the clonal selection algorithm (CLONALG), described in Section 3.2 to explain how the immune network model responds to non self antigens i.e. becomes activated. The recognition of cell receptors by other cell receptors results in network suppression. This is modeled by eliminating all but one of the self-recognizing cells. Given a set of patterns (**P**) to be recognized, the basic algorithm runs as follows:

1. Randomly initialize the network population;
2. For each antigenic pattern in **P** apply the CLONALG algorithm that will return a set of memory cells (**M***) and their co-ordinates for the current antigen;
3. Determine the affinity (degree of matching) among all the individuals of **M***;
4. Eliminate all but one of the individuals in **M*** whose affinities are greater than a given threshold. The purpose of this process is to eliminate redundancy in the network by suppressing self-recognizing elements;
5. Concatenate the remaining individuals of the previous step with the remaining individuals found for each antigenic pattern presented. This will result in a large population of memory individuals **M**;
6. Determine the affinity of the whole population **M** and suppress all but one of the self-recognizing elements. This will result in a reduced final population of memory cells that recognize and follow the spatial distribution of the antigens.
7. Repeat Steps 2 to 6 until a pre-defined stopping criterion is met, such as a minimum pattern recognition or classification error.

Affinity in this case can be taken to mean the degree of recognition or match, between the elements of the artificial immune system itself (self), and among them and the environment (non self).

4. APPLICATIONS OF ARTIFICIAL IMMUNE SYSTEM

Immune system at a fundamental level may be viewed as a learning system that readily accepts new input patterns of arbitrary length, maintains a database of previously encountered patterns, and reintroduction recognizes learned patterns efficaciously. The various algorithms of Artificial Immune Systems are suitably applicable in different fields. Negative Selection algorithm is used for fault detection, and computer security, thus utilizing self/non self recognition aspect. Similarly, Artificial Immune Network can be used in clustering, classification, data analysis, and data mining. Clonal Selection algorithm is generally used in optimization problems.

Biggest advantage of immune based learning compared to other current approaches is expected to be its ease of adaptation in dynamic environments. Table 1 shows the survey various applications of AIS and their advantages in the field of pattern recognition.

Table 1. Applications of AIS in Pattern Recognition fields

ARTIFICIAL IMMUNE SYSTEM	
APPLICATION	**ADVANTAGE**
1. Image pattern differentiation, categorization, and recognition [24].	Low cost automated monitoring, and object detection application.
2. Intrusion Detection System (IDS) [23].	Ability to learn and generalize the training results gives a possibility to create on their basis the internal information protection systems that can detect unknown computer attacks.

3. Information Compression and data clustering and to solve multi model function optimization problem. (Immune Network Model) [25].	Automatic detection of population size, combination of local with global search, defined convergence criterion, and capability of locating and maintaining stable local optimal solution.
4. Optimization problems (Clonal Selection Algorithm) [8].	It is highly parallel and presents a fine tractability in terms of computational cost and it can reach the optimal solution.
5. Anomaly Detection [22].	Distributed, symmetric and improved quality of checking quality can be traded off against the cost of performing check.
6. Hardware Fault Tolerance [26].	Highly dependable distributed systems with a very high degree of redundancy.

5. CONCLUSION

An overview of the Artificial Immune system along with its biological background has been presented in the paper. The three classes of artificial immune system algorithms to perform pattern recognition: 1) negative selection, 2) clonal selection, and 3) immune network models, have been reviewed. We have also described the role of artificial immune system in the field of pattern recognition. These algorithms can be directly applied to solve various problems in the various fields of pattern recognition, or to complement their potentialities. A survey on its varied application and advantages in various fields has been carried out and presented.

It can be concluded from the review presented that Artificial Immune System approach towards pattern recognition is significantly efficient. There are myriad fields where these algorithms can be applied and useful results derived. Besides, we have seen that the AIS approach when combined with other computational paradigms, such as neural network, genetic algorithm, and fuzzy logic yields even more beneficial results. This review enlightens us with the hidden potential that lies in the field of biologically inspired AIS and thus gives us a motivation to explore it further.

6. ACKNOWLEDGEMENTS

We are grateful to Abhishek Sharma (Manipal College of Pharmaceutical Sciences, Manipal University) for explaining and editing the biology part, thereby helping us realize this work.

7. REFERENCES

[1] Forrest, S., Somayaji, A & Ackley, D. H. 1997. Building Diversity Computer Systems. In proceedings of the 6th workshop on Hot Topics in Operating Systems, 66-72.

[2] Ayara, M., Timmis, J., De Lemos, R., De Castro, L., Duncan, R. 2002. Negative Selection: How to Generate Detectors. In proceedings of 1st ICARIS.

[3] Igawa K. and Ohashi H. 2008. A negative selection algorithm for classification and reduction of the noise effect. Applied Soft Computing. Vol. 9(1), 431-438.

[4] Gonzalez L. and Cannady J. 2004. A Self-Adaptive Negative Selection Approach for Anomaly Detection. Congress on Evolutionary Computation. Vol. 2, 1561-1568.

[5] Zeng J., Li T., Liu X., Liu C., Peng L. and Sun F. 2007. A Feedback Negative Selection Algorithm to Anomaly Detection. Third International Conference on Natural Computation. Vol. 3 (Aug 2007), 604-608.

[6] Xia F., Zhu Y., and Gao Y. 2007. Shape-space based negative selection algorithm and its application on power transformer fault diagnosis. In Proceedings of the IEEE International Conference on Robotics and Biomimetics, (Dec 2007), 2149 - 2154.

[7] Zhengbing H., Ji Z., and Ping M. 2008. A Novel Anomaly Detection Algorithm Based on Real-Valued Negative Selection System. Workshop on Knowledge Discovery and Data Mining, (Jan 2008), 499-502.

[8] De Castro, L. N. & Von Zuben, F. J. 2000. The Clonal Selection Algorithm with Engineering Applications. In proceedings of GECCO'00, (July 2000), 36-37.

[9] L.N. De Castro & F.J. Von Zuben. 2001. aiNet: An Artificial Immune Network for Data Analysis. In Data Mining: A Heuristic Approach, chapter 12, USA: Idea Group Publishing, pp. 231–259.

[10] Timmis, J. 2000. Artificial Immune Systems: A novel data analysis technique inspired by the immune network theory. Ph.D. Dissertation, Dept. of Computer Science, University of Wales. Available from: link (Assessed on: April 26, 2012).

[11] Timmis, J., Neal, M., Hunt, J. 2000. An artificial immune system for data analysis. Biosystems. Vol. 55(1-3), 143-150.

[12] Knight, T and Timmis, J. 2001. AINE: An immunological approach to data mining. In Proceedings of IEEE ICDM. ISBN: 0-7695-1119-8.

[13] De Castro, L. N., Von Zuben, F. J., Deus, G.A. The Construction of a Boolean Competitive Neural Network Using Ideas from Immunology. Neurocomputing. Vol. 50(Jan 2003), 51-85.

[14] De Castro, L.N, Timmis, J. 2002. Artificial immune systems: a novel paradigm for pattern recognition. In Artificial Neural Networks in Pattern Recognition. University of Paisley, UK, pp. 67-84, 2002.

[15] De Castro, L.N, Timmis, J. 2002. Artificial Immune Systems: A New Computational Intelligence Approach. Berlin, Germany: Springer-Verlag.

[16] Timmis, J, Neal, M. 2001. A resource limited artificial immune system for data analysis. Knowledge-Based Systems. Vol. 14(324), 121-130.

[17] Sumathi, S., Paneerselvam, S. 2010. Computational Intelligence Paradigms Theory & Applications using MATLAB. CRC Press 1 edition.

[18] Al-Enezi, J.R., Abbod, M.F. and Alsharhan, S. Artificial immune system- models, algorithms and applications. International Journal of Research and Reviews in Applied Sciences (IJRRAS). Vol. 2, 118-131.

[19] Nossal, G. J. V. 1993. The Molecular and Cellular Basis of Affinity Maturation in the Antibody Response. Cell. Vol. 68(1), 1-2.

[20] Langman, R. E. & Cohn, M. 1986. The 'complete' idiotype network is an absurd immune system. Immunology Today. Vol. 7(4), 100-101.

[21] Jerne, N. K. 1974. Towards a Network Theory of the Immune System. Ann Immunol (Inst. Pasteur) 125C, pp. 373-389.

[22] Forrest, S., Perleson, A., Allen, L., and Cherukuri, R. 1994. Self-Nonself discrimination in a computer. In Proceedings of IEEE Computer Security Symposium on Research in Security and Privacy. May 1994, 202–212.

[23] Kim J., Bentley P., Aickelin U., Greensmith J., Tedesco G., Twycross J. 2007. Immune system approaches to intrusion detection—a review. NATURAL COMPUTING. Vol. 6(4), 413-466.

[24] Wang H., Peng D., Wang W., Sharif H., Wegiel J., Nguyen D., Bowne R., Backhaus C. Artificial Immune System based image pattern recognition in energy efficient Wireless Multimedia Sensor Networks. Military Communications Conference, MILCOM 2008. IEEE, 1–7.

[25] De Castro, L.N., and Timmis, J. 2002. An artificial immune network for multimodal optimization. In proceedings of the 2002 Congress on Evolutionary Computation. Vol. 1, 699–704.

[26] Bradly, D. W. & Tyrrell, A. M. 2000. Immunotronics: Hardware Fault Tolerance Inspired by the Immune System. Lecture Notes in Computer Science, 1801, pp. 11-20.