Expert Systems with Applications 37 (2010) 5285–5294

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Chaotic-based hybrid negative selection algorithm and its applications in fault and anomaly detection

Ilhan Aydin *, Mehmet Karakose, Erhan Akin

Computer Engineering Department, Firat University, 23119 Elazig, Turkey

ARTICLE INFO

Keywords:
Artificial immune system
Negative selection
K-nearest neighbor
Anomaly and fault detection

ABSTRACT

This paper proposes a new negative selection algorithm method that uses chaotic maps for parameter selection. This has been done by using of chaotic number generators each time a random number is needed by the original negative selection for mutation and generation of initial population. The coverage of negative selection algorithm has been improved by using chaotic maps. The proposed algorithm utilizes from clonal selection to obtain optimal non-overlapping detectors. In many anomaly or fault detection systems, training data don't represent all normal data and self/non-self space often varies over the time. In the testing stage, when any test data cannot be detected by any self or non-self detector, the nearest detectors are found by K-Nearest Neighbor (K-NN) method and the nearest detector is mutated as a new detector to detect this new sample. Proposed chaotic-based hybrid negative selection algorithm (CHNSA) has been analyzed in the broken rotor bar fault detection and Fisher Iris datasets.

1. Introduction

Artificial immune system (AIS) is an emerging kind of soft computing method inspired by the principles and processes of the human immune system (de Castro & Timmis, 2003). From the computational perspective, the most important features of AIS are learning, diversity, maintenance and memory. Three main components of AIS are negative selection algorithm (NSA), clone selection and immune network model (Igawa & Ohashi, 2009). Application domains of AIS are various and include anomaly and fault detection (Aydin, Karakose, & Akin, 2008; Forrest, Perelson, Allen, & Cherukuri, 1994), classification (Leung, Cheong, & Cheong, 2007), pattern recognition and optimization problems (de Castro & Zuben, 2002).

AIS can offer robust information processing capabilities for solving complex problems. It has learning and adaptability features (Leung et al., 2007). Each AIS component is applied to a different engineering or scientific problem. NSA is appropriate for anomaly and fault detection problems. Clonal selection paradigm has been applied to many engineering tasks such as clustering, data mining and optimization. Immune network model is applied to pattern recognition problems. Although the whole field of AIS is in its early stage, it has already used in many practical applications. However, the new algorithms have been developed; they have several problems such as scalability, parameter tuning and efficiency.

NSA is based on the principles of self/non-self discrimination in the natural immune systems. It was firstly developed to detect computer viruses (Forrest et al., 1994). Recently, NSA has been widely applied in such engineering problems as aircraft fault detection (Dasgupta, KrishnaKumar, Wong, & Berry, 2004), induction motor fault detection (Aydin et al., 2008; Gan, Zhao, & Chow, 2009; Xinmin, Baoxiang, & Yong, 2007) and information security (Dasgupta & Gonzalez, 2002). In original NSA, candidate detectors are generated in a random manner, and then exposes to a negative censoring mechanism. Only the qualified detectors that don't match any self sample are inserted to the detector set. Unfortunately, these randomly generated detectors have three problems. Firstly, these randomly generated detectors cannot be guaranteed to cover the non-self space. Secondly, many detectors generated in a random manner are similar. In other words, some non-self data are detected by two or more detectors. Thirdly, self and non-self space often varies over time. Particle swarm optimization was proposed to optimize detectors of negative selection algorithm (Gao, Ovaska, & Wang, 2007). This method uses a multi-phase particle swarm optimization and anti-collision technique. Each subswarm represents a detector group, which consists of a given number of detectors. But their methods uses fixed radius for each detectors. The real valued negative selection was proposed in order to coverage the abnormal space (Zhou & Dipankar, 2004). They were introduced a new scheme of detector generation and matching mechanism for negative selection algorithm with variable

* Corresponding author. Tel.: +90 424 237 00 00x6308; fax: +90 424 241 55.
E-mail addresses: iaydin@firat.edu.tr (I. Aydin), mkarakose@firat.edu.tr (M. Karakose), eakin@firat.edu.tr (E. Akin).

0957-4174/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2010.01.011

properties. The negative selection was applied to multi-class classification problems (Igawa & Ohashi, 2009). Detectors are generated for each self and non-self data set using clonal selection principles and a new data reduction technique is proposed to reduce the noise effect. In their algorithm, when a test sample cannot be recognized by any detectors, each detector radius is enlarged by a fixed multiplier until this sample is recognized by at least one training sample. The problem is that more than one enlarged detectors covers the self space. In addition, many detectors overlap the others. Therefore, one test data can be recognized by multiple detectors and produced detectors are similar. The common goals in various negative selection algorithms are to minimize number of detectors, maximize coverage of non-self space and generate efficient detector set.

Recently, chaotic sequences have been adopted instead of random sequences and very interesting and somewhat good results have been shown in many applications such as secure transmission, telecommunication and cryptography (Suneel, 2006; Wong, Man, Li, & Liao, 2005); nonlinear circuits (Arena, Caponetto, Fortuna, Rizzo, & La Rosa, 2000), and image processing (Gao, Zhang, Liang, & Li, 2006). Chaos is a bounded unstable dynamic behavior that exhibits sensitive dependence on initial conditions and includes infinite unstable periodic motions in nonlinear systems. Many chaotic maps in the literature possess certainty, ergodicity and the stochastic property (Alatas & Akin, 2008).

In this paper several of the aforementioned problems are discussed when the AIS is applied to the problems of fault and anomaly detection. A new solution is proposed in the form of a novel negative selection algorithm, chaotic-based hybrid negative selection, to overcome some of the limitations of the previous approaches. Proposed method offers an important paradigm shift in the clonal selection idea and chaotic map construction, as the initial population generation and mutation in negative selection is redefined and the well known K-NN method is incorporated in the proposed solution as an adaptive testing stage. The proposed method provides two advantages, which are minimum overlapping of detectors and adaptability of detectors for the test data. The selected affinity function guarantees minimum overlapping between detectors and maximum coverage of non-self data. In the testing stage, nearest detector is mutated to detect unrecognized self/non-self sample using K-NN method. Therefore, adaptation and learning principles of detectors are ensured for unseen samples before.

The rest of paper is organized as follows. First, in Section 2, we will briefly discuss the main principles of the artificial immune system. Section 3 explains the generation and detection steps of our new chaotic based hybrid negative selection algorithm. Section 4 provides details of the tests performed and results obtained, and Section 5 concludes our paper.

2. Artificial immune system

Immune system has a complex structure with capability of learning, memory and novelty detection characteristics. This system protects our body against foreign cells called antigen. AIS developed in 1990 as a new branch in soft computing (de Castro & Timmis, 2003). The popular algorithms of AIS are negative selection, clonal selection and immune network model (Dasgupta, 2006). In this study, we have used negative and clonal selection algorithm. So, we discuss the essential principles of these two algorithms.

2.1. Clonal selection algorithm

Any pathogen that can be recognized by the adaptive immune system is known as antigen. When an immune system is exposed to an antigen, the antibodies spread on surfaces of B lymphocytes are produced (Zhou & Dipankar, 2004). Each B-cell is specific to a given antigen. Antibodies are able to recognize certain type of antigens. If a new antigen enters the body, the immune system clones the most stimulated lymphocytes. In real immune system, the level of binding between the antigen and antibody is performed on the molecular level. In AIS, this binding can be modeled by means of shape space (Gong, Jiao, Zhang, & Du, 2009). Each antigen and antibody can be represented as a vector. Clonal selection algorithm can be interpreted as an evolutionary method with three major principles. These principles are dynamically adjustable population, explanation and exploration of search space and natural selection (de Castro & Zuben, 2002). In clonal selection, cells that recognize the antigen proliferate. The evolution of B-cells inspired many different algorithmic implementations. A general scheme for clonal selection is CLONALG (de Castro & Zuben, 2002). This algorithm is applied to pattern recognition and optimization problems. It utilizes clonal selection and affinity maturation. The main steps of this algorithm are given in Fig. 1.

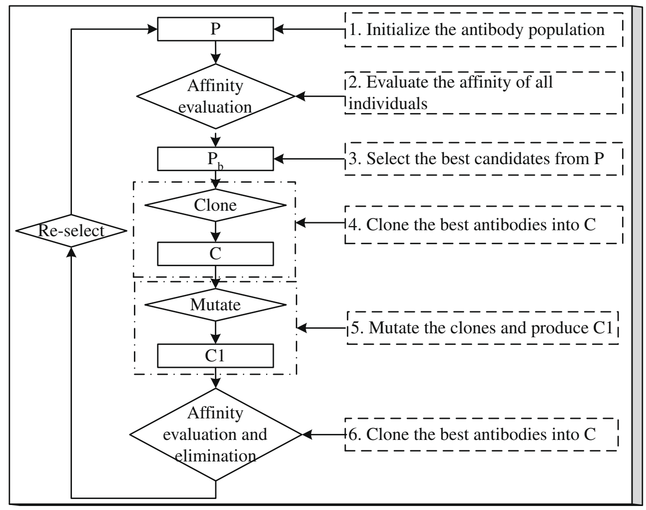

Fig. 1. Steps of clonal selection method.

Clonal selection algorithm has a unique mutation operator. The mutation rate of an individual is inversely proportional to its fitness by means of different mutation variations. In other words, the better fitness the antibody has, the less it may expose to the mutation. The similarity between candidates is minimized using elimination step (Yoo & Hajela, 1999). With such a diverse antibody population, clonal selection method can avoid being caught to local minima. Thus, we are going to utilize the clonal selection method to obtain optimum NSA detectors.

2.2. Negative selection algorithm

One of immune system function is to recognize all cells and distinguish them as self or non-self (Dasgupta, 2006). This function is achieved by two different type's lymphocyte: B and T cells. These cells are produced in the bone marrow. T cells are passed a censoring mechanism in thymus called as a negative selection. Only those who do not match the self cells of body are allowed to leave and others are destroyed there (Igawa & Ohashi, 2009). These maturated T cells protects the body from foreign antigens.

Inspired by biological negative selection, an algorithm was proposed to detect computer viruses (Forrest et al., 1994). This algorithm is based on principles of self/non-self discrimination. Fig. 2a shows the concept of self and non-self space for real valued