Proceedings of the Sixth IASTED International Conference
BIOMEDICAL ENGINEERING
February 13–15, 2008, Innsbruck, Austria
ISBN Hardcopy: 978-0-88986-721-5 / CD: 978-0-88986-722-2

ENHANCED FEATURE SELECTION ALGORITHM USING ANT COLONY OPTIMIZATION AND FUZZY MEMBERSHIPS

Rami N. Khushaba, Akram AlSukker, Ahmed Al-Ani, and Adel Al-Jumaily
Faculty of Engineering
University of Technology, Sydney
PO Box 123, Broadway, NSW 2007, Australia
rkhushab@eng.uts.edu.au, alsukker@eng.uts.edu.au, ahmed@eng.uts.edu.au , adel@eng.uts.edu.au

ABSTRACT
Feature selection is an indispensable pre-processing step when mining huge datasets that can significantly improve the overall system performance. This paper presents a novel feature selection method that utilizes both the Ant Colony Optimization (ACO) and fuzzy memberships. The algorithm estimates the local importance of subsets of features, i.e., their pheromone intensities by utilizing fuzzy c-means (FCM) clustering technique. In order to prove the effectiveness of the proposed method, a comparison with another powerful ACO based feature selection algorithm that utilizes the Mutual Information (MI) concept is presented. The method is tested on two biosignals driven applications: Brain Computer Interface (BCI), and prosthetic devices control with myoelectric signals (MES). A linear discriminant analysis (LDA) classifier is used to measure the performance of the selected subsets in both applications. Practical experiments prove that the new algorithm can be as accurate as the original method with MI, but with a significant reduction in computational cost, especially when dealing with huge datasets.

KEY WORDS
Ant colony optimization, feature selection

1. Introduction

Feature Selection (FS) has a very important role in data processing. It aims to reduce the feature set dimensionality through selecting a subset of features that performs the best under some classification problem. This can be done by eliminating irrelevant and redundant features. The search methods utilized in feature selection techniques are divided into two categories, according to their dependency on the classification algorithms, those are: filters and wrappers. Filter based FS methods are in general faster than wrapper based methods [1], due to the fact that filter based FS methods depend on some type of estimation of the importance of individual feature or subset of features. On the other hand, wrapper based FS methods are more accurate as the importance of feature subsets is measured using a classification algorithm.

A search strategy is needed to explore the feature space. A full search (exhaustive search) of all possible subsets will guarantee finding the best solution. However, exhaustive search is not viable when dealing with a large (or even moderate) number of variables. The selection of S features from the available N features will produce C combinations ($C = N! / (S! \times (N-S)!)$). If $N = 25$ and $S = 10$, then $C = 3{,}268{,}760$. Hence, a powerful search algorithm becomes a necessity.

Various search algorithms that differ in their optimality and computational cost have been developed to partially search the solution space. These methods include: Sequential Search (SS), Tabu Search (TS) [2], Simulated Annealing (SA) [3, 4], Genetic algorithms (GA) [5, 6], Ant Colony Optimization (ACO) [7-13], and Particle Swarm Optimization (PSO) [14, 15].

Ant colony optimization is a promising approach to solve discrete optimization problems, it was initially used to solve the well known travelling salesman problem[16]. However, there is little work in the literature that utilizes ACO in feature selection. ACO based feature selection has been adopted in dimensionality reduction problems in a small number of applications such as: face, speech, texture, and medical diagnostic [7-13]. The literature shows that ACO can achieve better results than other state of the art search algorithms, such as GA and PSO. The performance (accuracy and speed) of ACO depends on the local measure used to update pheromone trials. This paper presents a new implementation of ACO using a modified fuzzy membership as a local measure to speed up the search procedure. The results of this algorithm will be compared with another ACO variant in two biomedical problems: brain computer interface (BCI) and prosthetic devices control. Initial results show noticeable reduction in computational complexity while maintaining similar performance.

This paper is structured as follow: ACO and one of its variants are described in section 2. The proposed fuzzy membership ACO is introduced in section 3. Section 4 describes the BCI and prosthetic devices control problems. Experimental results are presented in Section 5, and a conclusion is given in Section 6.

2. Ant Colony Optimization

In real ant colonies, a pheromone, which is an odour substance, is used as an indirect communication medium. When a source of food is found, the ants lay some pheromone to mark the path. The quantity of the laid pheromone depends upon the distance, quantity and quality of the food source. While an isolated ant that moves at random detects a laid pheromone, it is very likely that it will decide to follow its path. This ant will itself lay a certain amount of pheromone, and hence enforce the pheromone trail of that specific path. Accordingly, the path that has been used by more ants will be more attractive to follow. In other words, the probability with which an ant chooses a path increases with the number of ants that previously chose that path. This process is hence characterised by a positive feedback loop [16].

Dorigo et. al. [17] adopted this concept and proposed an artificial colony of ants algorithm, which was called the Ant Colony Optimisation (ACO) metaheuristic, to solve hard combinatorial optimisation problems. The ACO was originally applied to solve the classical Travelling Salesman Problem (TSP) [16], where it was shown to be an effective tool in finding good solutions. The ACO has also been successfully applied to other optimization problems including data mining [18], telecommunications networks [19], vehicle routing [20], face recognition [21], and others . However, there are a small number of applications of ACO in features selection especially in BCI, while there were no attempts at all for applying it in MES classification problems. Therefore, in this paper, the ACO will be modified to best serve the purpose of feature selection in multi-channel problems (which includes BCI and MES recognition). This is achieved by constructing a number of parallel colonies and will attempt to collectively search for features that work well together.

The original ACO algorithm that was developed for the TSP is characterized by:

- At each iteration, the pheromone values are updated by all the m ants that have built a solution in the iteration itself. The pheromone τ_{ij}, associated with the edge joining cities i and j is updated as follows:

$$\tau_{ij} \leftarrow (1-\rho).\tau_{ij} + \sum_{k=1}^{m} \Delta\tau_{ij}^{k}. \tag{1}$$

Where ρ is the evaporation rate, m is the number of ants, and $\Delta\tau_{ij}^{k}$ is the quantity of pheromone laid on edge *(i,j)* by ant *k*:

$$\Delta\tau_{ij}^{k} = \begin{cases} Q/L_k & \text{if ant } k \text{ used edge } (i,j) \text{ in its tour,} \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

where Q is a constant and L_k is the length of the tour constructed by ant *k*.

- The probability of choosing the next node is given by:

$$p_{ij}^{k} = \begin{cases} \dfrac{\tau_{ij}^{\lambda}.\eta_{ij}^{\omega}}{\sum_{c_{il}\in N(s^p)} \tau_{il}^{\lambda}.\eta_{il}^{\omega}} & \text{if } c_{ij} \in N(s^p) \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

where $N(s^p)$ represents the set of feasible nodes. λ and ω are constants to control the relative importance of the pheromone versus the heuristic information η_{ij}, which is given by:

$$\eta_{ij} = \frac{1}{d_{ij}} \tag{4}$$

where d_{ij} is the distance between cities i and j.

The feature selection problem differs from TSP as the distance between cities are fixed in TSP. When adding one more city, the change in the objective function is affected only by the distance between last two cities. In contrast to TSP, adding a feature to an existing subset of features can have an impact on the overall performance. A relevant feature will produce a better subset, and hence improve the performance, while an irrelevant feature may degrade the performance of the original subset. When adding a feature to the current feature subset the local performance measure should take into account the relationship with all previously selected features and not only the last one.

Different local measures have been implemented in the literature. Mutual information (MI) was used in [7, 8, 13]. Mutual information based ACO from [7] is adapted as base line comparison, since the theoretical background behind this method is easily understood and is well justified. This method was termed ACOMIEF in order to distinguish it from the proposed new method called ACOFUZZY. In the ACOMIEF, the selection of next node (feature) is based on the following selection measure:

$$SM_i^{S_j} = \begin{cases} \dfrac{(\tau_i)^{\eta}(LI_i^{S_j})^{K}}{\sum_{g\notin S_j}(\tau_g)^{\eta}(LI_g^{S_j})^{K}} & \text{if } i \in S_j \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

where $LI_i^{S_j}$ is the local importance of feature f_i given the subset S_j. The parameters η and K control the effect of trial intensity and local feature importance respectively.

$$LI_i^{S_j} = I(C;f_i) \times \left[\frac{2}{1+\exp(-\alpha D_i^{S_j})} - 1\right] \tag{6}$$

where

$$D_i^{S_j} = \min_{f_s\in S_j}\left[\frac{H(f_i) - I(f_i,f_s)}{H(f_i)}\right] \times \frac{1}{|S_j|}\sum_{f_s\in S_j}\left[\beta\left(\frac{I(C;\{f_i,f_s\})}{I(C;f_i)+I(C;f_s)}\right)^{\gamma}\right] \tag{7}$$

The parameters α, β, and γ are constants, $H(f_i)$ is the entropy of f_i , $I(f_i;f_s)$ is the mutual information between f_i and f_s ,$I(C; f_i)$ is the mutual information between the class labels and f_i, and $|S_j|$ is the cardinal of S_j. For a detailed explanation of the method, the reader can be referred to [1] .

As mentioned earlier, the major drawback of this method is the high computational time needed to estimate the mutual information between each pair of features (with or without the class label), especially when dealing with large feature sets, in which the case the computational cost increases dramatically. As an alternative, this paper suggests the utilization of fuzzy logic in the estimation of the pheromone intensities by estimating the information contents of variables (features). This is in turn the novel contribution of this paper, were up to the authors knowledge there were no such attempts in the literature to estimate the pheromone intensities by employing fuzzy clustering and fuzzy memberships. Therefore, the introduced method replaces MI in equation (7) by fast fuzzy set membership where the first term is represented by fuzzy correlation and the second term by fuzzy membership function as will be described in section three.

3. Fuzzy C-mean based Subsets Importance Evaluation.

In information theory, the best way to estimate the information contents of any feature or subset of features is through the use of the concepts of mutual information and entropies. The known approach in literature for estimating the mentioned quantities employs the concept of probabilities generated by using histograms. Although this approach is known to present quite powerful estimation, especially when using sufficient number of instances, but the computational cost for such probabilities increases dramatically particularly when dealing with high dimension datasets. This is the exact case with most of biosignals driven systems. Due to the nature of most biosignals generated by human body, a multi-channel approach to the problem is usually adopted to provide more information, which however will increase the complexity of the problem. The additional computational cost proposed by the statistical approach will in turn decrease the chance for such systems to be clinically viable.

An alternative method to estimate the importance of feature subsets is through the utilization of fuzzy memberships rather than the statistical probabilities. Thus, utilizing a possibilistic approach instead of the probabilistic one. The fuzzy c-means (FCM) algorithm is utilized to compute the membership of each feature and subset of features, thus replacing the $I(C;f_i)$ and $I(C;f_s)$ by $\mu_C(f_i)$ and $\mu_C(f_s)$ that is the membership of feature f_i and the subset f_s in all the classes. Also the entropy $H(X)$ will be replaced by fuzzy entropy $FE(X)$. The fuzzy c-means algorithm attempts to cluster measurement vectors by searching for local minima of the generalized within group sum of squared errors functions (WGSSE). It was proposed by Trivedi and Bezdeck and is given by [22]:

$$J_m(U,v)=\sum_{k=1}^{n}\sum_{i=1}^{c}(\mu_{ik})^m \left|x_k-v_i\right|_A^2 , \quad 1<m<\infty \tag{8}$$

where

- c : is the number of clusters,
- n : is the number of vectors,
- x_k : is the $k'th$ measurement vector, $x_k \in R^n$
- v_i : is the $i'th$ centroid vector,
- m : is the fuzzy coefficient,
- $|\,.\,|_A$: is an inner product norm,

$|Q|_A^2 = Q^T AQ$, and A is a d *d positive definite matrix where is the dimension of the pattern vectors. When $m=1$, the objective function J_m in (3.1) is the classical WGSSE function and the algorithm reduces to the crisp k-means clustering algorithm. For $m>1$ under the assumption that $x_k \neq v_i$, (U,v) may be a local minimum of J_m only if (rephrase):

$$\mu_{ik}=\frac{1}{\sum_{j=1}^{C}\left(\frac{\left\|x_k-v_i\right\|_A}{\left\|x_k-v_j\right\|_A}\right)^{2/(m-1)}} \qquad \forall i,k \tag{9}$$

and

$$v_j=\frac{\sum_{k=1}^{n}(\mu_{ik})^m x_k}{\sum_{k=1}^{n}(\mu_{ik})^m}, \qquad \forall i \tag{10}$$

The FCM algorithm computes the membership of each pattern in all clusters (the clusters are represented by centroid vectors v_i) and then normalize the membership of each specific pattern x_k in all clusters. If this process is to be applied along each feature f_i rather than each pattern, then we can easily get the membership of each feature in each cluster $\mu_{ci}(f_j)$. Now by simply summing the memberships of the feature in each of the clusters we will end up with the total membership of the feature with all the clusters $\mu_C(f_j)$. In the same manner the value $\mu_C(f_s)$ representing the membership of each subset of features in all clusters can be computed following the same approach. For the calculation of the fuzzy entropy we simply followed the approach presented earlier by Khushaba et al [23]. From now on we will have a new term instead of the probability that is termed as the match degree D_C that is given by:

$$D_C=\frac{\sum_{x_d\in c}\mu_c(f_j)}{\sum_{x_d\in C}\mu_C(f_j)} \tag{11}$$

The above equation means that the match degree D_C equals to the summation of membership of feature f_j in class c divided by the membership of feature f_i in all C classes. The fuzzy entropy of the elements of class c is then equal to:

$$FE_c = -D_c \log D_c \tag{12}$$

Then in order to compute the entropy $H(X)$ we have to compute the fuzzy entropy along the universal set.

$$FE = \sum_{c=1}^{C} FE_c \tag{13}$$

The above calculated terms are now normalized and simply used to replace equation (1.6) with the following:

$$D_i^{S_j} = \min_{f_s \in S_j} \left[\frac{FE(f_i) + FE(f_s) - Corr(f_i, f_j)}{FE(f_i) + FE(f_s)} \right] * \frac{1}{|S_j|} \exp\left[\beta \left(\frac{\mu_C(f_i, f_j)}{\mu_C(f_i) + \mu_C(f_j)} \right)^{\gamma} \right] \tag{14}$$

where $Corr(f_i, f_j)$ represents the correlation coefficient between f_i and f_j.

4. Datasets Description

4.1 Brain Computer Interface

BCI data was taken from the Department of Medical Informatics, University of Technology, Graz, Austria. EEG signals were recorded for three right handed females with 56 Ag/AgCl Electrodes, with reference electrode on the right ear. Subjects were placed in an armchair and asked to imagine right or left finger movements according to stimuli on screen.

4.2 Prosthetic devices control with myoelectric signals

The MES dataset used to test the proposed method was acquired by the University of New Brunswick in Canada [24]. The dataset consisted of ten motions associated with three degrees of freedom (DOF's) of the wrist, two different hand grips, and a rest state. In particular they were: forearm pronation, forearm supination, wrist flexion, wrist extension, radial deviation, unlar deviation, key grip, chuck grip, hand open, and a rest state. Each session of the database consisted of two trials or two repetitions of each motion. Six subjects (AW, KS, LH, MW, SM, and WM) were prompted to complete medium force isometric contractions of 5 seconds duration followed by a brief rest period. Each record was 256 ms in duration (256 points sampled at 1024 Hz), pre-filtered between 10-500 Hz using the 4th order Bessel band pass filter with a gain of 2000 and a Common Mode Rejection Ratio (CMRR) greater than 96 db/channel.

5. Experimental results

As an application, the ACOFUZZY was first considered for the BCI problem. Then the prosthetic control problem was also included as an extra application to prove the effectiveness of the proposed method when dealing with different biosignals. In general, the wavelet packet transform (WPT) was employed for feature extraction in both applications.

In BCI, the WPT has been used to extract features from the EEG signal of each channel. Three features have been used that represent energy of the frequency bands 4-8, 8-16 and 16-24 Hz. It has been found that these three features represent a good compromise between computational cost and classification accuracy as shown in [25]. More details on experiment set-up can be found in [26].

In order to select features that maximize the classification accuracy, the number of desired features was varied to be selected between 3 and 99 in the BCI problem. While for the MES classification the chosen range was between 3 and 60. The chosen range of the desired number of features was made based on the total dimensionality of the original feature set. Both the ACOMIEF and ACOFUZZY search procedures are implemented using the average accuracy of a ten-fold cross-validation. LDA classification results are used as a fitness function. Both methods are implemented using the following parameters: number of ant = 50, number of generations = 100, α=1, β=1.65, γ=3 , η=1, K=1, ρ=0.85. When using the whole feature set generated by WPT from the EEG channels (consisting of 168 features) the achieved classification accuracy using all features was 70.45%. While the results of applying both the ACOMIEF and ACOFUZZY methods shown in Fig.1 revealed that both methods achieved an accuracy of 82.9% for the same problem when selecting only part of the original feature set. It is obvious that the performance of the two methods is very close to each other showing some points in which the ACOFUZZ did actually achieve some enhancements over the ACOMIEF results, and others in which the ACOMIEF prevailed.

In the MES classification problem, In order to prove the real power of such methods it was decided to add more complication to the problem that can mach with real time applications. The MES records were decomposed into four levels generating 16 subspaces and those were chosen as features along each channel (256 features for 16 channels). The classification results employing LDA are shown in Fig.2 in which it is very hard to differentiate between the results from both methods. The achieved classification accuracy of both methods is 99.67% when selecting only part of the 256 features generated by WPT. Both feature selection methods were able to remove irrelevant and redundant features which explain the high classification performance.

An important point to mention about the application of the ACOFUZZY in the MES classification problem is that no majority vote post-processing was considered while computing the classification results in Fig. 2. This was chosen to prove the real power of those methods in dimensionality reduction.

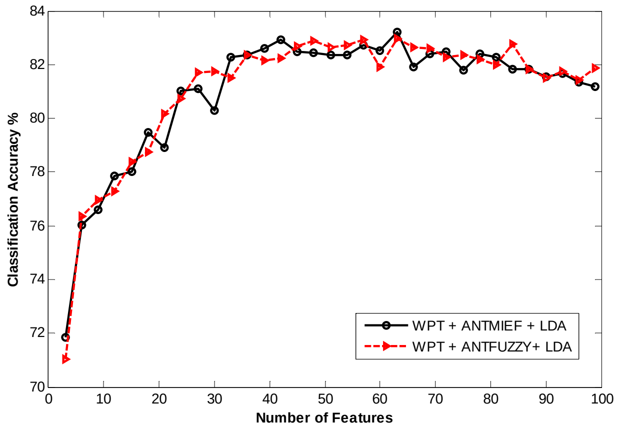

Figure 1 BCI classification accuracy for a selected number of features

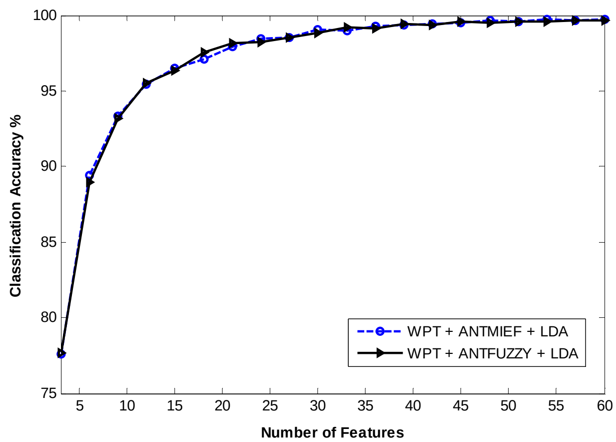

Figure 2 MES classification accuracy for a selected number of features

Both ACOMIEF and ACOFUZZY algorithms provided close results using almost all feature sets. However, the strength of any feature selection algorithm is based on the accuracy achieved, which is the same for both methods, and their speed, which favours the proposed method. The time required to calculate mutual information and fuzzy memberships and entropies used in equations 1.6 and 3.7 were tested using the two dataset (168 features and 256 features). The two methods were implemented using Matlab and they were ran on a conventional PC. In the BCI experiment, the time required to evaluate these values were 5.19 and 197.65 seconds for the fuzzy approach and statistical one respectively. In the MES experiment, these values were 430.13 and 6129.1 seconds respectively. The proposed approach of replacing the MI with the memberships of features reduce the time required to estimate local measures in ACO, i.e. a time reduction of 97.37% and 92.98% were achieved in BCI and MES respectively while maintaining the same classification results, which indeed is an important achievement.

6. Conclusion

This paper presented a developed version of an existing ant colony optimization based feature selection technique that employs the mutual information concept. The new algorithm proposed the use of a fuzzy logic based approach into information content estimation. Both the original (ACOMIEF) and the developed (ACOFUZZY) algorithms were able to produce comparable high classification accuracies using a subset of features instead of using the whole feature set. Both methods achieved close classification performance using the same number of selected features. However, the new fuzzy logic based method outperforms the mutual information algorithm in terms of computational cost, and hence can be used to implement a more viable ACO model. More experiment will be conducted in the future to select features from larger datasets, which will be useful in indicating the advantages, limitations, and other possible applications of the proposed method.

Acknowledgements

The authors would like to acknowledge the support of Levi Hargrove and Dr. Bernard Hudgins from the University of New Brunswick for providing the MES dataset.

References

[1] A. Al-Ani, M. Deriche, and J. Chebil, "A new mutual information based measure for feature selection," Intelligent Data Analysis, vol. 7, pp. 43-57, 2003.

[2] A. B. M. A. Tahir, F. Kurugollu, and A. Amira, "Feature selection using tabu search for improving the classification rate prostate needle biopsies," presented at Proceedings of the Pattern Recognition, 17th International Conference on (ICPR'04), Washington, DC, USA, 2004.

[3] M. Filippone, F. Masulli, and S. Rovetta, "Supervised Classification and Gene Selection Using Simulated Annealing," presented at International Joint Conference on Neural Networks, IJCNN '06., pp.3566-3571, 2006.

[4] F. M. M. Filippone, and S. Rovetta, "Unsupervised gene selection and clustering using simulated annealing," in Lecture Notes in Computer Science, vol. 3849, A. P. I. Bloch, and A. Tettamanzi, Ed.: Springer, pp.229-235, 2005.

[5] O. C. H. Frohlich, and B. Scholkopf, "Feature selection for support vector machines by means of genetic algorithms," in Department of Mathematics and Computer Science, vol. Diploma Thesis: Philipps-University Marburg, 2002.

[6] M. T. Harandi, M. N. Ahmadabadi, B. N. Araabi, and C. Lucas, "Feature selection using genetic algorithm and it's application to face recognition," presented at IEEE

Conference onCybernetics and Intelligent Systems, pp.1368-1373, 2004.
[7] A. Al-Ani, "Feature subset selection using ant colony optimization," International Journal of Computational Intelligence, vol. 2, pp. 53 – 58, 2005.
[8] Z. Chun-Kai and H. Hong, "Feature selection using the hybrid of ant colony optimization and mutual information for the forecaster," Proceedings of International Conference on Machine Learning and Cybernetics. , vol. 3, pp. 1728-1732, 2005.
[9] G. Hai-Hua, Y. Hui-Hua, and W. Xing-Yu, "Ant colony optimization based network intrusion feature selection and detection," presented at Proceedings of 2005 International Conference on Machine Learning and Cybernetics, pp.3871-3875 Vol. 6, 2005.
[10] Y. Han and P. Shi, "An improved ant colony algorithm for fuzzy clustering in image segmentation," Neurocomputing, vol. 70, pp. 665-671, 2007.
[11] H. R. Kanan, K. Faez, and M. Hosseinzadeh, "Face recognition system using ant colony optimization-based selected features," presented at IEEE Symposium on Computational Intelligence in Security and Defense Applications, CISDA., pp.57-62, 2007.
[12] R. K. S. a. S. Ramakrishnan, "A hybrid approach for feature subset selection using neural networks and ant colony optimization," Expert Systems with Applications, vol. 33, pp. 49-60, 2007.
[13] C. Z. a. H. Hu, "Ant colony optimization combining with mutual information for feature selection in support vector machines," Australian Conference on Artificial Intelligence, pp. 918-921, 2005.
[14] E. K. Tang, P. N. Suganthan, and X. Yao, "Feature Selection for Microarray Data Using Least Squares SVM and Particle Swarm Optimization," Proceedings of the IEEE Symposium on Computational Intelligence in Bioinformatics and Computational Biology, CIBCB '05, pp. 1-8, 2005.
[15] H. A. Firpi and E. Goodman, "Swarmed feature selection," Proceedings of the 33rd Applied Imagery Pattern Recognition Workshop (AIPR'04), pp. 112-118, 2004.
[16] M. Dorigo, V. Maniezzo, and A. Colorni, "Ant system: optimization by a colony of cooperating agents," IEEE Transactions on Systems, Man and Cybernetics, Part B, vol. 26, pp.29-41, 1996.
[17] M. Dorigo and T. Stützle, "The Ant Colony Optimization Metaheuristic: Algorithms, Applications, and Advances," in Handbook of Metaheuristics, 2003, pp. 250-285.
[18] C. G. Ajith Abraham, Vitorino Ramos, "Swarm intelligence in data mining," in Studies in Computational Intelligence, vol. 34, J. Kacprzyk, Ed.: Springer, 2006.
[19] D. F., D. C. G., and G. L.M, "Using Ant Agents to Combine Reactive and Proactive Strategies for Routing in Mobile Ad Hoc Networks," International Journal on Computational Intelligence and Applications (IJCIA), Special Issue on Nature-Inspired Approaches to Networks and Telecommunications, vol. 5, June 2005.
[20] R. Montemanni, L. M. Gambardella, A. E. Rizzoli, and A. V. Donati1, "Ant colony system for a dynamic vehicle routing problem," Journal of Combinatorial Optimization, vol. 10, 2005.
[21] X. G. P. Y. Y. L. W.-L. Yan, "The SEMG analysis for the lower limb prosthesis using wavelet transformation," EMBC 2004. Conference Proceedings. 26th Annual International Conference of the Engineering in Medicine and Biology Society., vol. 1, pp. 341 - 344 2004.
[22] M. M. Trivedi and J. C. Bezdeck, "Low-level segmentation of aerial images with fuzzy clustering," IEEE Transactions on Systems, Man and Cybernetics, vol. SMC-16, pp. 589- 598, 1986.
[23] R. N. Khushaba, A. Al-Jumaily, and A. Al-Ani, "Novel feature extraction method based on fuzzy entropy and wavelet packet transform for myoelectric control," 7th International Symposium on Communications and Information Technologies (ISCIT 2007), Sydney, 2007, pp.
[24] L. Hargrove, K. Englehart, and B. Hudgins, "A comparison of surface and intramuscluar myoelectric signal classification," IEEE Transactions on Biomedical Engineering, vol. 54, pp. 847-853, 2007.
[25] A. Al-Ani and A. Al-Sukker, "Effect of feature and channel selection on EEG classification," Proceedings of the 28th IEEE EMBS Annual International Conference, pp. 2171-2174, Aug 30-Sept 3, 2006.
[26] H. Ramoser, J. Muller-Gerking, and G. Pfurtscheller, "Optimal spatial filtering of single trial EEG during imagined hand movement," IEEE Transactions on Rehabilitation Engineering, vol. 8, pp. 441-446, 2000.