Applied Soft Computing 31 (2015) 210–222

Contents lists available at ScienceDirect

Applied Soft Computing

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

On the evolution of ellipsoidal recognition regions in Artificial Immune Systems

Seral Ozsen [1], Cuneyt Yucelbas *,[1]

Department of Electrical and Electronics Engineering, Selcuk University, Konya, Turkey

ARTICLE INFO

Article history:
Received 15 November 2013
Received in revised form 29 January 2015
Accepted 1 March 2015
Available online 16 March 2015

Keywords:
Classification
Artificial Immune Systems
Nonlinear classification
Ellipsoidal recognition regions
Clonal selection principle

ABSTRACT

Using different shapes of recognition regions in Artificial Immune Systems (AIS) are not a new issue. Especially, ellipsoidal shapes seem to be more intriguing as they have also been used very effectively in other shape space-based classification methods. Some studies have done in AIS through generating ellipsoidal detectors but they are restricted in their detector generating scheme – Genetic Algorithms (GA). In this study, an AIS was developed with ellipsoidal recognition regions by inspiring from the clonal selection principle and an effective search procedure for ellipsoidal regions was applied. Performance evaluation tests were conducted as well as application results on some real-world classification problems taken from UCI machine learning repository were obtained. Comparison with GA was also done in some of these problems. Very effective and comparatively good classification ratios were recorded.

1. Introduction

Since its beginning, shape space representation scheme has taken a general acceptance in Artificial Immune Systems (AIS) community. In many AIS based algorithms, system units which are usually called as Antibodies (*Abs*) have some recognition regions and any input datum in an *Ab*'s region is recognized by these *Abs*. Thus, input space should be carefully covered by these *Abs*. So far, except from some studies [1–4], spherically shaped recognition regions – this is why a term of recognition ball is used – have been utilized. In algorithms which use spherical recognition regions, a threshold that is equal to the radius of ball should be passed to enter the recognition region of an *Ab* and this threshold is same in all directions. However, two data points can be very near in one direction whereas they are far from each other in another direction. So, for an effective algorithm, threshold may be different with respect to the direction and this is only possible by using ellipsoidal recognition regions.

Classifying data with ellipsoidal detectors is not a new finding. Some studies used this issue in classification and clustering problems. For example in [5], authors generated Ellipsoidal Adaptive Resonance Theory (E-ART) and Ellipsoidal ART-MAP (E-ART-MAP). They concluded that, depending on the problem, E-ART and E-ART-MAP can be good classifiers with respect to their fuzzy counterparts. In another study, minimum volume ellipsoids (MVE) covering data in a class were found and Hopfield Neural Network was utilized to find these ellipsoids [6]. Authors of [7] proposed MVE clustering as an alternative clustering technique to k-means for data clusters with ellipsoidal shapes. They saw that very effective clustering performances were obtained with ellipsoidal k-means and it is worth to continue studying. Many other studies were conducted related with MVE [8–12]. A similar classification method which uses ellipsoids is the study of [13,14]. In their study, authors developed an algorithm to find best ellipsoidal regions covering the input space. From the findings on some benchmark data, their system can said to be good and comparable with state-of-art works. Authors of [15] presented an effective stream clustering algorithm called Hyper-Ellipsoidal Clustering for Evolving data Stream (HECES) by making a few changes in the recently proposed Hyperellipsoidal Clustering for Resource-Constrained Environments (HyCARCE) algorithm [16], which is a strong clustering technique, in particular designed for low dimensional static data. Hsiao et al. in [17] presented a neural network based on the Ellipsoidal Function Modulated Adaptive Resonance Theory (ART) (EFM-ART). In [18], a novel method was introduced to diagnose power transformer faults based on Ellipsoidal Basis Function (EBF) neural network and this method was compared with Radial Basis Function (RBF) neural network. In another study, a convex quadratic programming representable Minimum Mahalanobis Enclosing Ellipsoid (QP-MMEE) was presented for generally unbalanced dataset classification [19]. Forghani et al. [20] investigated an Extended Support Vector Data Description (ESVDD) which describes data by using a hyper-ellipse and as a result, ESVDD can represent data better than SVDD in the input space.

Similar to the above studies conducted in machine learning area, some researchers in AIS field have also used the idea of ellipsoidal recognition regions in their algorithms. In their study, authors of [1,2] have used Genetic Algorithm (GA) to evolve ellipsoidal detectors in negative selection algorithm. However they used negative selection as an inspiration source from the immunology, most of the work is done by GAs in that ellipsoidal recognition regions were evolved by GA. So, the origin of these studies can be regarded to GAs more than immune system. Some other studies can also be cited here like them but they are all the same with regard to their origin – GA finds best detectors [3,4]. In the study of [3], two synthetic datasets (star and multi-cluster) were used with four types of detector shapes which are hyper-ellipses, hyper-rectangles, hyper-spheres and mixed shapes. According to ROC curves of this study, different results were obtained for each detector shape, and low error rates were produced by mixed shaped detectors as compared to other single shaped detectors in both datasets. In the study of [4] however, 2-dimensional synthetic data was used with six types of shapes which are cross, triangle, circle,

* Corresponding author at: Department of Electrical and Electronics Engineering, Selcuk University, 42072 Konya, Turkey. Tel.: +90 332 223 3712; fax: +90 332 241 0635.

E-mail addresses: seral@selcuk.edu.tr (S. Ozsen), cyucelbas@selcuk.edu.tr (C. Yucelbas).

[1] Tel.: +90 332 2232035; fax: +90 332 2410635.

http://dx.doi.org/10.1016/j.asoc.2015.03.014
1568-4946/

stripe, intersection and pentagram. Hart in [21] presented a study to improve and expand the work of [22]. As a result, choosing recognition area can affect the elimination of antigen and the memory capacity of emergent networks. In work of [22], they mentioned the effects of changeable forms of the recognition area of cells in an idiotypic network simulation. Stibor et al. [23] investigated the behavior of the negative selection algorithm on artificial datasets by different-sized detectors. Classification performances of negative selection, positive selection and statistical anomaly detection techniques were analyzed on a high-dimensional KDD (Knowledge Discovery and Data Mining) dataset.

Whereas GA can be an effective search technique to find optimum ellipsoidal shapes, it is a generic search procedure with random mutation and recombination procedures. So, finding optimum ellipsoids can take time and sometimes ellipsoids which are indeed not optimum but seem to be locally optimum can be found (phenomenon called as catching to local optimum). Because of these negative points in GA, we developed an AIS algorithm that uses ellipsoidal recognition regions which are evolved with clonal selection principle in AIS. To fasten the search process, directed mutations depending on affinities are proposed in changing ellipsoidal shapes. At first, performance of developed system was evaluated and compared with GA on some artificially generated data. Then, comparisons with state-of-art works and GA were done on some real world classification problems. These problems are Pima Indians Diabetes Disease classification problem, Statlog Heart Disease classification problem and BUPA Liver Disorders classification problem whose datasets were taken from the UCI machine learning repository. Besides of these problems, the proposed system was also run for 5 more datasets which were again taken from UCI. Results were given in tabulated form.

2. Development of ellipsoidal clonal selection algorithm

2.1. Mathematical basis for ellipsoids

2.1.1. Definition of ellipsoids in N-dimensional space

Before giving the details of the developed Ellipsoidal AIS system, preliminary information related with ellipsoidal shapes in higher dimensional space are given in the following.

The equation for two-dimensional ellipse is given in the following [24]:

$$\frac{(x-x_0)^2}{a^2}+\frac{(y-y_0)^2}{b^2}=1 \tag{1}$$

where (x_0, y_0) are the center points and a and b are the lengths of semi-axes of x and y respectively. The following equation shows a matrix formulation of Eq. (1) [25]:

$$(x-w)^T \mathrm{V}\Lambda \mathrm{V}^T (x-w)=1 \tag{2}$$

Eq. (2) gives the same formula in (1) if,

$$x=\begin{bmatrix} x \\ y \end{bmatrix},\ \Lambda=\begin{bmatrix} 1/a^2 & 0 \\ 0 & 1/b^2 \end{bmatrix},\ w=\begin{bmatrix} x_0 \\ y_0 \end{bmatrix} \text{ and } \mathrm{V}=\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

If, $\mathrm{V}\Lambda\mathrm{V}^T$ in Eq. (2) is represented with $\sum$, then general form of an ellipsoid for n dimension can be written as:

$$(x-w)^T \sum (x-w)=1 \tag{3}$$

where w is a $n\times 1$ vector representing the center of the ellipsoid and $\sum$ is a real symmetric positive-definite $n\times n$ matrix. Here, V is a $n\times n$ matrix whose columns are orthonormal eigenvectors of $\sum$ and Λ is a $n\times n$ diagonal matrix whose entries are eigenvalues associated with the eigenvectors in V. The ith column of matrix V stands for the orientation of the ellipsoid in ith dimension. Besides, Λ defines the lengths of ellipsoid's semi-axes as the following:

$$\ell_i=\frac{1}{\sqrt{\Lambda_{i,i}}} \tag{4}$$

where ℓ_i is the length of the ith semi-axes [25].

Changing the orientation of the semi-axes means rotating the ellipsoid. A rotation in n-space is defined by a $n\times n$ orthonormal matrix. V is an orthonormal matrix, and defines this rotation. If $x-w$ is a point on the surface of some ellipsoid, then $\mathrm{V}(x-w)$ is a point on the surface of an ellipsoid that has been rotated by V. This is shown with the substitution $(x-w)\rightarrow \mathrm{V}(x-w)$, leading to:

$$\mathrm{V}(x-w)^T \mathrm{V}\Lambda\mathrm{V}^T \mathrm{V}(x-w)=1 \tag{5}$$

This equation can be re-arranged as:

$$(\mathrm{V}^T\mathrm{V}(x-w))^T \Lambda \mathrm{V}^T \mathrm{V}(x-w)=1 \tag{6}$$

Since V is orthonormal, $\mathrm{V}^T\mathrm{V} = \mathrm{I}$. Thus, Eq. (6) simplifies to Eq. (2). Rotation preserves the relative positions of points on the ellipsoid. Hence, if point p is on a semi-axis in the un-rotated ellipsoid, then Vp is on the rotated ellipsoid. Therefore, the orientation of the semi-axes is the columns of V [25].

2.1.2. Volume of ellipsoid

Tee showed that the volume of an n dimensional ellipsoid is calculated by the Eq. (7) [26]:

$$\text{Volume}=\Omega_n\prod_{i=1}^{n}\frac{1}{\sqrt{\Lambda_{j,i}}} \tag{7}$$

where $\Lambda_{i,i}$ is $1/(\ell_i)^2$ in which ℓ_i is the length of ith semi-axes. Ω_n is the volume of an n-dimensional unit hyper-sphere. Smith and Vamanamurthy [27] show that the volume of an n-dimensional unit hypersphere is calculated by the following equation:

$$\Omega_n=\frac{\pi^{n/2}}{\Gamma(1+(1/2)n)} \tag{8}$$

Here in Eq. (8), $\Gamma(\,)$ is the Gamma function. The $\Gamma(\,)$ function is a mathematical extension of the factorial function from positive integers to real numbers [28].

To determine whether a p point is inside the ellipsoid or not, the following squared Mahalanobis distance can be used [29]:

$$(x-w)^T \sum (x-w) < 1 \tag{9}$$

If the criterion in Eq. (9) holds, the point p is inside the ellipsoid.

2.1.3. Gram-Schmidt orthonormalization

To obtain an orthonormal V matrix from a randomly generated matrix, Gram-Schmidt orthonormalization technique is utilized [30].

Let $\{x_1, \ldots\ldots, x_n\}$ be a set of n linearly independent vectors and let $\{y_1, \ldots\ldots, y_n\}$ be the orthogonal set of vectors to be determined [30]. Then,

$$\begin{aligned}
y_1 &= x_1 \\
y_2 &= x_2-\frac{\langle x_2, y_1\rangle}{\langle y_1, y_1\rangle}y_1 \\
y_3 &= x_3-\frac{\langle x_3, y_2\rangle}{\langle y_2, y_2\rangle}y_2-\frac{\langle x_3, y_1\rangle}{\langle y_1, y_1\rangle}y_1 \\
&\ \ \vdots \\
y_k &= x_k-\sum_{i=1}^{k-1}\frac{\langle x_k, y_i\rangle}{\langle y_i, y_i\rangle}y_i
\end{aligned} \tag{10}$$

After obtaining $\{y_1, \ldots\ldots, y_n\}$ orthogonal set by using Eq. (10), each vector in this set is divided to its length to obtain the orthonormal set of $\{z_1, \ldots\ldots, z_n\}$:

$$z_i=\frac{y_i}{\left\|y_i\right\|} \tag{11}$$