Hybrid ACO-PSO Based Approaches for Feature Selection

Kamilia Menghour*,   Labiba Souici-Meslati

LISCO Laboratory, Badji Mokhtar-Annaba University,
P.O. Box 12, 23000, Annaba, Algeria
* Corresponding author's Email: k_menghour@yahoo.fr

Abstract: Feature selection attempts to find the most discriminative information aiming to design an accurate learning system. Feature selection has been the focus of interest for a long time and many works had been done. Recently, the tendency of research in this domain is oriented to the bio-inspired methods. In this paper, we propose hybrid bio-inspired approaches applied to the feature selection problem. The approaches are based on two swarm intelligence methods: ant colony optimization (ACO) and particle swarm optimization (PSO). The performances of these approaches are compared with simple bio-inspired feature selection methods based on ant colony optimization, particle swarm optimization and genetic algorithm. Our experimental results show the efficiency of the proposed approaches in the reduction of selected features number and improvement of classification performance.

Keywords: feature selection (FS); ant colony optimization (ACO); particle swarm optimization (PSO); hybrid bio-inspired approach.

1. Introduction

The effectiveness of a system is matched to the feature extraction step and the samples representation. In many applications, data sets are represented as vectors of extremely high dimension. This motivates the research on feature (or attribute) selection: it is the problem of choosing a small subset of features that ideally is necessary and sufficient to describe the target concept [1]. It can be defined as selecting a subset of M features from a set of N features, M <N, such that the value of a criterion function is optimized over all subsets of size M [2].

Dash and Liu [2] propose four basic steps in a typical feature selection method. A generation procedure: produces subsets of features for evaluations. This procedure can start with no features or with all features and attributes are iteratively added (forward) or removed (backward). The procedure can also start with a random subset of features and attributes are iteratively added, removed or reproduced by a certain procedure. The evaluation function measures the goodness of a subset under examination. Without a stopping criterion, the feature selection process may run exhaustively or forever through the space of subsets. The validation procedure is not a part of the feature selection process itself, but a feature selection method must be validated.

Different methods have been developed and used for feature subset selection using several search strategies and evaluation functions. Liu and Yu [3] have developed three dimensions to categorize feature selection methods: search strategies (complete sequential and random), evaluation criteria (Filter, Wrapper, and Hybrid) and data mining tasks (classification or clustering). Based on the evaluation criteria, several authors prefer separating the feature selection methods into two approaches Filter and Wrapper, according to their dependence or independence of the induction algorithm [3, 4]. The filter model relies on general characteristics of the data to evaluate and select feature subsets without involving any algorithm. The wrapper model requires one predetermined algorithm and uses its performance as the evaluation criterion. It searches for features better suited to the algorithm aiming to improve the performance, but it

also tends to be more computationally expensive than the filter model [5].

Since 1990, several collective behaviors inspired algorithms (like social insects, bird flocking and ant colonies) have been proposed. The application areas of these algorithms refer to well-studied optimization problems like NP-hard problems, clustering and data mining. Particle Swarm Optimization (PSO) and Ant Colonies Optimization (ACO) are currently the most popular algorithms in the swarm intelligence (SI) domain.

Feature selection can be considered as an optimization problem with many competing criteria, which may be, for example, minimizing the number of selected features or maximizing the percentage of correctly classified training samples.

The problem of feature selection has long been an active research topic. The earliest feature selection approaches were based on classical methods and search algorithms such as dynamic programming and branch and bound [6]. Recently, the tendency of research on feature selection is oriented to bio-inspired methods, especially swarm based ones, since the fact that several meta-heuristic algorithms, derived from the behavior of biological systems in the nature have been proposed as powerful methods for global optimizations.

In our study, we were particularly attracted by the hybridization of bio-inspired methods for feature selection. In this context, according to the literature, we have noticed a lack of hybrid ACO-PSO based approaches for feature selection. In this paper, we propose three feature selection approaches based on the hybridization of ant colony optimization and particle swarm optimization.

The basic idea of our first approach ACO-PSO1 is to use the ACO and PSO simultaneously to explore the search space. The performances of the selected subsets are evaluated with the classification accuracy, after each iteration and the best subset found is used as the best solution to both ACO and PSO. In our second approach, ACO-PSO2, the ACO is employed to explore the search space. When the performance of the best subset selected has not significantly improved for several generations, the searching process is switched to the PSO algorithm. Our third approach ACO-PSO3 is an adaptation of the hybrid PSO-ACO algorithm proposed by Shuang et al [7] to the feature selection problem.

The rest of this paper is organized as follows: section 2 introduces ant colony optimization and particle swarm optimization. Section 3 presents feature selection methods using ACO and PSO algorithms. In section 4, we present some hybrid bio-inspired feature selection approaches and a

synthesis of some swarm-based hybrid approaches applied to several application domains. The hybrid approaches proposed in this paper are detailed in section 5. The experimental results obtained with the proposed approaches are presented and compared with ACO and PSO based methods in section 6. Finally, section 7 concludes the paper.

2. Background

2.1. Ant Colony Optimization (ACO)

Ant colony optimization was introduced in early 1990s by Dorigo and his colleagues [8]. ACO algorithm is a novel nature inspired meta-heuristic for the solution of hard combinatorial optimization problems. The main inspiration source of ACO is the foraging behavior of the real ants, and more specifically, the indirect communication between ants within the colony via the secretion of chemical pheromones [9].

The first ACO algorithm, Ant System (AS), was developed by Dorigo et al [10] and was introduced using the Travelling Salesman Problem (TSP). Following the original AS, various improvements were made which gave rise to several other ant algorithms, including Max-Min AS (MMAS) [11], Rank-based AS (ASrank) [12] and Ant Colony System (ACS) [13]. The main differences between AS and these extensions are the way in which the pheromone update is performed as well as some additional details in the management of the pheromone trials (see Table 1).

2.2. Particle Swarm Optimization (PSO)

Particle swarm optimization (PSO) algorithm is a relatively new branch of evolutionary computation techniques, which includes stochastic search algorithms inspired by the mechanics of natural selection and genetics to emulate evolutionary behaviors in biological systems [14]. The PSO is an evolutionary computation technique developed by Kennedy and Eberhart in 1995, based on the behavior of swarming animals such as birds and fish [15].

In PSO, hundreds or thousands of particles search the optimum while communicating with other particles. Each particle p has two state vectors: position x_i^k and velocity v_i^k. These state vectors are simply updated as follows [16]:

$$v_i^k = wv_i^k + c_1 r_1 (p_i^k - x_i^k) + c_2 r_2 (p_g^k - x_i^k) \quad (1)$$
$$x_i^k = x_i^k + v_i^k \quad (2)$$

c1 and c2 are positive constants called cognitive learning rate and social learning rate respectively. r1 and r2 are two random functions in the range [0,1] and w is the inertia weight. p_i^k and p_g^k are the best particle position and the global best position respectively.

The original PSO is designed for real value problems. Now, the algorithms have been extended to tackle discrete problems. The term BPSO 'Binary PSO' or 'discrete PSO' appeared when PSO was used to solve discrete problem [17].

In 1997, Kennedy and Eberhart developed a binary version of PSO for solving combinatorial optimization problems [18], in which the particles take the values of binary vectors of length N and the velocity is interpreted as a probability to change a bit from 0 to 1, or from 1 to 0 when updating the position of particles [14].

Table 1. Different extensions of ACO

Algorithm	Updating	Updating formula
Ant System (AS)	Deposit on all arcs visited by all ants	$\tau_{ij}(t+1) = (1-\rho).\tau_{ij}(t) + \Delta\tau_{ij}(t)$ (3) where $\Delta\tau_{ij}(t) = \sum_{k=1}^{m}\Delta\tau_{ij}^k(t)$ (4) and $\Delta\tau_{ij}^k(t) = \begin{cases} \frac{Q}{L^k(t)} & if\ (i,j) \in T^k(t) \\ 0 & if\ (i,j) \notin T^k(t) \end{cases}$ (5)
Rank based Ant System (ASrank)	Ants sorted by increasing tour length; deposit weighted according to the rank of each ant	$\tau_{ij} = \rho\tau_{ij} + \Delta\tau_{ij}^{best} + \Delta\tau_{ij}^{rank}$ (6) where $\Delta\tau_{ij}^{rank} = \begin{cases} \sum_{\mu-1}^{\sigma-1}(\sigma-\mu)\frac{Q}{L_\mu} & if\ the\ \mu^{th}best \\ & ant\ travels\ on\ edge\ (i,j) \\ 0 & otherwise \end{cases}$ (7) and $\Delta\tau_{ij}^{best} = \begin{cases} \sigma\frac{Q}{L_{best}} & if\ edge\ (i,j)is\ part\ of \\ & the\ best\ solution\ found \\ 0 & otherwise \end{cases}$ (8)
Max–Min Ant System (MMAS)	Deposit only either by the iteration best-ant, or the best-so-far ant; interval $\{\tau_{min}, \tau_{max}\}$	$\tau_{ij} \leftarrow (1-\rho).\tau_{ij} + \Delta\tau_{ij}^{best}$ (9) Where $\Delta\tau_{ij}^{best} = \begin{cases} \frac{1}{L_{best}} & if\ the\ best\ ant\ used \\ & edge\ (i,j)\ in\ its\ tour \\ 0 & otherwise \end{cases}$ (10)
Where: ρ is the evaporation rate, Q is a constant, L_μ is the tour length of the μth best ant, L_k is the tour length of the kth ant, L_{best} is the length of the tour of the best ant, μ ranking index, σ number of elitist ants.		

3. ACO and PSO based feature selection

Most of the existing feature selection algorithms suffer from the problems of stagnation in local optima. In order to avoid these types of problems, ant colony optimization, genetic algorithms and particle swarm optimization randomize the search space stochastically and also use the information of previous history to explore the search space which provides global optimal with proper tuning of the parameters. However, unlike GA, ACO and PSO have no evolution operators like crossover and mutation. But both have memory, which is important for the algorithms. Compared with GA where individuals are in rivalry, the efficiency of PSO and ACO is in the cooperative behavior of agents

(particles or ants) [19].

In the following subsections, we present some feature selection methods based on ACO and PSO.

3.1. ACO based feature selection

Feature selection problem is an example of a difficult discrete problem which can be represented as a graph problem; for this reason, the ACO approach is of interest to solve it.

In most ACO based feature selection methods, nodes in the graph are used to represent features, with edges between them denoting the choice of the next feature. Ants traverse on the graph to look for a path containing part of the nodes which indicate the feature [20].

In the literature, many successful ACO based feature selection algorithms have been proposed. Some of them use hybrid techniques (filter and wrapper) to estimate the heuristic information and overall performance. For example, Kabir *et al* [21,22] propose a new hybrid ACO algorithm for feature selection, called ACOFS. ACOFS uses a hybrid search technique that combines the advantages of wrapper and filter approaches. This algorithm modifies the standard pheromone update and heuristic information measurement rules based on the above two approaches.

Most existing ACO based feature selection algorithms are wrapper approaches which use the classifier performance to evaluate the selected subsets. Kanan and Faez [23] propose an ACO-based feature selection algorithm which adopts the classifier performance and the length of the selected feature vector as heuristic information for ACO. Khushab *et al* [24] use a combination of ACO and a differential evolution operator for feature selection in the Brain Computer Interface and multifunctional myoelectric control problems. Aghdam *et al* [25] propose an ACO algorithm for text feature selection, this algorithm was enhanced by adding a heuristic measure based on statistics and a local search in [26].

Gomez et al. [27] propose a hybrid model multi-colony ACO and Rough Set Theory (RST) to solve distributed feature selection problem. The algorithm uses a multi-colony ACO as a search method. The approach is based on the interchange of pheromone between the colonies and the RST offers the heuristic function to measure the quality of one feature subset.

Huang [28] proposes a novel hybrid ACO-based model that hybridizes the ant colony optimization and support vector machines to select the small and suitable features subset, to simultaneously optimize the feature subset and the SVM kernel parameters.

In [29] a new feature selection based ACO for regression problems was proposed. The algorithm uses a binary representation and integrates this into an ACS framework introducing a new research direction in ant algorithms.

In [30] an advanced binary ACO is presented. Each node in the graph has two sub-nodes, one for selecting and the other for deselecting the features. Several statistical measures are used as heuristics for edges visibility in the graph. In [31] the author present a feature selection approach based on the graph clustering approach and ant colony optimization. The approach starts with the representation of features as a graph, then the features are divided into different clusters. Finally, a novel search strategy based on the ant colony optimization is developed to select the final subset of features.

Tabakhi and Moradi [32] present a novel unsupervised feature selection based ACO method. This method is based on a filter approach which analyzes the relevance and the redundancy of features. Furthermore, a novel heuristic information measure is proposed to improve the accuracy by considering the similarity between subsets of features.

3.2. PSO based feature selection

The PSO application to the feature selection problem requires the use of the binary PSO variant where each particle's position is represented as a binary string of length N, where N is the total number of features. Therefore, each particle's position is a features subset. For example, if {X, Y, Z and W} is the features set and if a particle position is (1,1, 0, 1), then the features subset is {X,Y,W}.

The particles positions are mapped into boolean values using a logistic regression function (see formula 11) and randomly generated threshold (see formula 12).

$$s\left(v_i^k\right) = \frac{1}{1 + exp(-v_i^k)} \qquad (11)$$

$$x_i^k = \begin{cases} 1 & if \quad (rand < s\left(v_i^k\right)) \\ 0 & otherwise \end{cases} \qquad (12)$$

Formula 11 normalizes velocities into [0.1] rang, whereas the formula 12 replaces particle updating position formula (see formula 2). Here, particle positions are coded with binary values corresponding to the state of the feature xi in the k-th particle solution. A solution is obtained from the

binary vector of particle positions by the selection of the set of features with position set to 1 [33].

A large number of Binary PSO algorithms have been proposed for selecting features. Kothari *et al* [34] present a survey paper on particle swarm optimization in feature selection. The paper presents a comparative study of PSO implementations and reviews the success of PSO in various fields of science.

Hidaka and Kurita [16] implement PSO to select the powerful subset of non-neighboring rectangular features (NNRFs) from the various candidates. Niiniskorpi *et al* [35] show that the PSO is effective in the identification of high-performing voxel subsets for functional magnetic resonance imaging (fMRI) volume classification. Li-Yeh Chuang *et al* [36] used an improved binary particle swarm optimization (IBPSO) to implement feature selection, and the K-nearest neighbor (K-NN) method serves as an evaluator of the IBPSO for gene expression data classification problems. Moreover, in [37] a PSO based feature selection algorithm was proposed for an automatic speaker verification system. Chen *et al* [38] integrate PSO algorithm and 1-NN (k-nearest neighbor with k=1) classifiers to address the feature selection problem. The PSO algorithm is adopted to select important features, and the 1-NN classifiers are employed as a fitness of the PSO algorithm to evaluate the efficiency of selected features.

Table 2. Summary of some works using ACO and PSO for feature selection

Authors and reference	SI method	FS Approach	Application Domain
Basiri *et al*., 2008 [39]	ACO	Wrapper	Protein post-synaptic activity prediction
Kanan and Faez, 2008[23]	ACO	Wrapper	Face recognition system
Chuang *et al*., 2009 [36]	PSO	Wrapper	Gene expression data classification
Deng *et al*., 2009 [40]	ACO	Wrapper	Classification
Gomez *et al*., 2009 [27]	ACO	Wrapper	Classification
Huang, 2009 [28]	ACO	Wrapper	Classification
Kabir *et al*., 2009 [41]	ACO	Wrapper	Classification
Lai *et al*., 2009 [14]	PSO	Filter	Spam filtering
Niiniskorpi *et al*.,2009 [35]	PSO	Hybrid	FARMI pattern classification
Bae *et al*., 2010[42]	PSO	Wrapper	Classification
Chen *et al*., 2010 [43]	ACO	Filter	Classification
Nemati and Basiri, 2010[37]	PSO	Wrapper	Automatic speaker verification
Kabir *et al*., 2012 [21]	ACO	Hybrid	Classification
Meena *et al*., 2012 [26]	ACO	Wrapper	Text categorization
Shunmugapriya *et al*., 2012 [44]	ACO	Wrapper	Classification
Xue *et al*., 2012 [45]	PSO	Filter	Classification
Bing *et al*., 2013 [46]	PSO	Wrapper	Classification
Cervante *et* al., 2013 [47]	PSO	Filter	Classification
Chen *et al*., 2013 [20]	ACO	Wrapper	Image Classification
Kabir *et al*., 2013[22]	ACO	Hybrid	Classification
Vieira *et al*., 2013[48]	PSO	Wrapper	Mortality prediction of septic patients
Xue *et al*., 2013[49]	PSO	Wrapper	Classification
Inbarania *et al*., 2014 [50]	PSO	Filter	Medical diagnosis problems
Nguyen *et al*., 2014 [51]	PSO	Hybrid	Classification
Abd-Alsabour, 2015 [29]	ACO	Wrapper	Regression problems
Kashef and Nezamabadi-pour, 2015 [30]	ACO	Wrapper	Classification
Moradi and Rostami, 2015 [31]	ACO	Filter	Classification
Tabakhi and Moradi, 2015 [32]	ACO	Filter	Classification

Some feature selection approaches address the problem as a multi-objective one. Xue *et al* [45] propose a filter multi-objective feature selection algorithm based on BPSO for classification. Two multi-objective FS algorithms are developed by applying mutual information and entropy as two different filter evaluation criteria in the proposed framework. Cervante *et al* [47] propose a novel feature selection algorithm based on a multi-objective PSO and probabilistic rough set theory with the goal of obtaining a set of non-dominated features subsets.

Furthermore, in [20] an ACO-based feature selection algorithm is presented to reduce the memory requirement and computation time. This algorithm uses the classifier performance and feature set size to guide the search, and optimizes the feature set in terms of its size and classifier performance. In [49] a new initialization strategy for the PSO algorithm is presented. This strategy based on the ideas of forward and backward selection methods. The authors also proposed a new pbest and gbest updating mechanism to overcome the limitation of the traditional updating mechanism in order to ensure the feature subset with the highest classification performance and the smallest number of features.

In [50] the authors introduced supervised feature selection methods based on a hybrid approach to solve the medical diagnosis problems. The approach combines the strength of Rough Set Theory and PSO.

Nguyen *et al* [51] propose a new feature selection approach based on particle swarm optimization (PSO). The local search is based on a typical backward elimination method to improve the gbest during the search process. The approach uses a mutual information filter measure to incorporate with the wrapper fitness function aiming to take the advantages of both filter and wrapper approaches.

4. Hybrid bio-inspired approaches

4.1. Hybrid bio-inspired approaches for feature selection

Using hybrid models became a common practice in feature selection to find the optimal subset of features and improve the prediction accuracy. Several bio-inspired hybrid methods were proposed such as ACO and GA hybridization [52, 53, 54, 55].

Basiri and Nemati [52] proposed a new hybrid ACO-GA feature selection algorithm for text categorization. The classifier performance and the length of selected feature subset are adopted as heuristic information. Sheikhan and Mohammadi [53] developed a hybrid model for short-term load forecasting. GA and ACO are combined in this model to explore the space of all subsets of given feature set and multi-layer perceptron is used for hourly load prediction. Sheikhan and Mohammadi [54] use a hybrid GA-ACO feature selection method to achieve a small and efficient feature subset using the IEEE load dataset.

Another application using hybrid methods based on ACO and artificial neural networks (ANNs) to find the optimal feature subset is presented in [56]. The proposed hybrid model is evaluated using medical diagnosis data sets.

Furthermore, Jin et al. [57] propose sort-based BPNN-PSO (BPNN: back propagation neural network) to select some critical attributes for improving the generalization performance and reducing computation cost of the BBNN. In the proposed feature selection method, input output correlation is applied for calculating feature importance.

Yang *et al* [58] improved binary particle swarm optimization was embedded in a genetic algorithm to serve as a local optimizer for each generation in the feature selection problem.

4.2. Swarm based hybrid approaches

Recently, many authors proposed hybrid approaches to solve several optimization problems in different application domains. Since we focus on ACO and PSO, we have studied some hybrid swarm based approaches using either ACO or PSO as part of their proposed hybridization. The table 3 presents our summary of some recent swarm based hybrid approaches.

5. Proposed hybrid ACO-PSO based feature selection approaches

In this section, we describe the three proposed feature selection approaches based on PSO-ACO hybridization. In the proposed approaches, we have used the MMAS variant of ACO algorithm (see section 2.1) and the naïve Bayes classifier.

Table 3. Summary of some swarm based hybrid approaches

Authors and reference	Proposed hybridization	Hybridization principle	Application domain
Juang, 2004 [59]	PSO- GA	The individuals in the new generation are created, not only by crossover and mutation operation as in GA, but also by PSO.	Recurrent network design problem
Shi *et al.*, 2005[60]	PSO-GA	Run the two algorithms (PSO, GA) simultaneously and select P individuals from each system for exchanging after the designated N iterations	Optimization of nonlinear functions
Li *et al.*, 2008, [61]	GA-PSO	After ranking the individuals in the actual generation the top individuals are selected as the elites and reproduce them directly to the next generation, the individuals followed are evolved with PSO and their best positions are updated and the bottom individuals are evolved with IGA.	Antenna array pattern synthesis
Valdez *et al.*, 2008 [62]	PSO-GA	Use fuzzy logic to integrate the results of both ACO and PSO methods.	Mathematical function optimization
Niknam and Amiri, 2010 [63]	Fuzzy adaptive PSO-ACO and Kmeans	-The selection of the Gbest particle for each individual is according to the ACO best path selection methodology. -PSO–ACO algorithm is used as the initial condition of the kmeans algorithm.	Nonlinear partitional clustering problem
Shuang *et al.*, 2011 [7]	ACO-PSO	The search mechanism of PSO algorithm is introduced into ACO algorithm, with combining the local exploration and the global exploitation into the pheromone update rules of ACO.	Travelling salesman problem
Kiran *et al.*, 2012 [64]	ACO-PSO	The best solution is assigned to global best solution of the system according to comparing best solutions which are found by ACO and PSO	Turkish energy demand estimation
Mahi *et al.*, 2015 [65]	ACO-PSO	The PSO algorithm is used for detecting optimum values of parameters α and β which are used for city selection operations in the ACO algorithm and determines significance of inter-city pheromone and distances	Traveling salesman problem

5.1. ACO-PSO1 approach

ACO-PSO1 (Fig. 1) is an improvement of the approach proposed by Nemati *et al* [55] where our idea is to replace genetic algorithms (GA) by PSO.

In [55], the authors present a hybrid feature selection algorithm based on ACO and GA hybridization. ACO offers a critical advantage of local searching, not found in GA. On the other hand, GA considers a global perspective by operating on the complete population from the very beginning. Therefore, ACO and GA can nullify each other's drawbacks when hybridized.

Using genetic algorithm to perform feature selection requires much iteration. If the learning algorithm is an iterative process, the computational cost of the whole process would be very expensive, since learning algorithm is used, in each iteration, to evaluate the fitness for each individual.

Like GA, PSO is also an evolutionary algorithm. Compared to GA, PSO does not need complex operators as crossover and mutation that GA does, it requires only primitive and simple mathematical operators, and is computationally inexpensive in terms of both memory and time [66, 67]
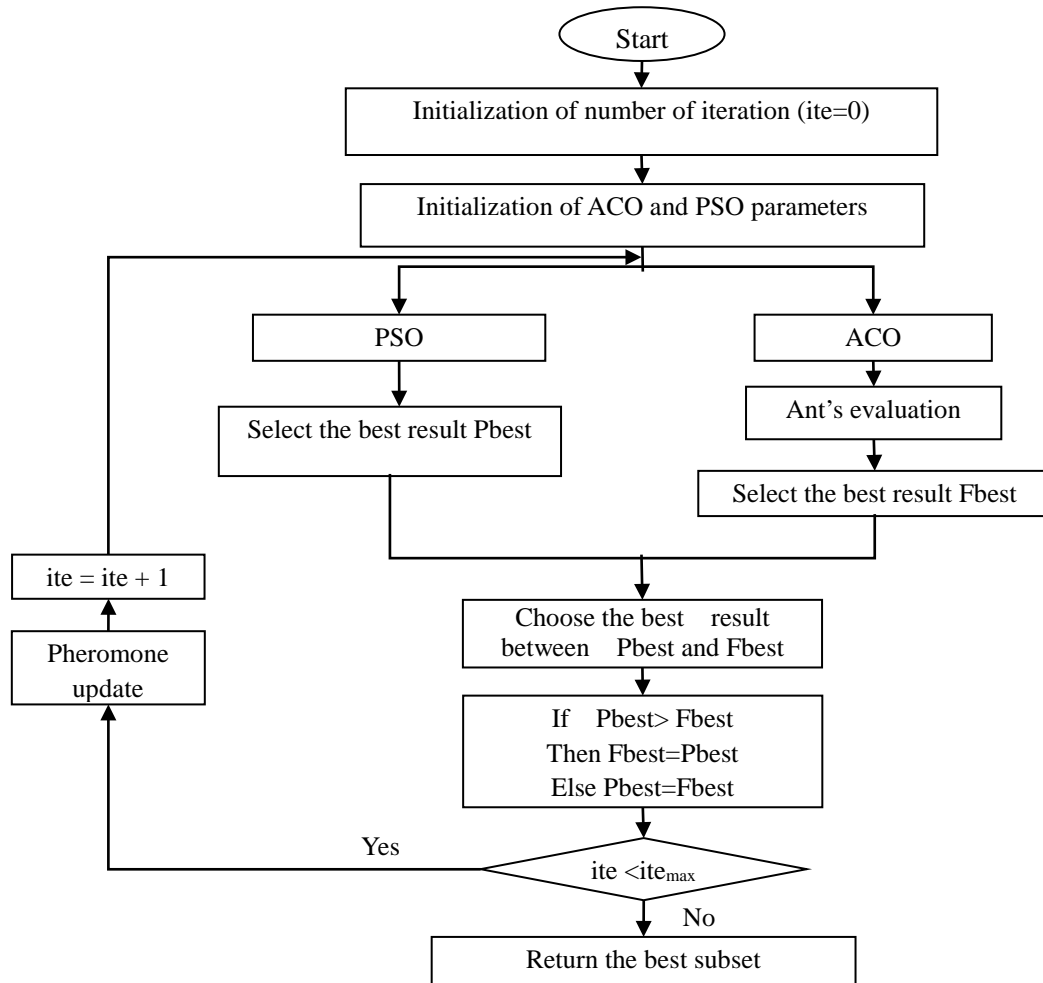
Figure 1. ACO-PSO1 hybrid approach

The main steps of ACO-PSO1 are as follows:
Step 1: Initialization.
- – Initialize the PSO parameters: Generate the position and the velocity of each particle randomly.
- Initialize the ACO parameters:
  - – Determine the population of ants. p equals the initial number of features.
  - – Set the intensity of pheromone trail associated with each feature.
- Determine the maximum of allowed iterations.

Step 2: Run ACO and PSO in the same time
- ACO: Generation of ants and evaluation of each one.
  - – Each ant (Ai, i = 1, 2, ..., p) is randomly assigned to one feature and it should visit all features and build solutions.
  - – In this step, the classifier error is used as an evaluation measure.
  - – At each iteration, the ants follow these

steps to build a subset:
- o At first, the subset contains only one feature (the search strategy used for feature selection is forward strategy).
- o **Calculate la probability of the kth ant making the transition from feature i to feature j

$$p_{ij}^k(t) = \frac{[\tau_{ij}(t)]^\alpha \cdot [\eta_{ij}]^\beta}{\sum_{l \, \epsilon \, J_i^k} [\tau_{il}(t)]^\alpha \cdot \, [\eta_{il}]^\beta} \quad (13)$$

With $\tau_{ij}(t)$ the pheromone quantity, η_{ij} is the heuristic desirability of choosing feature j when at feature i.

$$\eta_{ij} = \frac{1}{corr_{ij}} \quad (14)$$

Where $corr_{ij}$ define the correlation between features i and j.
- o Add a feature that has the max value of $p_{ij}^k(t)$ subset.
- o Test the classification error of current subset. If the error is reduced compared to the original error (the

error found with the full set of features), the ant terminates its traversal and outputs this feature subset as a solution. Otherwise goto **.

- PSO: Particles fitness evaluation
  - The predictive accuracy is used to measure the fitness of an individual.

$$fitness_p^t = \frac{1}{BER_p^t}$$

  Where BER_p^t is the achieved classification error using the subset selected by particle p at the iteration t.
  - The best fitness value for each particle I; s pbest defines the best fitness value found by the particle over all previous iterations.

$$pbest_p^t = Max(fitness_p^t)$$

  - Best fitness value within a group of pbest is the global fitness value gbest.

$$gbest_p^t = Max(pbest_p^t)$$

  - Update the positions and the velocity of each particle.

Step 3: Evaluation of the selected subset.
- Evaluate the importance of the selected subset of each ant (Fbest) and particles (Pbest) through the classifier performance and select the best one.
  - If Pbest> Fbest $\rightarrow$ Fbest=Pbest
  - Else $\rightarrow$ Pbest=Fbest

Step 4: Check the stop criterion.
- If the number of iterations is greater than the maximum iteration, exit, otherwise continue.

Step 5: Pheromone updating, and goto step2.

5.2. ACO-PSO2 approach

The fundamental idea of this hybrid approach is that the ACO is employed at the beginning of the searching process. When the fitness function value is not changed for several generations, or change value is very small (for example: less than 0.0001), the searching process is switched to the PSO algorithm; the best subsets found by the ants population of the ACO is the initial population of PSO.

The overall process of ACO-PSO2 approach can be seen in fig. 2. The process begins by generating a number of ants equal to the number of features within the data, these ants are then placed randomly

on the graph (i.e. each ant starts with one random feature). Each ant starts path construction at a different feature. From these initial positions, they traverse the edges probabilistically (using the equation11) until a traversal stopping criterion is satisfied (the current classifier error is less than the classifier error found with the features full set). The resulting subsets are gathered and then evaluated. If a certain number of iterations is reached, or the fitness function value is not changed for some generations, or change value is very small, then the searching process is switched to the PSO. Otherwise, the pheromone is updated, a new set of ants is created and the process iterates once more.

In this step, the position and velocity of each particle are generated according to the best results found by the ants. The position is represented by binary string form; the bit values 1 and 0 represent a selected and not-selected feature, respectively.

The predictive accuracy is used to measure the fitness of an individual. The best fitness value for each particle is pbest defines the best fitness value fined by the particle over all previous iterations and the best fitness value within a group of pbest is the global fitness value gbest. After fitness evaluation, the position and velocity of each particle are updated. If the algorithm reaches a certain number of iterations or the best fitness value is not changed after 10 successive iterations, then the process halts and outputs the best feature subset encountered during all iterations. If the conditions are not fulfilled, the process iterates once more.

5.3. ACO-PSO3 approach

Our ACO-PSO3 approach is an adaptation of the PS-ACO algorithm proposed by Shuang et al, [7] to the feature selection problem. The authors [7] presented a new hybrid algorithm integrating the ideas of PSO and ACO algorithms and tested it in the well-known travelling salesman problem. In this new algorithm, the pheromone updating rules of ACO are combined with the local and global search mechanisms of PSO. On one hand, the search space is expanded by the local exploration; on the other hand, the search process is directed by the global experience.

In PS-ACO, the pheromone trails increment $\Delta\tau_{ij}$ is improved with the idea of PSO algorithm and it is defined by:

$$\Delta\tau_{ij} = \sum_{k=1}^{n}\Delta\tau_{ij} + c1.r1.\sum_{k=1}^{n}\Delta\tau_{ij}^{local\ k} + \ldots$$
$$\ldots + c2.r2.\Delta\tau_{ij}^{global}$$
$$(15)$$

where $\Delta\tau_{ij}$ has the same definition as (5),

$\Delta\tau_{ij}^{local\ k}$ is the local update value of ant k at edge (vi, vj), and it is defined by:

$$\Delta\tau_{ij}^{local\ k} = \frac{Q}{f_{best\_so\_far}^{k}} \qquad (16)$$

where $f_{best\_so\_far}^{k}$ is the fitness value of the best solution toured by ant k from the beginning.

$\Delta\tau_{ij}^{global}$ is the global update value of edge (vi, vj) which is given by:

$$\Delta\tau_{ij}^{global} = \frac{Q}{f_{best\_so\_far}} \qquad (17)$$

where $f_{best\_so\_far}$ is the fitness value of the best solution toured by all ants from the beginning.
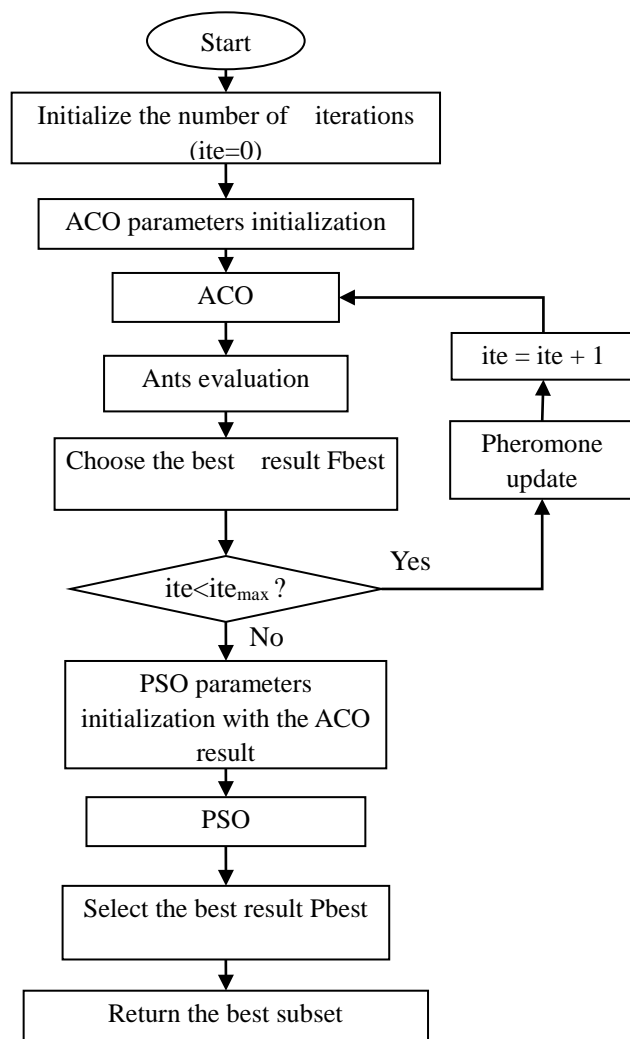

Figure 2. ACO-PSO2 hybrid approach

6. Experimental results

To evaluate the proposed hybrid approaches, we have conducted a series of experiments. In our earlier works, we have proposed simple bio-inspired feature selection approaches such as the ACO with its different variants and the PSO.

In our experiments we have used the Naïve Bayes classifier to evaluate the performance of the selected subsets. The spider package (object-oriented machine learning package) has been used for classification [68].

The parameters of ACO and PSO were determined experimentally. They are presented in the tables 4and 6 respectively and are used in the hybrid proposed approaches. The maximum number of iterations is 50. It was determined experimentally.

Table 4. ACO parameters

Parameters	Value
Number of ants	Features number in the database
Number of nodes	Features number in the database
α	1
β	5

Table 5. PSO parameters

Parameters	Value
Swarm size	Features number in the database
W	0.6571
c1	1.6319
c2	0.6239

6.1. Databases

Recently the majority of feature selection works are evaluated on public classifications databases. The most popular public databases collection is the UCI Machine Learning Repository [69].

Table 6. Databases Description

Database	Features	Features types	Samples
SpamBase	57	Integer, Real	4601
Breast Cancer	9	Integer	699
German	24	Categorical, Integer	1000
Hepatitis	19	Categorical, Integer, Real	155
Liver	6	Categorical, Integer, Real	345
Musk	166	Integer	476

Six datasets available at the UCI machine learning repository are used in all of our

experiments. These datasets are: Spambase, Breast cancer, German, Hepatitis, Liver and Musk. Table 6 presents a brief description of these databases.

6.2. Results

In order to evaluate the performance of the proposed hybrid feature selection approaches, a comparison is made with the ACO and PSO based feature selection performances obtained in our previous works [70]. Table 7-12 present the results in terms of number of selected features and classification error for the six databases. We have used the Balanced Error Rate (BER is the average of the error rate of the positive class and the rate of the negative class) to evaluate the performance of the hybrid approaches.

From tables 7-12, we can observe that all hybrid approaches have produced significant results both in terms of reduction of the selected features number and improvement of classification performance on all of the used databases. The hybrid ACO-PSO1 approach has achieved the best results in terms of classification performance with the majority of databases, followed by the simple PSO approach and ACO-PSO 2. The simple PSO approach produced the best results with 2 databases (German and Musk). However, the better error reduction was achieved for the Spambase. Except the German and Spambase datasets, the hybrid ACO-PSO2 didn't produce a better result than the simple ACO approach.

Table 7. Performance comparison between various feature selection approaches for Spambase database

Method	BER	Nbr of selected features	The selected features
Before selection	0.3128	57 (all features)	/
ACO	0.1992	8	36, 47, 25, 41, 48, 33, 39, 27
PSO	0.1545	16	2, 15, 24, 25, 26, 27, 30, 31, 32, 33, 35, 38, 42, 43, 46, 48
ACO-PSO1	0.1362	34	7, 8, 20, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 37, 38, 39, 40, 41, 45, 47, 50, 51, 53
ACO-PSO2	0.1937	21	2, 4, 8, 12, 14, 25, 27, 30, 32, 33, 34, 36, 38, 40, 43, 44, 45, 46, 47, 48, 55
ACO-PSO3	0.1951	15	45, 38, 30, 47, 36, 14, 55, 2, 43, 27, 39, 33, 48, 41, 25

Table 8. Performance comparison between various feature selection approaches for Breast Cancer database

Method	BER	Nbr of selected features	The selected features
Before selection	0.4433	9 (all features)	/
ACO	0.3033	2	9, 1
PSO	0.3483	2	1, 5
ACO-PSO1	0 .2550	2	1, 7
ACO-PSO2	0.3033	2	1, 9
ACO-PSO3	0 .3033	2	9, 1

Table 9. Performance comparison between various feature selection approaches for German database

Method	BER	Nbr of selected features	The selected features
Before selection	0.4990	24 (all features)	/
ACO	0,3680	2	1, 9
PSO	0.2775	8	1, 2, 8, 13, 14, 16, 21, 22
ACO-PSO1	0.2825	8	1, 4, 9, 12, 14, 16, 23, 24
ACO-PSO2	0.2855	4	1, 4, 16, 21
ACO-PSO3	0.3355	2	4, 1

Table 10. Performance comparison between various feature selection approaches for Hepatitis database

Method	BER	Nbr of selected features	The selected features
Before selection		19 (all features)	/
ACO	0.3836	2	5, 7
PSO	0.2760	7	2, 5, 9, 12, 15, 17, 19
ACO-PSO1	0.2701	3	1, 2, 5
ACO-PSO2	0.2930	8	4, 5, 7, 10, 11, 15, 16, 17
ACO-PSO3	0.3836	3	9, 7, 5

Table 11. Performance comparison between various feature selection approaches for Liver database

Method	BER	Nbr of selected features	The selected features
Before selection	0.5	6 (all features)	/
ACO	0.5	6	1, 2, 3, 4, 5, 6
PSO	0.4182	1	5
ACO-PSO1	0.4010	4	2, 3, 4, 5
ACO-PSO2	0.4071	5	2, 3, 4, 5, 6
ACO-PSO3	0 .5	6	1, 2, 3, 4, 5, 6

Table 12. Performance comparison between various feature selection approaches for Musk database

Method	BER	Nbr of selected features	The selected features
Before selection	0.3725	166(all features)	/
ACO	0.2510	16	25,147,90,75,62,111,9,132,18,165,103,36,73,137,104,163
PSO	0.2277	26	1,15,16,18,20,23,26,36,37,48,50,51,52,53,60,67,78,80, 82,92,94,97,103,104,106,124,129,132,136,145,147,151, 155,160
ACO-PSO1	0.2474	52	1,6,8,12,13,17,24,25,28,29,32,34,46,49,50,52,53,54,56, 73,74,76,83,84,91,96,97,98,102,104,106,108,115,116, 122,123,124,129,131,133,134,135,140,147,148,150,152, 155,161,162,165,166
ACO-PSO2	0.2309	28	8,11,14,15,18,36,38,40,59,62,67,73,75,90,94,104,126, 130,131,132,135,137,138,144,147,161,163,164
ACO-PSO3	0.2527	41	15,32,38,161,40,164,152,147,90,75,62,107,70,5,149,106, 127,59,79,125,20,92,19,157,29,103,165,83,67,4,135,130, 18,132,143,104,137,73,36 ,14,163

7. Conclusion and perspectives

In this paper, we presented bio-inspired hybrid approaches for feature selection. The proposed approaches are based on the integration of ant colony optimization (ACO) and particle swarm optimization (PSO) and their performances are compared with simple ACO and PSO based feature selection methods. The proposed approaches are tested on six public databases.

Our experimental studies show that the proposed approaches can efficiently reduce the number of features and improve the classification accuracy. The ACO-PSO1 can achieve the best result in terms of classification accuracy rate on the majority of databases.

More studies can be done in the future, although the proposed algorithm showed significant performances using naïve bayes classifier, the

classification accuracy may be improved using other classifiers like support vector machine. Furthermore, a variety of novel swarm intelligence algorithms such as those based on bee colonies, fireflies or bats were successfully applied to many optimization problems. We plan to use some of these algorithms in feature selection and we will also evaluate other interesting swarm based hybridizations.

References

[1] K. Kenji, and A.R. Larry, "Learning Inductive: "The feature selection problem: traditional methods and new algorithm", *AAAI-92 Proceedings*, pp. 128-134, 1992.

[2] M. Dash, and H. Liu, "Feature Selection Methods for Classification", *Intelligent Data Analysis: An International Journal,* Vol. 1, No. 3, pp. 131-156, 1997.

[3] H. Liu, and L. Yu, "Toward Integrating Feature Selection Algorithms for Classification and Clustering", in *Knowledge and Data Engineering,* Vol. 17, No. 4, pp. 491-502, 2005.

[4] D. Cakmakov, and Y. Bennani, *Feature selection for pattern recognition*, Skopje, Informa, 2002.

[5] Y. Liu, Z. Qin, Z. Xu, and X. He, "Feature Selection with Particle Swarms", In: J. Zhang, J.-H. He, and Y. Fu, (Eds). *Computational and Information Science*, Heidelberg: Springer-Verlag, Vol. 3314, pp. 425–430, 2004.

[6] M.A. Hall, *Correlation-based Feature Selection for Machine Learning*. PhD thesis, Dept. of Computer Science, Univ. of Waikato, Hamilton, New Zealand, 1998.

[7] B. Shuang, J. Chen, and Z. Li, "Study on hybrid PS-ACO algorithm", *Applied Intelligence*, Vol. 34, No.1, pp. 64-73, 2011.

[8] M. Dorigo, and L. M. Gambardella, "A study of some properties of Ant-Q", In PPSN IV: *Proceedings of the 4th international conference on parallel problem solving from nature*, pp. 656–665, 1996.

[9] R.J. Mullen, D. Monekosso, S. Barman, and P. Remagnino, "A review of ant algorithms", *Expert Systems with Applications,* Vol. 36, No. 6, pp. 9608-9617, 2009.

[10] M. Dorigo, and L.M. Gambardella, "Ant Colony System: a cooperative learning approach to the traveling salesman problem", *IEEE Trans. Evol. Comput.*, Vol. 1, No. 1, pp. 53-66, 1997.

[11] T. Stützle, and H.H. Hoos, "MAX–MIN ant system", *Future Generation Computer Systems*, Vol.16, No.8, pp. 889-914, 2000.

[12] B. Bullnheimer, R.F. Hartl, and C. Strauss, "A new rank-based version of the ant system: a computational study", C*entral European Journal for Operations Research and Economics* 7, pp. 25-38, 1999.

[13] L.M. Gambardella, and M. Dorigo, "Solving symmetric and asymmetric TSPs by ant colonies", In *Proc. of IEEE Int. Conf. on Evolutionary Computation*, New York, pp. 622-627, 1996.

[14] C.C. Lai, C-H. Wu, and M-C. Tsai, "Feature selection using particle swarm optimization with application in spam filtering", *International Journal of Innovative Computing, Information and Control.*, Vol. 5, No. 2, pp. 423-432, 2009.

[15] B. Panigrahi, Y. Shi, and M. Lim, *Handbook of Swarm Intelligence: Concepts, Principles and Applications*, Springer, Berlin, Germany, 2011.

[16] A. Hidaka, and T. Kurita, "Non-Neighboring Rectangular Feature Selection using Particle Swarm Optimization", *19th International Conference on Pattern Recognition*, Tampa, FL, USA, pp. 1-4, 2008.

[17] J. Kennedy, and R.C. Eberhart, "A discrete binary version of the particle swarm algorithm", In: *Proceedings of the International Conf. on Systems, Man and Cybernetics*, Piscataway: IEEE Press, pp. 4104-4109, 1997.

[18] A. Zainal, M. Maarof, and S. Shamsuddin, "Feature Selection using Rough-DPSO in Anomaly Intrusion Detection", *ICCSA* , Vol. 4705, pp. 512-524, 2007.

[19] F. Moussouni, S. Brisset, and P. Brochet, "Comparison of Two Multi-Agent Algorithms: ACO and PSO for the Optimization of a Brushless", DC Wheel Motor, *Studies in Computational Intelligence*, Vol. 119, pp. 3-10, 2008.

[20] B. Chen, L. Chen, and Y. Chen, "Efficient ant colony optimization for image feature selection", *Signal Processing,* Vol.93, No. 6, pp. 1566-1576, 2013.

[21] M.M. Kabir, M. Shahjahan, and K. Murase, "A new hybrid ant colony optimization algorithm for feature selection", *Expert Systems with Applications,* Vol.39, No 3, pp. 3747-3763, 2012.

[22] M. Kabir, M. Shahjahan, K. Murase, "Ant Colony Optimization Toward Feature Selection", In H. Barbosa (Ed.), *Ant Colony Optimization - Techniques and Applications,* InTech, pp.3-44, 2013.

[23] H.R. Kanan, and K. Faez, "An improved feature selection method based on ant colony optimization (ACO) evaluated on face recognition system", *Applied Mathematics and Computation*, Vol. 205, No. 2, pp. 716-725, 2008.

[24] R.N. Khushaba, A. Al-Ani, A. AlSukker, and A. Al-Jumaily, "A Combined Ant Colony and Differential Evolution FS Algorithm", In: M.Dorigo, et al. (Eds.), *ANTS*, Vol. 5217, pp. 1-12, 2008.

[25] M.H. Aghdam, N. Ghasem-Aghaee, and M.E. Basiri, "Text feature selection using ant colony optimization", *Expert Systems with Applications*, Vol.36, No. 3, pp. 6843-6853, 2009.

[26] M.J. Meena, K.R. Chandran, A. Karthik, and A.V. Samuel, "An enhanced ACO algorithm to select features for text categorization and its parallelization", *Expert Systems with Applications,* Vol. 39, No.5, pp. 5861-5871, 2012.

[27] Y. Gomez, R. Bello, A. Nowe, E. Casanovas, and J. Taminau, "Multi-colony ACO and Rough Set Theory to Distributed Feature Selection Problem", In: S. Omatu, et al. (Eds.), *IWANN* 2009, Part II, Vol. 5518, pp. 458–461, 2009.

[28]C-L. Huang, "ACO-based hybrid classification system with feature subset selection and model parameters optimization", *Neurocomputing*, Vol.73, No 1-3, pp. 438-448, 2009.

[29]N. Abd-Alsabour, "Binary Ant Colony Optimization for Subset Problems", In: S. Dehuri et al. (Eds.); *Multi-objective Swarm Intelligence*, *Studies in Computational Intelligence,* Vol. 592, pp. 105-121, 2015.

[30]S. Kashef, and H. Nezamabadi-pour, "An advanced ACO algorithm for feature subset selection", *Neurocomputing,* Vol. 147, pp. 271-279, 2015.

[31]P. Moradi, and M. Rostami, "Integration of graph clustering with ant colony optimization for feature selection"*, Knowl. Based Syst*, Vol. 48, pp. 144-161, 2015.

[32]S. Tabakhi, and P. Moradi, "Relevance–redundancy feature selection based on ant colony optimization", *Pattern Recognition*, Vol. 48, No. 9, 2015.

[33]M.A. Esseghir, G. Goncalves, and Y. Slimani, "Adaptive Particle Swarm Optimizer for Feature Selection", In: C. Fyfe, et al. (Eds.), *IDEAL 2010*, Vol. 6283, pp. 226-233, 2010.

[34]V. Kothari, J. Anuradha, S. Shah, and P. Mittal, "A Survey on Particle Swarm Optimization in Feature Selection", In: P.V. Krishna, M.R. Babu, and E. Ariwa, (Eds.), *ObCom 2011, Part II,* Vol. 270, pp. 192–201, 2012.

[35]T. Niiniskorpi, M.B. Aberg, and J. Wessberg, "Particle swarm feature selection for fMRI pattern classification", *Proceedings of the 2nd ICBSSP (BIOSIGNALS),* Porto, Portugal, pp. 279-284, 2009.

[36]L.Y. Chuang, H.W. Chang, C.J. Tu, and C.H. Yang, "Improved binary PSO for feature selection using gene expression data", *Computational Biology and Chemistry,* Vol.32, No.1, pp 29-38, 2008.

[37]S. Nemati, and M.E. Basiri, "Particle Swarm Optimization for Feature Selection in Speaker Verification", In: C. Di Chio, et al. (Eds.), *EvoApplications 2010, Part I,* Vol. 6024, pp. 371-380, 2010.

[38]L.F. Chen, C-T. Su, K-H. Chen, and P-C. Wang, "Particle swarm optimization for feature selection with application in obstructive sleep apnea diagnosis", *Neural Comput & Applic*, Vol. 21, No. 8, pp. 2087-2096, 2011.

[39]M.E. Basiri, A.N. Ghasem, and M.H. Aghdam, "Using Ant Colony Optimization-Based Selected Features for Predicting Post-Synaptic Activity in Proteins", In: E. Marchiori, and J.H. Moore, (Eds.), *EvoBIO*, Vol. 4973, pp. 12-23, 2008.

[40]T. Deng, C. Yang, Y. Zhang, and X. Wang, "An Improved Ant Colony Optimization Applied to Attributes Reduction", In: B. Cao, C. Zhang, and T. Li, (Eds.) *Fuzzy Info. and Engineering*, Springer, Vol. 54, pp. 1-6, 2009.

[41]M.M. Kabir, M. Shahjahan, and K. Murase, "Involving New Local Search in Hybrid Genetic Algorithm for Feature Selection", In: C.S. Leung, M. Lee, and J.H. Chan, (Eds.), *ICONIP,* Part II, Vol.

5864, pp. 150-158, 2009.

[42]C. Bae, W.C. Yeh, Y.Y. Chung, and S.L. Liu, "Feature selection with Intelligent Dynamic Swarm and Rough Set", *Expert Systems with* Applications, Vol. 37, No. 10, pp. 7026-7032, 2010.

[43]Y. Chen, D. Miao, and R. Wang, "A rough set approach to feature selection based on ant colony optimization", *Pattern Recognition Letters,* Vol. 31, No. 3, pp. 226-233, 2010.

[44]P., Shunmugapriya, S. Kanmani, S. Devipriya, J. Archana, and J. Pushpa, "Investigation on the Effects of ACO Parameters for Feature Selection and Classification", In: V.V. Das, and J. Stephen, (Eds.), *CNC 2012*, Vol. 108, pp. 136-145, 2012.

[45]B. Xue, L. Cervante, L. Lin Shang, and M. Zhang, "A Particle Swarm Optimisation Based Multi-objective Filter Approach to Feature Selection for Classification", In: P. Anthony, M. Ishizuka, and D. Lukose (Eds.), *PRICAI 2012,* Vol. 7458, pp. 673-685, 2012.

[46]X. Bing, Z. Mengjie, and N.B. Will, "Novel Initialisation and Updating Mechanisms in PSO for Feature Selection in Classification"*, EvoApplications, LNCS,* Vol. 7835, pp. 428-438, 2013.

[47]L. Cervante, B. Xue, L. Shang, and M. Zhang, "A Multi-Objective Feature Selection Approach Based on Binary PSO and Rough Set Theory", In: M. Middendorf, and C. Blum, (Eds.), *EvoCOP 2013*, Vol. 7832, pp. 25–36, 2013.

[48]S.M. Vieira, L.F. Mendonca, G.J. Farinha, J.M.C. Sousa, "Modified binary PSO for feature selection using SVM applied to mortality prediction of septic patients", *Applied Soft Computing,* Vol. 13, No. 8, pp. 3494-3504, 2013.

[49]B. Xue, "Zhang, M. and Browne, W.N. "Novel Initialisation and Updating Mechanisms in PSO for Feature Selection in Classification", *EvoApplications,* Vol. 7835, pp. 428–438, 2013.

[50]H.H. Inbarania, A.T. Azarb, and G. Jothic, "Supervised hybrid feature selection based on PSO and rough sets for medical diagnosis", *computer methods and programs in biomedicine,* Vol. 113, No. 1, pp. 175-185, 2014.

[51]B.H. Nguyen, B. Xue, I. Liu, and M. Zhang, "Filter based Backward Elimination in Wrapper based PSO for Feature Selection in Classification", In *proc. of IEEE Congress on Evolutionary Computation (CEC)* July 6-11, 2014, Beijing, China, pp. 3111-3118, 2014.

[52]M.E. Basiri, and S. Nemati, "A novel hybrid ACO-GA algorithm for text feature selection", In *Proc. of IEEE Congress on Evolutionary Computation*, pp. 2561-2568, 2009.

[53]M. Sheikhan, and N. Mohammadi, "Neural-based electricity load forecasting using hybrid of GA and ACO for feature selection", *Neural Comput Appl*, Vol. 21, No. 8, pp. 1961-1970, 2011.

[54]M. Sheikhan, and N. Mohammadi, "Time series prediction using PSO-optimized neural network and hybrid feature selection algorithm for IEEE load data", *Neural Comput & Applic*, Vol. 23, No. 3, pp.

1185-1194, 2012.

[55] S. Nemati, M.E. Basiri, N. Ghasem-Aghaee, and M.H. Aghdam, "A novel ACO–GA hybrid algorithm for feature selection in protein function prediction", *Expert Systems with Applications*, Vol. 36, No.10, pp. 12086-12094, 2009.

[56] R.K. Sivagaminathan, and S. Ramakrishnan, "A hybrid approach for feature subset selection using neural networks and ant colony optimization", *Expert Systems with Applications*, Vol. 33, No. 1, pp. 49-60, 2007.

[57] C. Jin, S-W. Jin, and L-N. Qin, "Attribute selection method based on a hybrid BPNN and PSO algorithms", *Applied Soft Computing*, Vol. 12, No. 8, pp. 2147-2155, 2012.

[58] C.S. Yang, L-Y. Chuang, C-H. Ho, and C-H. Yang, "Microarray Data Feature Selection Using Hybrid GA-IBPSO", In: O. Castillo et al. (eds.), *Trends in Intelligent Systems and Computer Engineering*, Springer, Vol. 6, pp. 243-253, 2008.

[59] C-F. Juang, "A hybrid of genetic algorithm and particle swarm optimization for recurrent network design", *IEEE Trans. Syst.*, Man, Cybern. B, Cybern., Vol. 34, No. 2, pp. 997–1006, 2004.

[60] X.H. Shi, Y.C. Lianga, H.P. Leeb, C. Lub, and L.M. Wang, "An improved GA and a novel PSO-GA-based hybrid algorithm", *Information Processing Letters*, Vol. 39, No.5, pp. 255-261, 2005.

[61] W. T. Li, X. Shi, W.L. Xu, and Y.Q. Hei, "Improved GA and PSO culled hybrid algorithm for antenna array pattern synthesis", *Progress In Electromagnetics Research*, Vol. 80, pp. 461-476, 2008.

[62] F. Valdez, P. Melin, and O. Castillo, "A New Evolutionary Method Combining Particle Swarm Optimization and Genetic Algorithms Using Fuzzy Logic". In: O. Castillo, et al. (Eds.), *Soft Computing for Hybrid Intel. Systems*, Vol. 154, pp. 347-361, 2008.

[63] T. Niknam, and B. Amiri, "An efficient hybrid approach based on PSO, ACO and k-means for cluster analysis", *Applied Soft Computing*, Vol. 10, No. 1, pp. 183-197, 2010.

[64] M.S. Kiran, E. Ozceylan, M. Gunduz, and T. Paksoy, "A novel hybrid approach based on Particle Swarm Optimization and Ant Colony Algorithm to forecast energy demand of Turkey", *Energy Conversion and Management,* Vol. 53, No. 1, pp. 75-83, 2012.

[65] M. Mahi, Ö.K. Baykan, H. Kodaz, "A new hybrid method based on Particle Swarm Optimization, Ant Colony Optimization and 3-Opt algorithms for Traveling Salesman Problem", *Applied Soft Computing,* Vol. 30, pp 484-490, 2015.

[66] X. Wang, J. Yang, N. Peng, and X. Teng, "Finding Minimal Rough Set Reducts with Particle Swarm Optimization", *Rough Sets, Fuzzy Sets, Data Mining, and Granular Computing*, Vol. 3641, pp. 451-460, 2005.

[67] X. Wang, J. Yang, X. Teng, W. Xia, and R. Jensen, "Feature selection based on rough sets and particle swarm optimization", *Pattern Recognition Letters*, Vol. 28, No. 4, pp. 459-471, 2007.

[68] Spider: object-oriented machine learning library. [online] http://people.kyb.tuebingen.mpg.de/spider/

[69] UCI Machine Learning Repository. [online] http://archive.ics.uci.edu/ml/datasets

[70] K. Menghour, and L. Souici-Meslati, "Classical and swarm based approaches for feature selection in spam filtering", *International Journal of Advanced Intelligence Paradigms, IJAIP*, InderScience Publishers, Vol. 6, No. 3, pp. 214-234, 2014.