Journal of Computational Information Systems 7: 13 (2011) 4756-4763
Available at http://www.Jofcis.com

First Study of Fuzzy Cognitive Map Learning Using Ants Colony Optimization

Zhengjian DING†, Dai LI†, Juanjuan JIA

School of Computer and Communication, Lanzhou University of Technology, Lanzhou 730050, China

Abstract

An approach for Fuzzy Cognitive Maps (FCMs) learning, which is based on the minimization of a properly defined objective function using the Ants Colony Optimization algorithm, is presented. The proposed approach is used in a well-established process control industry issues, and the results are availability. This novel approach overcomes some deficiencies of other learning algorithms and enriches the swarm intelligence algorithms in FCMs learning, thus, improves the efficiency and robustness of Fuzzy Cognitive Maps.

Keywords: Fuzzy Cognitive Maps; Ants Colony Optimization; Swarm Intelligence; Soft Computing

1. Introduction

Fuzzy Cognitive Maps (FCMs) are a soft computing tool proposed by Kosko as an expansion of cognitive maps which are widely using to represent social scientific knowledge [1]. One of the most beneficial aspects of the FCMs is its potential for use in decision support as a prediction tool [5]. Given an initial state of a system, represented by a set of values of its constituent concepts, an FCM can simulate its evolution over time to predict its future behavior.

However, the mathematical proof of the established developments is still required further testing on systems for a higher degree of complexity. Moreover, the elimination of deficiencies, such as the abstract estimation of the initial weight matrix and relying on the subjective reasoning of experts' knowledge, will greatly improve the function of FCMs. Under this background, the development of learning algorithm is a kind of exciting research subject [8].

This paper puts forward a new method of FCM learning, which is the Ants Colony Optimization (ACO) method, based on a swarm intelligence algorithm. More specifically, this method we proposed is in the way of Papaious' PCO learning method[8], which is selected to determine the proper weight matrix in this system by minimizing the value of objective function. ACO is selected because of its efficiency and effectiveness on a plethora of applications in science and engineering, and its direct applicability. The

† Corresponding author.
Email addresses: dingzj1952@sina.com (Zhengjian DING).
† Corresponding author.
Email addresses: lidai0628@sina.com (Dai LI).

December, 2011

proposed method is used to illustrate the application in an industrial process control issues with its promising results.

2. Overview of Fuzzy Cognitive Map

An FCMs (such as the Figure. 1) consists of concepts nodes, C_i, $i = 1, \ldots, N$, where N is the total number of concepts. Each concept node represents a factor of the system, and it is characterized by a value $A_i \in [0, 1]$, $i = 1, \ldots, N$. The concepts are interconnected with weighted arcs, which imply the relationships among them. Each interconnection between two concepts C_i and C_j has a weight W_{ij}, which is proportional to the strength of the causal link between C_i and C_j. At each step, the value A_i of the concept C_i is influenced by the values of the concepts nodes connected to it, and it is updated according to the rule:

$$A_i^{(k+1)} = f\left(A_i^{(k)} + \sum_{\substack{j \neq i \\ j=1}}^{N} A_j^{(k)} w_{ji} \right), \tag{1}$$

Where the k stands for the iteration counter.

And $w_{ji} \in [0, 1]$ is the weight of the arc connecting the concept C_j to the concept C_i. There are three possible types of causal relationships among concepts [6, 8]:

- $w_{ji} > 0$, positive causality between concepts C_i and C_j;
- $w_{ji} < 0$, negative causality between concepts C_i and C_j;
- $w_{ji} = 0$, no relationship between concepts C_i and C_j.

The function f is the sigmoid function:

$$f(x) = \frac{1}{1 + e^{-\lambda x}}, \tag{2}$$

Where $\lambda > 0$ is a parameter that determines its steepness in the area around zero.

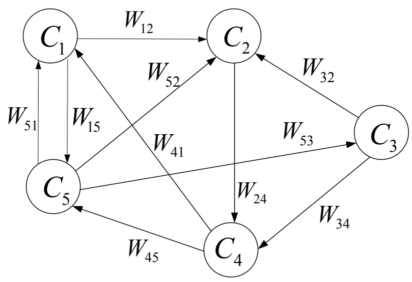

Fig.1 A Simple Fuzzy Cognitive Map

3. The Proposed Learning Method by ACO

3.1. The ACO Algorithm

Ant colony algorithm is optimization algorithm inspired by real ants foraging behavior in the wild. When looking for food, the preliminary exploring area nearby nest in a random manner. When an ant finds a food source, it calculates the quantity, quality of the food and carries some food back to the nest. In return, the ant release chemical pheromone trail on the ground. The quantity of pheromone deposited, which may lie on the quantity and quality of the food, will lead other ants to the food source. The communication between the ants indirectly, they can seek out the shortest path between home and a food sources by the pheromone trail. This feature of real ant colonies is exploited in artificial ant colonies in order to solve difficult combinatorial optimization problems. In ant colony algorithm, artificial ants probabilistically establish the solutions by taking into account dynamical artificial pheromone trails. The central component of ACO algorithm is the pheromone model including the state transition rule and updating rule, which is used to probabilistically sample the search space [2, 9]. The ACO problem can be defined as follows:

Definition 1. A model $P = (S, \Omega, T)$ of an ACO problem consists of Dorigo and Blum (2005)[10]:

●A search (or solution) space S is defined over a finite set of discrete decision variables and a set Ω of constraints among the variables;

●An objective function T**:** $S \rightarrow R^+$ to be minimized.

The search space S is defined as follows: Given a set of n discrete variables U_i with values $a_i^j \in D_i = \left\{a_i^1, \cdots, a_i^{|D_i|}\right\}$, $i = 1, \ldots, N$, a variable instantiation, that is, the assignment of a value a_i^j to a variable U_i, is denoted by $U_i = a_i^j$. A feasible solution $s \in S$ is a complete assignment (an assignment in which each decision variable has a domain value assigned) that satisfies the constraints. If the set of constraints Ω is empty, then each decision variable can take any value from its domain independently of the values of the other decision variables. In this case, we call P an unconstrained problem model, otherwise a constrained problem model. A feasible solution $s^* \in S$ is called a *global optimum*, if $T(s^*) \leq T(s)$, $\forall\, s \in S$. The set of global optimums is denoted by $S^* \subseteq S$. To solve an ACO problem, one has to find a solution $s^* \in S^*$. The basic framework of the ACO algorithm is shown in Table1. At the beginning of the algorithm, the ACO problem model is inputted with initializing the parameters. The basic key-component of any ACO algorithm is a constructive heuristic for probability construction solutions. At each iteration, ants use a given pheromone model to work out the probability construction solutions of the optimization problem. Then, some of the solutions are used for performing a pheromone update. Sometimes, daemon actions can be used to implement centralized actions, not by a single ant. Examples include the local search to the constructed solutions, or the global information collection which can be applied to determine whether it is available or not to deposit additional pheromone to bias the search process from a non-local viewpoint. The general structure of the ACO algorithm is as Table 1.

3.1. The ACO Algorithm for FCMs Learning

The main goal of this learning method[6, 8] is to detect a weight matrix $W = [\, w_{ij} \,]$, $i, j = 1, \ldots, m$, which lead the weight retain their physical meaning. Let $C_1, \ldots, C_N$ be the concept of an FCM, and let $C_{out_1}, \ldots, C_{out_m}$ $(1 \leq m \leq N)$ be the output concept, while the remaining concepts are considered input or interior concepts. We restricted the value of these output concepts in strict bounds,

$$A_{out_i}^{\min} \leq A_{out_i} \leq A_{out_i}^{\max}, i = 1, 2, \cdots, m\ ,$$

The *Objective Function F* is following [6, 8, 11]:

$$F(W) = \sum_{i=1}^{m} H\left(A_{out_i}^{\min} - A_{out_i}\right)\left|A_{out_i}^{\min} - A_{out_i}\right| + \sum_{i=1}^{m} H\left(A_{out_i} - A_{out_i}^{\max}\right)\left|A_{out_i}^{\max} - A_{out_i}\right|, \tag{3}$$

Where H is the well-known Heaviside Function:

$$H(x)=\begin{cases}0, & x<0,\\ 1, & x\geq 0,\end{cases} \tag{4}$$

and A_{out_i} , $i = 1,\ldots,m$, are the steady state value of the output concepts which are obtained by iteration. Obviously, the global minimum value of objective function F, are weight matrices that lead the FCM to a desired steady state , all output concepts are bounded within the desired regions. The ACO works for the problem that the global minimum value of objective function. The algorithm stops, when the best value is obtained. Flowchart of the proposed procedure is as figure 2.

This Flowchart is the same as Papaious' PCO learning method beside the learning algorithm. The data type, amount, way and objective function values of FCMs' I/O delivered are also the same as Papaious' method. So the question comparing the two methods of FCMs learning transforms to compare the two optimization algorithms. Moussouni et al. had done some deeper study on comparing the two algorithms. Comparing the solution's accuracy and the computation time, ACO is exacter and faster. However, compared with ACO, the superiority of PSO is its easier performance and defining fewer parameters [7]. In the following part, the simulation results are reported and analyzed.

Table 1 The Framework of a Basic ACO Algorithm

Input: An ACO problem model *P=(S, Ω, T)*.

Set Parameters
Initialize Pheromone Trails
Solution Construction
While(*Not TERMINATION_CONDITION*) {
  Schedule Activities
    Ant Based Solution Construction
    Daemon Actions { Optional }
    Pheromone Update
  End Schedule Activities
}
End While

Output: The obtained best solution so far.

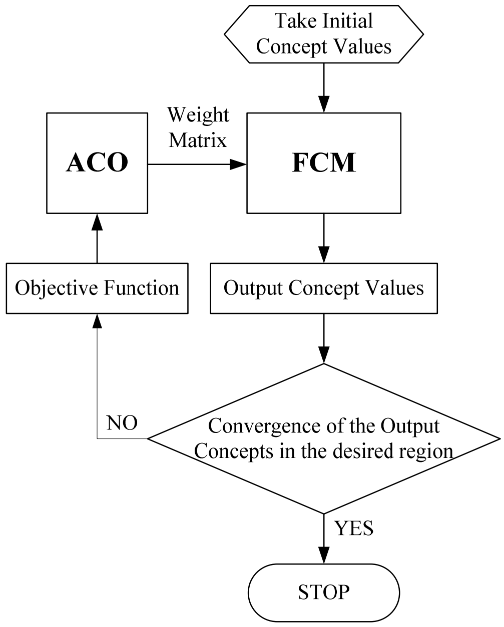

Fig.2 Flowchart of the Proposed Learning Procedure

4. Experiments and Results

A simple process control problem encountered in the chemical industry is selected to illustrate the workings of the proposed learning algorithm (Stylios and Groumpos [3]). Papaious' PCO learning method also takes this problem to illustrate its availability [6]. The ranges of the weights implied by the fuzzy regions, as they were suggested by experts, were:

$$\begin{aligned}&-0.50\leq W_{12}\leq -0.30,\quad -0.40\leq W_{13}\leq -0.20,\quad 0.20\leq W_{15}\leq 0.40,\quad 0.30\leq W_{21}\leq 0.40,\\ &0.40\leq W_{31}\leq 0.50,\quad -1.00\leq W_{41}\leq -0.80,\quad 0.50\leq W_{52}\leq 0.70,\quad 0.20\leq W_{54}\leq 0.40.\end{aligned} \qquad \text{Relation (1)}$$

Since the consideration of all eight constraints on the weights prohibits the detection of a suboptimal matrix, some of the constraints were omitted. More specifically, the constraints for the weights W_{12}, W_{13}, W_{15}, W_{52} and W_{54}, for which, the experts' suggestions regarding their values varied widely. So we consider that the corresponding weights be allowed to assume values in the range [－1, 0] or [0, 1], in order to avoid physically meaningless weight matrices[6]. While the three weights W_{21}, W_{31} and W_{41}, have little dispute about their range, we choose their range from Relation (1). At the same time, the output concepts for this problem are C_1 and C_5. The desired regions for the two output concepts, which are crucial for the proper operation of the modeled system, have been defined by the experts.

$$0.68 \le C_1 \le 0.70, \quad 0.78 \le C_5 \le 0.85. \qquad \text{Relation (2)}$$

The *Initial Concepts* are as follow:

$$C_1 = 0.10, \quad C_2 = 0.45, \quad C_3 = 0.39, \quad C_4 = 0.04, \quad C_5 = 0.01.$$

The results obtained through ACO were compared with that of PSO reported in [6]. We performed 100 independent experiments in very case. And we attempt to set the parameters, in order to make the iteration complete and seek out the best solution in 500 times function evaluations. The error goal for the optimization problem was set to 10^{-8}. In tables, statistics regarding the weights are reported severally in Case A, Case B, and Case C. The results suggest that ACO is a very promising approach for FCM learning.

The experiments have three cases, and the results are very interesting. At beginning, we consider that the ranges of the three weights W_{21}, W_{31} and W_{41}, are from Relation (1), in order to calculate the further range of them. And we reported the results in Table 2.

Table 2 Statistical Analysis for the Weights Obtained with PSO and ACO at Beginning

	PSO					***ACO***				
	Max	Min	Mean	Median	Stdev.	Max	Min	Mean	Median	Stdev.
W_{12}	−0.1000	−0.6662	−0.2832	−0.2389	0.1847	−0.0180	−0.2982	−0.1725	−0.1829	0.0347
W_{13}	−0.1000	−0.3965	−0.1595	−0.1166	0.0805	−0.0018	−0.2612	−0.0649	−0.0228	0.0653
W_{15}	1.0000	0.7143	0.9198	0.9611	0.0917	0.9340	0.7137	0.8042	0.7841	0.0489
W_{21}	0.4000	0.3753	0.3994	0.4000	0.0035	0.4000	0.3773	0.3916	0.3912	0.0038
W_{31}	0.5000	0.4679	0.4994	0.5000	0.0035	0.4989	0.4720	0.4857	0.4871	0.0082
W_{41}	−0.8000	−0.8048	−0.8000	−0.8000	0.0005	−0.8000	−0.8384	−0.8193	−0.8171	0.0075
W_{52}	1.0000	0.5196	0.8915	0.9599	0.1397	0.9919	0.6541	0.8416	0.8691	0.0797
W_{54}	0.2770	0.1000	0.1216	0.1000	0.0438	0.1147	0.0000	0.0452	0.0388	0.0390

After the iteration calculation, we find the three weights W_{21}, W_{31} and W_{41}, detected in the smaller range:

$$0.38 \le W_{21} \le 0.40, \quad 0.47 \le W_{31} \le 0.50, \quad -0.84 \le W_{41} \le -0.80. \qquad \text{Relation (3)}$$

We observed the range of them, and the two weights W_{21} and W_{31} have very similar range by each method, while the range of the W_{41} by ACO is little bigger than PSO.

- **Case A.** In this case, the three weights W_{21}, W_{31} and W_{41}, are constrained in the bounds defined in Relation (3), while the other weights are in [－1, 0] or [0, 1] without unconstrained. The statistics of the obtained results for this case are reported in Table 3.

Table 3 Statistical Analysis for the Weights Obtained with PSO and ACO for Case A.

	PSO					ACO				
	Max	Min	Mean	Median	Stdev.	Max	Min	Mean	Median	Stdev.
W_{12}	−0.1000	−0.5742	−0.2328	−0.1952	0.1408	−0.0008	−0.5517	−0.1741	−0.1737	0.0927
W_{13}	−0.1000	−0.3182	−0.1443	−0.1006	0.0651	−0.0009	−0.2828	−0.1028	−0.0900	0.0608
W_{15}	1.0000	0.7141	0.9106	0.9475	0.0976	0.9907	0.7109	0.7595	0.7399	0.0520
W_{21}	0.4000	0.4000	0.4000	0.4000	0.0000	0.4000	0.3802	0.3930	0.3934	0.0032
W_{31}	0.5000	0.4700	0.4991	0.5000	0.0051	0.5000	0.4748	0.4908	0.4909	0.0036
W_{41}	−0.8100	−0.8100	−0.8100	−0.8100	0.0000	−0.8001	−0.8339	−0.8105	−0.8101	0.0043
W_{52}	1.0000	0.6045	0.8828	0.9349	0.1295	0.9980	0.5432	0.8302	0.8337	0.0679
W_{54}	0.2324	0.1000	0.1132	0.1000	0.0297	0.2134	0.0001	0.0731	0.0682	0.0432

In this case, the three weights W_{21}, W_{31} and W_{41}, all have a very low Standard Deviation in the PSO and ACO (Table 3), and take almost same values in experiments. The range of the other weights we got by ACO is following:

$$\begin{aligned} &0.00 \le W_{12} \le -0.55, \quad 0.00 \le W_{13} \le -0.28, \quad 0.99 \le W_{15} \le 0.71, \\ &1.00 \le W_{52} \le 0.54, \quad 0.21 \le W_{54} \le 0.00. \end{aligned} \qquad \text{Relation (4)}$$

● **Case B.** The three weights W_{21}, W_{31} and W_{41}, we equaled to their Mean Values reported in Table 3 of Case A, while the remaining weights are constrained within the detected ranges in Relation (4). The statistics of the obtained results are reported in Table 4.

Table 4 Statistical Analysis for the Weights Obtained with PSO and ACO for Case B.

	PSO					ACO				
	Max	Min	Mean	Median	Stdev.	Max	Min	Mean	Median	Stdev.
W_{12}	−0.1000	−0.5700	−0.2366	−0.2220	0.0880	−0.0004	−0.4900	−0.1803	−0.1774	0.0595
W_{13}	−0.1000	−0.3040	−0.1349	−0.1223	0.0375	−0.0002	−0.2645	−0.0985	−0.0985	0.0306
W_{15}	1.0000	0.7147	0.8637	0.8674	0.0805	0.9898	0.7131	0.8546	0.8545	0.0406
W_{21}	0.4000	0.4000	0.4000	0.4000	0.0000	0.3900	0.3900	0.3900	0.3900	0.0000
W_{31}	0.5000	0.5000	0.5000	0.5000	0.0000	0.4900	0.4900	0.4900	0.4900	0.0000
W_{41}	−0.8100	−0.8100	−0.8100	−0.8100	0.0000	−0.8100	−0.8100	−0.8100	−0.8100	0.0000
W_{52}	1.0000	0.7226	0.9124	0.9223	0.0695	0.9998	0.5609	0.8395	0.8359	0.0576
W_{54}	0.2050	0.1000	0.1214	0.1166	0.0232	0.1547	0.0001	0.0584	0.0562	0.0219

In the Case B, we find the Standard Deviation of the weight W_{12}, W_{13}, W_{15}, W_{52} and W_{54}, by ACO is lower than that of PSO. Meanwhile they are all in an extensive range around their bounds by each method.

● **Case C.** In final case, the three weights W_{21}, W_{31} and W_{41}, are still equaled to the same values as in Case B, while the others are constrained within the ranges defined by their

$$\text{Mean Value} \pm \text{Standard Deviation}$$

Where the Mean Values and Standard Deviations are reported in Table 4. The obtained results are reported in Table 5, and box plots for the weights and concepts values are depicted in Figure 4. Furthermore, we made comparisons between PSO and ACO on Mean Value and Standard Deviations (Figure 3).

Table 5 Statistical Analysis for the Weights Obtained with PSO and ACO for Case C.

	PSO					***ACO***				
	Max	Min	Mean	Median	Stdev.	Max	Min	Mean	Median	Stdev.
W_{12}	−0.1525	−0.3177	−0.2268	−0.2298	0.0486	−0.1209	−0.2397	−0.1791	−0.1774	0.0197
W_{13}	−0.0976	−0.1714	−0.1321	−0.1320	0.0212	−0.0679	−0.1290	−0.0981	−0.0981	0.0100
W_{15}	0.9440	0.7848	0.8659	0.8654	0.0422	0.8952	0.8140	0.8548	0.8549	0.0134
W_{21}	0.4000	0.4000	0.4000	0.4000	0.0000	0.3900	0.3900	0.3900	0.3900	0.0000
W_{31}	0.5000	0.5000	0.5000	0.5000	0.0000	0.4900	0.4900	0.4900	0.4900	0.0000
W_{41}	−0.8100	−0.8100	−0.8100	−0.8100	0.0000	−0.8100	−0.8100	−0.8100	−0.8100	0.0000
W_{52}	0.9818	0.8433	0.9165	0.9127	0.0402	0.8970	0.7820	0.8408	0.8409	0.0188
W_{54}	0.1440	0.0983	0.1192	0.1180	0.0135	0.0802	0.0365	0.0578	0.0578	0.0073

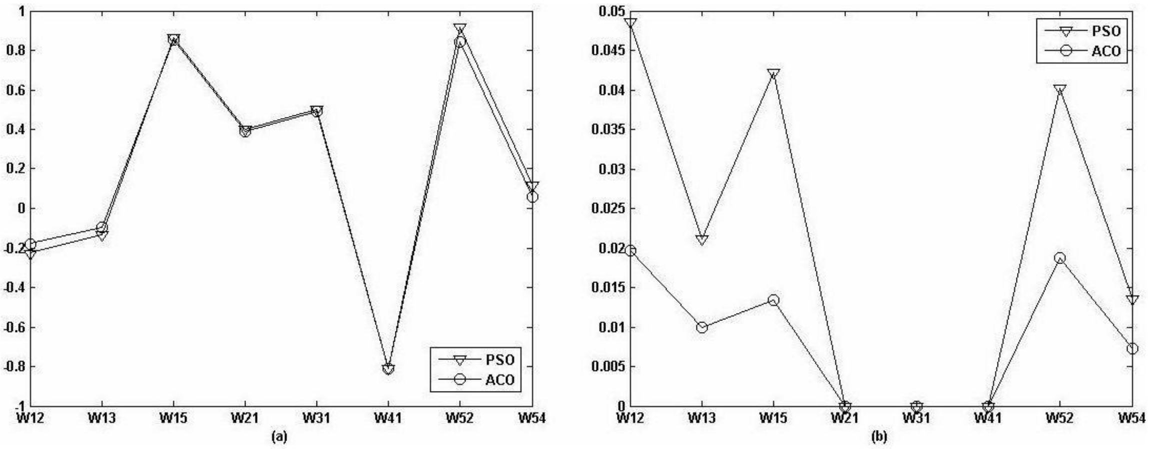

Fig.3 The Mean Values (a) and Standard Deviation (b) between PSO and ACO.

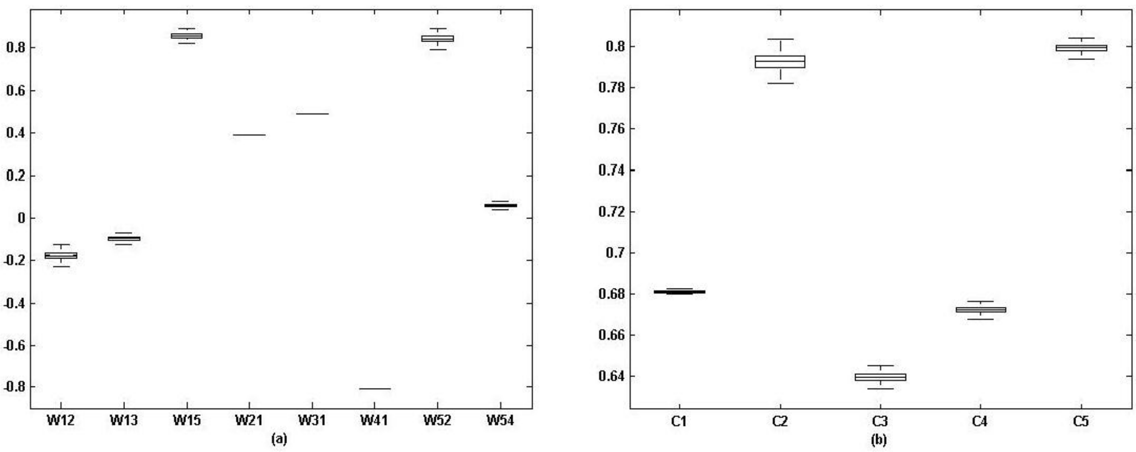

Fig.4 The Obtained Weights Values (a) and Concept Values (b) by ACO in Case C

In Figure 3, it's interesting that we can clearly observe the Mean Values of the weight W_{12}, W_{13}, W_{15}, W_{52} and W_{54}, are very similar by each method, and the values by ACO approach to zero. Furthermore, the

Standard Deviations of those weights by ACO are much smaller, so we also can clearly see that the obtained weights and concepts in convergence state are closer than that of PSO (Figure 4).

At the beginning of the whole experiments, the ACO need more iteration times find the best solution. With implementing the following experiments, it becomes faster and more accurate in fewer iteration times.

5. Conclusions and Further Directions

In this study, the Ants Colony Optimization algorithm is selected for Fuzzy Cognitive Maps learning, and demonstrating the feasibility and effectiveness of the swarm intelligence algorithms approach. It has been shown how ACO helps FCM detect the suboptimal weights.

The paper discusses relevant work, and proposes and tests a novel learning method, based on the ACO algorithm. The method is able to detect a suboptimal weight from the range provided by experts and led the concept go into the convergence state. First set of experiments aimed to detect the weights led the concept go into convergence state with the constraints of output concepts. In order to achieve this goal, the objective function was called. Later, a set of tests was performed to make the weights of FCMs close the Mean Values. The results show that the proposed learning method is very effective, and obtain weights can almost perfectly represent the experts' suggestion. In general, the proposed method achieved excellent quality for maps up to 8 nodes. Since the ACO algorithm for FCMs learning has been tested, the produced results could provide some guidelines for other swarm intelligence learning methods.

The future work will concern on the improvement of the proposed learning method, especially in terms of its scalability (computational complexity) and convergence.

References

[1] B. Kosko. Fuzzy Cognitive Maps. In *International Journal of Man-Machine Studies*, *24*, pages 65-75, 1986.

[2] C. Blum. Ant colony optimization: Introduction and recent trends. In *Physics of Life Reviews*, *2(4)*, pages 353-373, 2005.

[3] C. D. Stylios, P. P. Groumpos. The Challenge of Modeling Supervisory Systems Using Fuzzy Cognitive Maps. In *Journal Intelligent Manufacturing*, *9*, pages 339-345, 1998.

[4] C. M. Lin,Y. He, B. Y. Tang. Immune Genetic Learning of Fuzzy Cognitive Map. In *Journal of Donghua University(English Edition)* , *23(6)*, pages 10-14, 2006.

[5] E. I. Papageorgiou, C. D. Stylios, P. P. Groumpos. An integrated two-level hierarchical system for decision making in radiation therapy based on fuzzy cognitive maps. In *IEEE Transactions on Biomedical Engineering*, *50 (12)*, pages 1326-1339, 2003.

[6] E. I. Papageorgiou, K. E. Parsopoulos, C. D. Stylios, P. P. Groumpos, M. N. Vrahatis. Fuzzy cognitive maps learning using particle swarm optimization. In *Journal of Intelligent Information Systems*, *25:1*, pages 95-121, 2005.

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

[8] K. E. Parsopoulos, E. I. Papageorgiou, P P Groumpos, M. N. Vrahatis. A First Study of Fuzzy Cognitive Maps Learning Using Particle Swarm Optimization. In *Proceedings of the IEEE 2003 Congress on Evolutionary Computation*, pages 1440-1447, 2003.

[9] M. Dorigo, T. Stützle. *Ant Colony Optimization*. The MIT Press, 2004.

[10] M. Dorigo, C. Blum. Ant colony optimization theory: A survey. In *Theoretical Computer Science*, *344(2-3)*, pages 243-278, 2005.

[11] Z. Wang, B. Feng, J. Sun. Solve industry problem with novel fuzzy cognitive map learning algorithm. In *Computer Engineering and Applications*, *44(10)*, pages 217-220, 2008.