Fuzzy Clustering of ECG Beats Using a New Metaheuristic Approach

Berat Doğan[1], Tamer Ölmez[1]

[1]Istanbul Technical Universtiy, Department of Electronics and Communication Engineering, Istanbul, Turkey

{bdogan, olmezt}@itu.edu.tr

Abstract. This study proposes a new single-solution based metaheuristic, namely the Vortex Search algorithm (VS), for fuzzy clustering of ECG beats. The newly proposed metaheuristic is quite simple and highly competitive when compared to the population-based metaheuristics. In order to study the performance of the proposed method a number of experiments are performed over a dataset which is created by using the records selected from MIT-BIH arrhythmia database. The selected records includes six type of beats, namely, Normal Beat (N), Premature Ventricular Contraction (PVC), Fusion of Ventricular and Normal Beat (F), Atrial Premature Beat (A), Right Bundle Branch Block Beat (R) and Fusion of Paced and Normal Beat (f). The records are first preprocessed and then four morphological features are extracted for each beat type to form the training and test sets. By using the newly proposed method, fuzzy cluster centers of the training set is found. By using these clusters' centers a supervised classification method is then classified the test set to evaluate the clustering performance of the method. The results are compared to the fuzzy c-means algorithm (FCM), fuzzy c-means algorithm with particle swarm optimization (FCM-PSO2011) and fuzzy c-means algorithm with artificial bee colony (FCM-ABC). It is shown that, in spite of its simplicity, the newly proposed metaheuristic with fuzzy c-means algorithm (FCM-VS) is highly competitive and performs quite well when compared to FCM-PSO2011 and FCM-ABC methods.

Keywords: Fuzzy clustering, metaheuristics, particle swarm optimization, artificial bee colony, ECG, arrhythmia

1   Introduction

Clustering is an unsupervised learning method that assigns a given set of objects into disjoint groups or clusters based on a similarity criteria. The similarity or dissimilarity measure is usually calculated by a distance function. Thus, after the clustering process, similar patterns are grouped in the same cluster and dissimilar patterns are grouped in different clusters.

Perhaps the most popular clustering algorithm is the hard C-means (or k-means) algorithm in which a data point is assigned to exactly one cluster. If the clusters are

well separated in the feature space this does not lead a problem. However, for most of the real world problems the clusters overlap each other and thus, a data point can be in a location which is almost equally distant from two or more cluster centers. In hard C-means algorithm, by forcing such a point to exactly one cluster, the similarity of this point to other clusters is totally ignored. To avoid this problem fuzzy clustering methods are proposed. In fuzzy clustering methods a data point can belong to more than one cluster with different degrees of membership.[1].

The most widely-used fuzzy clustering method is the fuzzy c-means algorithm [2]. Fuzzy c-means algorithm is used in several clustering problems efficiently. However, it has two major drawbacks that highly affect its performance. Firstly, it is sensitive to initial positions of the clusters and secondly, it can be easily trapped into local minima. To avoid these drawbacks, recently some metaheuristic methods are combined to the fuzzy c-means algorithm [3-7].

In this study, we proposed a new metaheuristic approach to improve the clustering performance of the fuzzy c-means algorithm. The newly proposed metaheuristic approach is called as the "Vortex Search" algorithm (VS), which is a random search method with an adaptive step size adjustment scheme. When compared to the existing methods the VS algorithm is very simple and thus computationally inexpensive. The proposed method is tested on the ECG beat clustering problem and its performance is compared to the particle swarm optimization (PSO2011) [8] and artificial bee colony (ABC) algorithms [9].

ECG is a very important tool for the diagnosis of the heart beat abnormalities. Therefore, numerous methods (including fuzzy clustering methods) are proposed for automatically diagnosis of the heart beat abnormalities from ECG recordings [10-14].

Remaining part of this paper is organized as follows: In Section-2, fuzzy c-means clustering algorithm and the newly proposed metaheuristic, the VS algorithm, are introduced. Then, the methodology used to improve the clustering performance of the fuzzy c-means algorithm with the newly proposed metaheuristic approach is given. Section-3 covers the dataset formation from the ECG recordings. In Section-4 experimental results are given and finally, Section-5 concludes the work.

2 Methodology

2.1 Fuzzy c-means algorithm

Clustering is a method of unsupervised learning which aims to assign a given set of objects into disjoint groups or clusters based on similarity criteria, such that objects from the same clusters are similar and objects from different clusters are dissimilar. The similarity or dissimilarity measure is calculated by a distance function [1].

Let us say C_i, $i = 1,2,.....c$ are the cluster prototypes and c_i, $i = 1,2,.....c$ are the clusters' centers. By using a predetermined distance measure (i.e. Euclidian distance) a given algorithm creates C_i cluster prototypes each having a cluster center, c_i.

Probably the most popular clustering algorithm is the hard C-means (or k-means) algorithm which assigns each data point x_j in a given dataset $X = (x_1, x_2,...,x_n)$, $X \subseteq \Re^p$ to exactly one cluster. Such an assignment can be inadequate because some data points can be in a location which is almost equally distant from two or more cluster centers. By forcing such a point to exactly one cluster, the similarity of this point to other clusters is totally ignored. For this reason fuzzy clustering methods are proposed. In fuzzy clustering methods a data point can belong to more than one cluster with different degrees of membership which is useful especially when the clusters overlap each other [1].

Table 1. Fuzzy c-means algorithm

```
Select the fuzzifier exponent m(m>1) and initialize the
fuzzy partition matrix U=(u_ij) randomly.
while termination conditions not met do
    Compute the cluster centers c_i with Eq. 6
    Update the fuzzy partition matrix U=(u_ij) with  Eq.5
end while
```

Since a data point has membership to each cluster center c_i, the membership vector of this data point then can be defined as,

$$u_j = (u_{1j},...,u_{cj}) \tag{1}$$

If the data set contains n data points then a $c \times n$ fuzzy partition matrix $U = (u_{ij}) = (u_1,...,u_n)$ can be formed. Following constraints must be ensured for the fuzzy partition matrix.

$$\sum_{j=1}^{n} u_{ij} > 0, \ \forall i \in \{1,...,c\} \tag{2}$$

$$\sum_{i=1}^{c} u_{ij} = 1, \ \forall j \in \{1,...,n\} \tag{3}$$

First constraint implies that, all clusters must involve at least one data point, which means there is not any empty cluster. Second constraint ensures that the sum of the membership degree of a data to all clusters equals to 1. Both of these constraints prevent full membership of all data points to only one cluster.

Both hard and fuzzy c-means algorithm use objective function J, which is basically a mathematical criterion to evaluate the fitness of the resulted clusters. These func-

tions can be interpreted as cost functions that have to be minimized in order to find optimum clusters. Thus, for a given dataset X, fuzzy c-means algorithm performs the optimal partitioning when the following objective function is minimized:

$$J = \sum_{i=1}^{c} \sum_{j=1}^{n} u_{ij}^{m} \left\| x_j - c_i \right\|^2 \tag{4}$$

where, $m \, (m>1)$ is called the fuzzifier exponent, u_{ij} is the membership degree of the jth data point x_j to the ith cluster center c_i.

Bezdek et. al. [2] proposed the fuzzy c-means algorithm which iteratively optimizes the objective function J, by updating the fuzzy partition matrix $U = (u_{ij})$ and cluster centers c_i with the following formulas, at each iteration.

$$u_{ij} = \cfrac{1}{\sum_{l=1}^{c} \left(\cfrac{d_{ij}^2}{d_{lj}^2} \right)^{\frac{1}{m-1}}} \tag{5}$$

$$c_i = \cfrac{\sum_{j=1}^{n} u_{ij}^{m} x_j}{\sum_{j=1}^{n} u_{ij}^{m}} \tag{6}$$

In Table-1 a brief description of the fuzzy c-means algorithm is given. Algorithm can be terminated when the relative change in the cluster centers get very small or the objective function J cannot be minimized anymore.

2.2 The newly proposed metaheuristic: Vortex search algorithm

In the past two decades, metaheuristics (both single solution and population-based) are applied to many real-world optimization problems successfully. The success of these algorithms lie behind the good balance they provided between the explorative and exploitative behavior of the search. The explorative behavior of the search defines the global search ability of the algorithm whereas, the exploitative behavior defines the ability of finding the optimum around a near-optimal solution which can also be considered as the local search ability. Single-solution based metaheuristics are accepted to be more exploitation oriented, whereas population-based metaheuristics are more exploration oriented.

In this study, a new single-solution based metaheuristic namely, Vortex Search (VS) algorithm, and its application to fuzzy clustering of ECG beats is proposed. The

VS algorithm can be studied within the family of the search algorithms that comprises the Random Search (RS) and Pattern Search (PS) algorithms.

Table 2. A description of the proposed VS algorithm

Inputs: Initial solution s_0 is selected to be the center of the search space.
Initial radius r_0 is computed according to the upper and lower limits of the problem
Best solution found so far $s_{best} = \inf$
$t = 0;$
Repeat
    /* Generate candidate solutions by using Gaussian distribution with the center s_t and standard deviation (radius) r_t */
    Generate($C(s_t)$) ;
    /* If exceeded shift the $C(s_t)$ values into the boundaries*/
    $C(s_t) > upperlimit = rand \cdot (upperlimit - lowerlimit) + lowerlimit$
    $C(s_t) < lowerlimit = rand \cdot (upperlimit - lowerlimit) + lowerlimit$
    /* Select the best solution from $C(s_t)$ to replace the best solution
    found so far s_{best} */
    $s' = $ Select($C(s_t)$) ;
    if $f(s') < f(s_{best})$
            $s_{best} = s'$
     else
        keep the best solution found so far.
     end
    /* The next center is always shifted to the best solution found so far */
    $s_{t+1} = s_{best}$
    /* Decrease the radius for the next iteration */
    $r_{t+1} = $ Decrease(r_t)
    $t = t + 1;$
Until the termination condition is met.
Output: Best solution found s_{best}

The Random Search algorithm (which is also known as Fixed Step Size Random Search) is attributed to Rastrigin [16], who made an early presentation of RS along with basic mathematical analysis. RS works by iteratively moving to better positions in the search space which are sampled from an hyper sphere surrounding the current position. The Pattern Search (PS) algorithm is also a similar kind of algorithm which was proposed by Hooke and Jeeves [17]. The problem with the above mentioned algorithms is "the step size" which dramatically affects the performance of the algorithms. Thus, none of these algorithms could challenge to the population based metaheuristics. In the proposed VS algorithm this problem is avoided by using an adaptive step size adjustment scheme.

In the proposed VS algorithm, first a number of candidate solutions $C_0(s)$ is generated within an hyper sphere defined with a Gaussian distribution with the center (the initial solution, s_0) and the standard deviation, r_0. Here, the standard deviation r_0 can be considered as the initial radius of the hyper sphere. Then, a solution (which is the best one) $s' \in C_0(s)$ is selected and memorized from $C(s)$ to replace the current solution s_0. The memorized best solution s' is assigned to be the center of the second hyper sphere (the inner one). In the second step, the effective radius (r_0) of the old hyper sphere is reduced and r_1 (the radius of the new hyper sphere) is obtained. Then, a new set of solutions $C_1(s)$ is generated around the new center $s_1 = s'$. This process iterates until the termination condition is met. By this way, once the algorithm is terminated the resulting pattern looks like a vortex like structure which can be represented by a number of nested circles for a two dimensional optimization problem, where the center of the smallest circle is the optimum point found by the algorithm. That is why the algorithm is called as "Vortex Search".

In Table-2 a short description of the proposed VS algorithm is given. As it can be shown from Table-2, the proposed VS algorithm is quite simple. Different from the high-level representation of the single-solution based metaheuristics given in [18], the proposed VS algorithm uses a poor memory (in which always the best solution is memorized) and an additional step in which the radius is iteratively decreased. The way how the radius is decreased is of critical importance for the success of the algorithm. This process should be performed in such a way that allows the algorithm to behave in an explorative manner in the initial steps and in an exploitative manner in the further steps. To achieve this, the value of the radius must be tuned properly during the search process. In the VS algorithm, the inverse incomplete gamma function is used for this purpose.

The incomplete gamma function given in Eq.7 most commonly arises in probability theory, particularly those applications involving the chi-square distribution [19].

$$\gamma(a, x) = \int_0^x e^{-t} t^{a-1} dt \quad a > 0 \tag{7}$$

In conjunction with the incomplete gamma function, usually its complementary $\Gamma(a,x)$ is also introduced (Eq.8).

$$\Gamma(a,x) = \int_x^\infty e^{-t}t^{a-1}dt \quad a > 0 \tag{8}$$

Thus, it follows that,

$$\gamma(a,x) + \Gamma(a,x) = \Gamma(a) \tag{9}$$

where, $\Gamma(a)$ is known as the gamma function. There exist many studies in literature that proposed different methods for the numerical calculation of the incomplete gamma function [20-22]. MATLAB® also provides some tools for the calculation of functions including, gamma function (gamma), incomplete gamma function (gammainc), and inverse incomplete gamma (gammaincinv) function. The inverse incomplete gamma function (gammaincinv), computes the inverse of the incomplete gamma function with respect to the integration limit x and represented as gammaincinv(x,a) in MATLAB®.

Table 3. Statistical results of 30 runs for VS, PSO2011 and ABC algorithms on numerical function optimization

Function	Range	Dim	Min.		PSO2011	ABC	VS
Rastrigin	[-5.12, 5.12]	30	0	Mean	26,11016129	0	57,60799224
				StdDev	5,686650032	0	13,94980276
				Best	16,91429893	0	33,82857771
Rosenbrock	[-30, 30]	30	0	Mean	0,930212233	0,003535257	0,367860114
				StdDev	1,714978077	0,003314818	1,130879848
				Best	2,14999E-19	7,08757E-05	9,42587E-05
Griewank	[-600, 600]	30	0	Mean	0,004761038	0	0,032798017
				StdDev	0,008047673	0	0,018570459
				Best	0	0	0,00739604
Ackley	[-32, 32]	30	0	Mean	0,660186991	2,44545E-14	1,15463E-14
				StdDev	0,711496752	3,02083E-15	3,61345E-15
				Best	7,99361E-15	2,22045E-14	7,99361E-15
Schwefel	[-500, 500]	30	-12569.5	Mean	-8316,185447	-12569,48662	-11283,05416
				StdDev	463,9606712	1,85009E-12	352,1869262
				Best	-9466,201047	-12569,48662	-11799,62928
Sphere	[-100, 100]	30	0	Mean	0	2,78624E-16	6,5681E-265
				StdDev	0	2,56204E-17	0
				Best	0	2,23487E-16	8,181E-284

The inverse incomplete gamma function provides a good balance between the explorative and exploitative behavior of the search. In Figure-1 (a) and (b) an example plot of $(1/x) \cdot gammaincinv(x,a)$ (for $x = 0.1$ and a values ranging from 1 down to 0 with a step size of 0.01) and a representative pattern is given for a two dimensional optimization problem that has upper and lower limits between [-10,10] interval and a

global optimum around [0,0]. Here, $x = 0.1$ defines the value that the inverse incomplete gamma function will have at approximately half of the search. a values are determined with the number of iterations (for 100 iterations a values ranging from 1 down to 0 with a step size of 0.01). As the number of iteration increases the resolution of the search also increases and thus the nested circles given in Figure-1-b become denser.
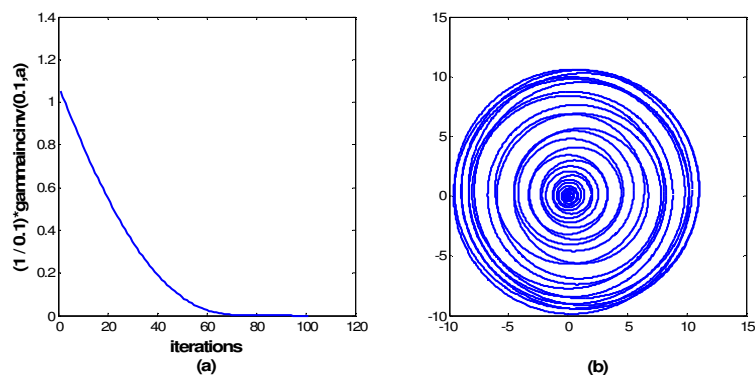

Fig. 1. $(1/x) \cdot gammaincinv(x, a)$ (for $x = 0.1$ and a values ranging from 1 down to 0 with a step size of 0.01) and the resulting pattern is given

Prior to the fuzzy clustering of ECG beats, the proposed VS algorithm is tested over some well-known benchmark numerical functions and results are compared to the PSO2011 and ABC algorithms. In Table-3, statistical results of 30 runs are provided. Here, population based metaheuristics (ABC, PSO2011) are selected to have a population size of 50 which is also the number of neighborhood solutions of the proposed VS algorithm.

The acceleration coefficients (c_1 and c_2) of the PSO2011 algorithm are both set to 1.8 and the inertia coefficient is set to 0.6. The limit value for the ABC algorithm is determined as limit = SN * D where, SN represents the number of food sources and D represents the dimension. Maximum number of function evaluations is selected as 500.000.

From Table-3 it can be shown that, in spite of its simplicity the proposed VS algorithm is highly competitive when compared to the population-based algorithms PSO2011 and ABC.

2.3   Fuzzy c-means algorithm with VS (FCM-VS)

As mentioned before, fuzzy c-means algorithm has two major drawbacks; first, it can be easily trapped into local minima and second, it is sensitive to initialization. To overcome these drawbacks in this paper fuzzy c-means algorithm is combined to the proposed VS algorithm. In FCM-VS cluster centers are encoded as a candidate solution for VS algorithm as follows:

$$\left\{c_1^i,...,c_k^i\right\}=\left\{s_1^i,...,s_k^i\right\} \tag{10}$$

where, $\left(i=1,...,n\right)$, n represents the number of dimension and k represents the number of clusters.

For a given training set the VS algorithm is initialized. After a certain number of iterations optimum cluster centers are found.
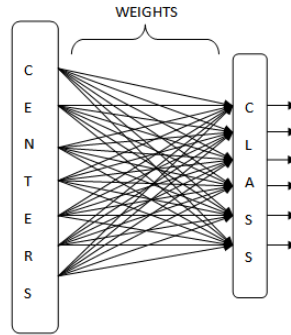

Fig. 2. Classification stage of the proposed system

Clustering is the unsupervised learning part of the proposed method, while to evaluate the clustering performance of the proposed method ECG beats are classified in a supervised manner. Fig.2 shows the classification stage of the proposed system. Here, the goal is to find the weights between the cluster centers and the output of the proposed system. The weight matrix W can be obtained by the pseudo-inverse method as follows:

$$W=\left(U^TU\right)^{-1}U^T*T \tag{11}$$

where, U represents the fuzzy partition matrix and T represents the target output matrix.
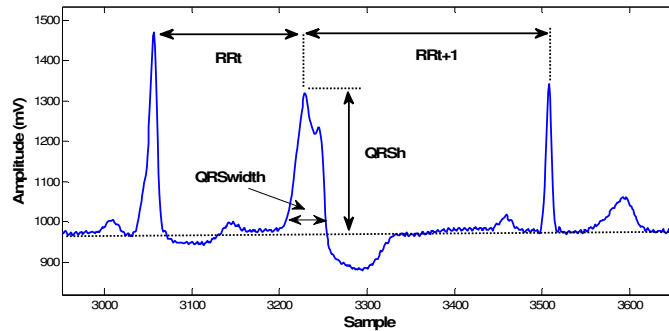

Fig. 3. Extracted features from ECG records

Once the weights are found, classification is performed by computing the fuzzy partition matrix U_t by using Eq.5 for the test set by using the centers found by the FCM-VS algorithm. Then the output of the proposed classifier is found easily as follows:

$$output = U_t * W \qquad (12)$$

3   The ECG Dataset and Extracted Features for Experiments

In order to test the clustering performance of the FCM-VS algorithm, a number of records including six types of beats, Normal Beat (N), Premature Ventricular Contraction (PVC), Fusion of Ventricular and Normal Beat (F), Artrial Premature Beat (A), Right Bundle Branch Block Beat (R) and Fusion of Paced and Normal Beat (f) are selected from the MIT-BIH arrhythmia database. Then, in the preprocessing stage noise removal and baseline correction is performed for each record. After the preprocessing of the selected records, R peaks of each QRS complex are detected. Then, ECG beats are collected by selecting 128 points from each side of the R peaks. After the normalization process, four morphological features, including RR_t, RR_{t+1} intervals, *QRSh* and *QRSwidth* are extracted from the beats. Extracted features can be shown in Fig. 3. A training set which includes 50 feature vectors for each class is formed for the training phase of the classification. Fig. 4 shows the distribution of the feature vectors within the training set.

Table 4. Test set

Beat type	Number of beats	MIT-BIH Records
N	836	113, 119, 213, 234
PVC	605	119, 208, 221, 223, 228, 233
F	248	208, 213
A	165	209, 220, 222
R	150	212
f	130	217

4   Results and Discussion

For the given training set, the proposed FCM-VS algorithm is initialized, and optimum clusters are found. By using the obtained cluster centers weights for the classification stage of the proposed system are computed. Classification performance of the proposed system is then tested over the test set given in Table 4.

In Fig.5a, classification result of the proposed method is given for different number of clusters (6, 10, 15, 20, 25, 30 and 35 clusters) and results compared to the FCM, FCM-PSO2011 and FCM-ABC algorithms. From this figure it can be shown that the average classification results of ten experiments for the proposed method is superior

to FCM, FCM-PSO2011and FCM-ABC algorithms. For a total number of 35 clusters the proposed method gives 93.58% sensitivity. In Fig.5b, average computational results of 10 experiments for the FCM-VS, FCM-PSO2011 and FCM-ABC are also given for different number of clusters. As expected, the average computational time for the proposed method is less when compared to the FCM-PSO2011 and FCM-ABC algorithms. Although the difference is not too much for 500 iterations, it dramatically increases as the number of iterations increases. Experiments are performed in Matlab2009b with an Intel Pentium 1.86 GHz Notebook with 2GB RAM.
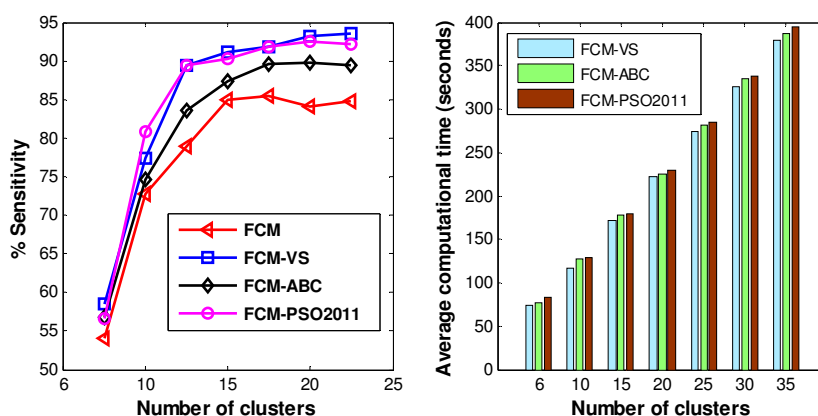

Fig. 4. (a) Average classification results for different number of clusters (b) average computational time for 500 iterations for different number of clusters

5 Conclusion

This study proposes a new metaheuristic approach for fuzzy clustering of ECG beats. The newly proposed method is quite simple when compared to the other population-based metaheuristics. In spite of its simplicity, the proposed method is highly competitive and gives quite good results per experiment. Moreover, due to its simplicity the computational cost of the proposed method is less than the other algorithms. These advantages make the newly proposed metaheuristic approach to be a good candidate for other problems, such as neural network optimization for classification applications, crisp clustering, and regression.

In the future studies, the proposed method will be applied some other datasets related to the biomedical, bioinformatics by using both clustering and some other supervised methods.

6   References

1. Berat Doğan, Mehmet Korürek, A new ECG beat clustering method based on kernelized fuzzy c-means and hybrid ant colony optimization for continuous domains, Applied Soft Computing, Volume 12, Issue 11, November 2012, Pages 3442-3451, ISSN 1568-4946
2. J.C. Bezdek, R. Ehrlich, and W. Full,   "FCM: The fuzzy c-means clustering algorithm", *Computers & Geosciences*, vol. 10, no. 2-3, pp.191 - 203 , 1984.
3. B. Biswal, P.K. Dash, S. Mishra, A hybrid ant colony optimization technique for power signal pattern classification, *Expert Systems with Applications*, Volume 38, Issue 5, May 2011, Pages 6368-6375, ISSN 0957-4174, 10.1016/j.eswa.2010.11.102.
4. Taher Niknam, Babak Amiri, An efficient hybrid approach based on PSO, ACO and k-means for cluster analysis, *Applied Soft Computing*, Volume 10, Issue 1, January 2010, Pages 183-197, ISSN 1568-4946, 10.1016/j.asoc.2009.07.001.
5. Dervis Karaboga and Celal Ozturk, Fuzzy clustering with artificial bee colony algorithm, *Scientific Research and Essays,* Vol. 5(14), pp. 1899–1902, 18 July, 2010
6. M.G.H. Omran, M. Clerc, 2011, <http://www.particleswarm.info/>, accessed 3 December 2013
7. D. Karaboga, B. Basturk, A. powerful, A powerful and efficient algorithm for numerical function optimization: artificial bee colony (abc) algorithm, Journal of Global Optimization 39 (3) (2007) 459–471.
8. Yüksel Özbay, Rahime Ceylan, Bekir Karlik, Integration of type-2 fuzzy clustering and wavelet transform in a neural network based ECG classifier, *Expert Systems with Applications*, Volume 38, Issue 1, January 2011, Pages 1004-1010, ISSN 0957-4174, 10.1016/j.eswa.2010.07.118.
9. Rahime Ceylan, Yüksel Özbay, Bekir Karlik, A novel approach for classification of ECG arrhythmias: Type-2 fuzzy clustering neural network, Expert Systems with Applications, Volume 36, Issue 3, Part 2, April 2009, Pages 6721-6726, ISSN 0957-4174, 10.1016/j.eswa.2008.08.028.
10. Dogan, B.; Korurek, M., "ECG beat clustering using fuzzy c-means algorithm and particle swarm optimization," *Signal Processing and Communications Applications Conference (SIU), 2012 20th* , vol., no., pp.1,4, 18-20 April 2012
11. Ince, T.; Kiranyaz, S.; Gabbouj, M.; , "A Generic and Robust System for Automated Patient-Specific Classification of ECG Signals," *Biomedical Engineering, IEEE Transactions on* , vol.56, no.5, pp.1415-1426, May 2009 doi: 10.1109/TBME.2009.2013934
12. Rastrigin, L.A. (1963). "The convergence of the random search method in the extremal control of a many parameter system". Automation and Remote Control 24 (10): 1337–1342.
13. Hooke, R.; Jeeves, T.A. (1961). ""Direct search" solution of numerical and statistical problems". Journal of the Association for Computing Machinery (ACM) 8 (2): 212–229
14. E.-G. Talbi Metaheuristics: From Design to Implementation Wiley & Sons, Hoboken, New Jersey (2009)
15. Larry C. Andrews, Special Functions of Mathematics for Engineers, SPIE Press, 1992
16. S. Winitzki Computing the incomplete Gamma function to arbitrary precision Computational Science and Its Applications – ICCSA 2003, of LNCS, Vol. 2667 Springer-Verlag, Berlin (2003), pp. 790–798
17. Giampietro Allasia, Renata Besenghi. Numerical calculation of incomplete gamma function by the trapezoidal rule. Numer. Math. [Numerische Mathematik] 50 (4):419{428, 1987