International Journal of Hybrid Information Technology
Vol.7, No.2 (2014), pp.29-38
http://dx.doi.org/10.14257/ijhit.2014.7.2.04

Ant Colony Optimization Algorithm Based on Dynamical Pheromones for Clustering Analysis

Xiaoyong Liu

Department of Computer Science, Guangdong Polytechnic Normal University, Guangzhou, Guangdong, 510665, China

liugucas@gmail.com

Abstract

This paper presents an improved clustering algorithm with Ant Colony optimization (ACO) based on dynamical pheromones. Pheromone is an important factor for the performance of ACO algorithms. Two strategies based on adaptive pheromones which improved performance are introduced in this paper. One is to adjust the rate of pheromone evaporation dynamically, named as ρ *, and the other is to adjust the strength of pheromone dynamically, named as* Q *. Two evaluation indices, Precision and Recall, are chosen to validity the improvement strategies. Numerical simulations demonstrate that the two strategies on pheromone can achieve better performance than basic ant colony algorithm and clustering algorithm with ant colony based on best solution kept.*

Keywords: *Ant Colony Algorithm, Clustering with ACO, Data Mining, Pheromone*

1. Introduction

Clustering analysis is an important method in the study of data mining. Clustering with ant colony optimization (ACO) is an active research field in recent years. The first clustering algorithm with ACO was provided by Deneubourgin 1991 [1]. The clustering algorithm simulates the foraging action of ants. Lumer and Faieta presented an improvement LF algorithm based Deneubourg's algorithm and applied to data analysis problem [2]. Labroche introduced a new algorithm that is based on the chemical recognition system of ants [3]. Ref [4] and Ref [5] proposed a new method, which combined ACO with K-means clustering. This algorithm firstly clustered datasets by K-means, and then updated pheromone according to the clustering result to guide artificial ants to choose path. Ref [6] used a fitness function to guide artificial ants clustering self adapting and self organization. Ref [7], Ref [8] and Ref [9] use ACOC to solve other problems, such as, macroscopic planning of highway transportation hub, module partition for mass customization, and intrusion detection. Ref [10] and Ref [11] presented an ant colony optimization methodology for optimally clustering N objects into K clusters. The new clustering algorithm employed distributed agents which mimic the way real ants find a shortest path from their nest to food source and back. Ref [12] provided a mechanism of the best solution kept to improve the performance of clustering analysis with ACO. In these work, pheromone evaporation factor ρ and pheromone intensity Q always are fixed parameters during the algorithm running. The chosen values of two parameters will affect the global search capability of ant colony algorithm, so this paper will adopt dynamical pheromones strategy to weaken the influence that is due to improper initial value of two parameters.

ISSN: 1738-9968 IJHIT
Copyright ⓒ 2014 SERSC

This paper is organized as follows. In Section 2 we first introduce Ant Colony System and clustering algorithm with ACO. We then discuss strategies of the two pheromone dynamic adjustment. Clustering algorithm with ant colony optimization based dynamical pheromones (RRACOC) is also presented. In Section 3, we illustrate our work using RRACOC algorithm. Then, experimental results are presented. Finally, Section 4 presents our conclusions and future work.

2. Clustering Algorithm with ACO based on Dynamic Pheromone

2.1. Ant Colony System

As shown in Figure 1, two ants start from their nest in search of food source at the same time to different directions. One of them chooses the path that turns out to be shorter while the other takes the longer sojourn. The ant moving in the shorter path returns to the nest earlier and the pheromone deposited in this path is obviously more than what is deposited in the longer path. Other ants in the nest thus have high probability of following the shorter route. These ants also deposit their own pheromone on this path. More and more ants are soon attracted to this path and hence the optimal route from the nest to the food source and back is very quickly established. Such a pheromone-meditated cooperative search process leads to the intelligent swarm behavior [5].

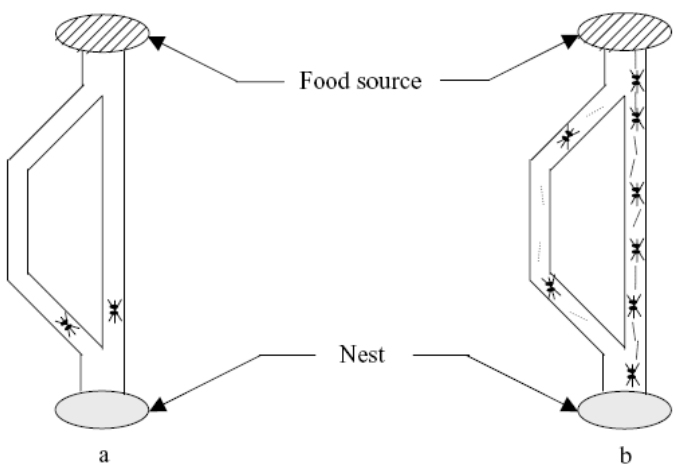

Figure 1. Movement of ant algorithm from nest-food source and back[5]

Ant colony system (ACS) algorithm was first proposed by Dorigo in 1997[7]. Informally, ACS works as follows: m ants are initially positioned on n cities chosen according to some initialization rule (*e.g.*, randomly). Each ant builds a tour (*i.e.*, a feasible solution to the TSP) by repeatedly applying a stochastic greedy rule (the state transition rule). While constructing its tour, an ant also modifies the amount of pheromone on the visited edges by applying the local updating rule. Once all ants have terminated their tour, the amount of pheromone on edges is modified again (by applying the global updating rule). As was the case in ant system, ants are guided, in building their tours, by both heuristic information (they prefer to choose short edges), and by pheromone information: An edge with a high amount of pheromone is a very desirable choice. The pheromone updating rules are designed so that they tend to give more pheromone to edges which should be visited by ants. In the general case, ACS

Copyright ⓒ 2014 SERSC

algorithm applies the artificial ants' concept, it is represented by the following steps:

Step 1: Initialization of parameters.

Step 2: Construction of solutions.

Step 3: Pheromone updating rule.

Step4: Return to Step 2 until a given stopping criterion satisfied.

2.1.1. State Transition Rule: In ACS the state transition rule is as follows: an ant positioned on node i chooses the city j to move to by applying the rule given by function (1).

$$p_{ij}^k(t) = \begin{cases} \dfrac{[\tau_{ij}(t)]^\alpha [\eta_{ik}(t)]^\beta}{\sum_{s \subset J_k(i)} [\tau_{is}(t)]^\alpha [\eta_{is}(t)]^\beta} & \text{if} \quad j \in J_k(i) \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

Where,

$p_{ij}^k(t)$ is the state transition probability with which the ant k moves from city i to city j in t-th iteration.

$\tau_{ij}(t)$ is pheromone value between city i and city j in t-th iteration.

$\eta_{ij} = \dfrac{1}{d_{ij}} \cdot$

$J_k(i)$ is the set of cities that ant k hasn't been visited when it is at city i.

α and β are the parameters and denote heuristic value.

2.1.2. Pheromone Updating Rule: While building a solution (*i.e.*, a tour) of the TSP, ants visit edges and change their pheromone level by applying the updating rule of function (2).

$$\tau_{ij}(t+1) = \rho \tau_{ij}(t) + \Delta \tau_{ij}(t) \tag{2}$$

Where,

$$\Delta \tau_{ij}(t) = \sum_{k=1}^{N} \Delta \tau_{ij}^k(t)$$

$$\Delta \tau_{ij}^k(t) = \begin{cases} \dfrac{Q}{L_k} & \text{if} \quad t \to t+1, \text{city } i \to \text{city } j \\ 0 & \text{otherwise} \end{cases}$$

N is number of ants.

Copyright ⓒ 2014 SERSC

Q is a constant and denotes the strength of pheromone.

L_k is length of tours that ant k.

ρ is residual pheromone value, $\rho < 1$.

2.2. Clustering algorithm with ACO

P. S. Shelokar [10-11] proposed a novel clustering algorithm, which was different with Deneubourg's algorithm. The algorithm converted the clustering problem into solving optimal value of the objective function. The objective function was Euclidean distance between the point of each sample and the center of the class. In general, the smaller the value of the function, the better the clustering results. The formula expressed as follows.

$$\min \varphi(U,W) = \sum_{j=1}^{K}\sum_{i=1}^{N}\sum_{v=1}^{n} W_{ij}\left\| x_{iv} - U_{jv} \right\|^2 \qquad (3)$$

Where, K is the number of classes that is predefined. N is samples in data set. n is dimension of each sample. x_{iv} is the v-demension value of the i-th sample. W is the $N \times K$ matrix. U is $K \times n$ matrix. The value of W and U are calculated by formula (4) and formula (5).

$$W_{ij} = \begin{cases} 1 & if\ i\ belongs\ to\ cluster\ j \\ 0 & else \end{cases} \qquad (4)$$

$$U_{jv} = \frac{\sum_{i=1}^{N} W_{ij} x_{iv}}{\sum_{i=1}^{N} W_{ij}} \qquad (5)$$

($i = 1,2,3,\ldots N$; $j = 1,2,3,\ldots K$; $v = 1,2,3,\ldots n$)

Pheromone update is similar with the basic ant colony algorithm during the each iteration of the process.

$$\tau_{ij}(t+1) = (1-\rho)\cdot\tau_{ij}(t) + \sum_{r=1}^{R} \Delta\tau_{ij}^{r}(t) \qquad (6)$$

Where, $\Delta\tau_{ij}^{r}(t) = Q / L_k$

In formula (4), $\Delta\tau_{ij}(t)$ is the variation of pheromone update in the t-th iteration. Q is pheromone intensity.

2.3. The strategy of dynamical pheromones

Strategies of the two pheromone dynamic adjustment are combined and applied to clustering algorithm with ant colony optimization in this paper [13-15]. With the increase of the number of iterations, ρ and Q will change dynamically by formula (7) and formula (8).

 Copyright ⓒ 2014 SERSC

$$\rho(t+1)=\begin{cases}0.9\cdot\rho(t) & if \quad \rho(t)\geq\rho_0 \\ \rho_0 & else\end{cases} \qquad (7)$$

$$Q(t)=\begin{cases}Q_0/4 & (t<T/4) \\ 3Q_0/4 & (T/4\leq t<3T/4) \\ Q_0 & (3T/4\leq t\leq T)\end{cases} \qquad (8)$$

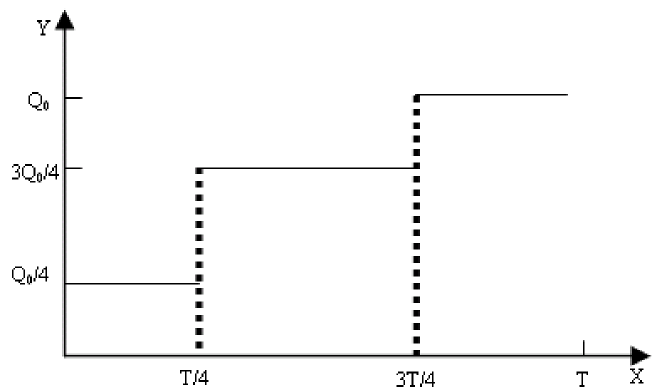

Figure 2. The image of Q(t)

Pheromone update formula (6) is amended as the formula (9) after combining the two pheromone dynamic adjustment strategies.

$$\tau_{ij}(t+1)=\left[1-\rho(t+1)\right]\cdot\tau_{ij}(t)+\sum_{r=1}^{R}\frac{Q(t+1)}{L_k} \qquad (9)$$

2.4. Clustering algorithm with ant colony optimization based dynamical pheromones

Clustering algorithm with ant colony optimization based dynamical pheromones (RRACOC) is depicted as follows.

Step1 Initializing parameters, such as, the number of ants- R , maximal iterations - T , evaporation rate- ρ , initial pheromone matrix- τ_{ij} , etc.

Step2 Using τ_{ij} to construct solutions.

Step3 Calculating w_{ij} and constructing matrix - W .

Step4 Calculating the value of the objective function $\varphi(U,W)$.

Step5 Updating pheromone matrix by formula (9),

$$t=t+1$$

Step6 Terminating the algorithm and outputting results when achieving the maximal iterations - T , otherwise, gotoing step2.

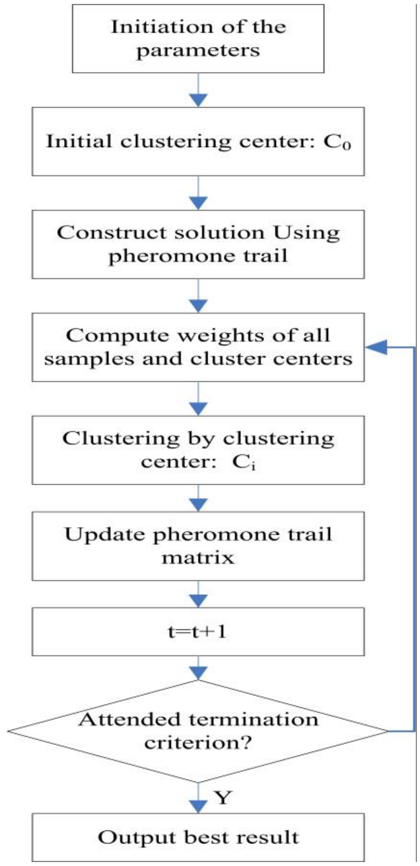

Figure 3. The flowchart of RRACOC

3. Simulation Experiment

3.1. Datasets and parameters setting

In order to verify the performance of pheromone dynamic adjustment combining strategy, this paper chooses two benchmark datasets, Iris and Wine from UCI Machine Learning Repository (Web Site: http://archive.ics.uci.edu/ml/).

Four algorithms, ACOC, EACOC, RRACOC and E_RRACOC, run ten times respectively in two datasets. Parameters setting of these algorithms are listed in Table 1.

Table 1. Parameters setting

	R	T	Rho_max	Rho_min	ρ	E	Q
ACOC	20	500	——	——	0.9	——	50
EACOC	20	500	——	——	0.9	4	50
RRACOC	20	500	0.9	0.001	——	——	50
E_RRACOC	20	500	0.9	0.001	——	4	50

Copyright ⓒ 2014 SERSC

Note:

(1) R is the number of ants. T is the maximal iterations. Rho_max and Rho_min are the maximal and the minimal value of ρ . E is the number of elite ants[12].

(2) **ACOC** is the traditional clustering algorithm with ACO. **EACOC** is the clustering algorithm with ACO based on best solution kept. **RRACOC** is clustering algorithm with ant colony optimization based dynamical pheromones. **E_RRACOC** is clustering algorithm with ant colony optimization based on dynamical pheromones and best solution kept.

3.2. Comparison of results

There are two methods to evaluate the performance of clustering algorithms, method of external quality and method of internal quality. This paper applies two indexes, Precision and Recall, from method of external quality to evaluate the clustering results of four algorithms.

Precision and Recall can be defined as follows,

$$\text{Precision} = N_{ij} / N_i, \quad \text{Recall} = N_{ij} / N_j$$

Where, N_{ij} is the number of i in cluster j . N_i is the number of samples in cluster i . N_j is the number of samples in cluster j .

After running ten times, the value of *Precision and Recall* of four algorithms are listed in Table 2 (iris dataset) and Table 3 (wine dataset) and depicted Figure 4-Figure 7.

Table 2. Precision and Recall of four algorithms in Iris dataset

		T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	Max.
Precision	ACOC	0. 612	0. 589	0. 751	0. 708	0. 695	0. 698	0. 641	0. 570	0. 644	0. 715	0. 751
	EACOC	0. 757	0. 621	0. 647	0. 547	0. 620	0. 625	0. 741	0. 609	0. 709	0. 737	0. 757
	RRACOC	0. 776	0. 612	0. 715	0. 557	0. 633	0. 561	0. 687	0. 730	0. 578	0. 681	0. 776
	E_RRACOC	0. 735	0. 755	0. 672	0. 769	0. 713	0. 788	0. 674	0. 672	0. 718	0. 639	0. 788
Recall	ACOC	0. 316	0. 295	0. 375	0. 353	0. 347	0. 351	0. 322	0. 288	0. 324	0. 359	0. 375
	EACOC	0. 381	0. 311	0. 325	0. 273	0. 312	0. 327	0. 373	0. 311	0. 360	0. 367	0. 381
	RRACOC	0. 390	0. 304	0. 370	0. 285	0. 321	0. 287	0. 343	0. 363	0. 291	0. 346	0. 390
	E_RRACOC	0. 370	0. 381	0. 348	0. 380	0. 357	0. 401	0. 334	0. 337	0. 365	0. 324	0. 401

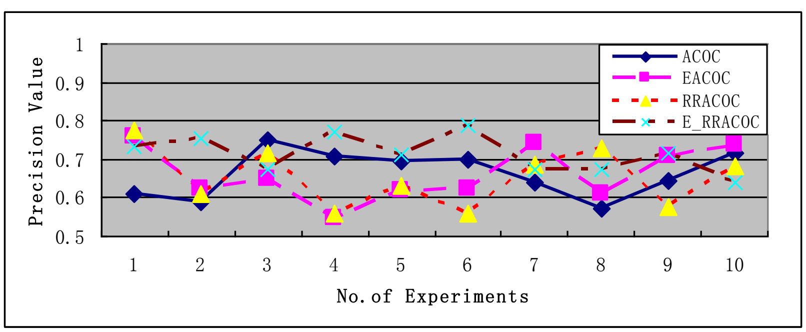

Figure 4. The value of Precision in Iris dataset

Copyright ⓒ 2014 SERSC

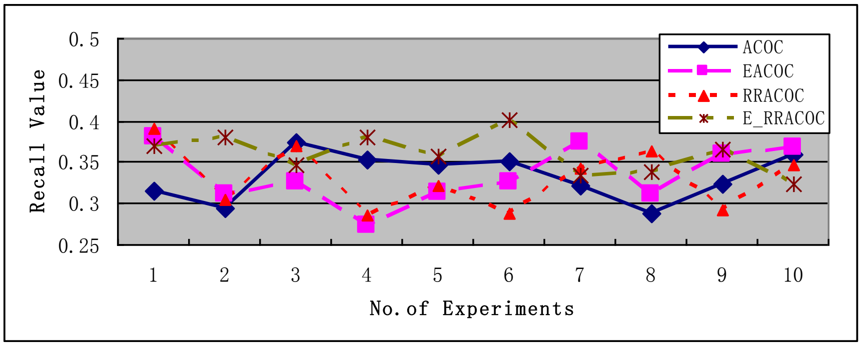

Figure 5. The value of Recall in Iris dataset

Table 3. Precision and Recall of four algorithms in Wine dataset

		T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	Max.
Precision	ACOC	0.554	0.611	0.741	0.724	0.622	0.700	0.683	0.620	0.663	0.657	0.741
	EACOC	0.663	0.612	0.671	0.655	0.598	0.701	0.700	0.663	0.759	0.600	0.759
	RRACOC	0.626	0.713	0.635	0.703	0.739	0.728	0.685	0.741	0.674	0.761	0.761
	E_RRACOC	0.684	0.650	0.559	0.781	0.749	0.535	0.601	0.667	0.699	0.608	0.781
Recall	ACOC	0.282	0.303	0.371	0.370	0.316	0.353	0.347	0.309	0.334	0.331	0.371
	EACOC	0.328	0.311	0.333	0.331	0.303	0.352	0.358	0.333	0.380	0.301	0.380
	RRACOC	0.315	0.360	0.328	0.353	0.371	0.359	0.369	0.371	0.337	0.382	0.382
	E_RRACOC	0.344	0.329	0.285	0.392	0.379	0.267	0.297	0.342	0.352	0.308	0.392

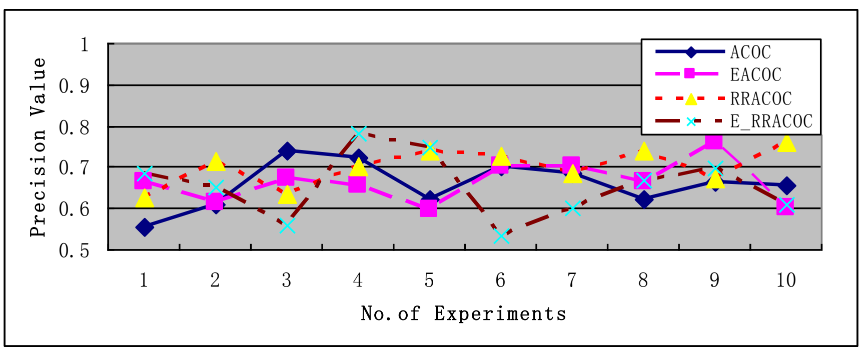

Figure 6. The value of Precision in Wine dataset

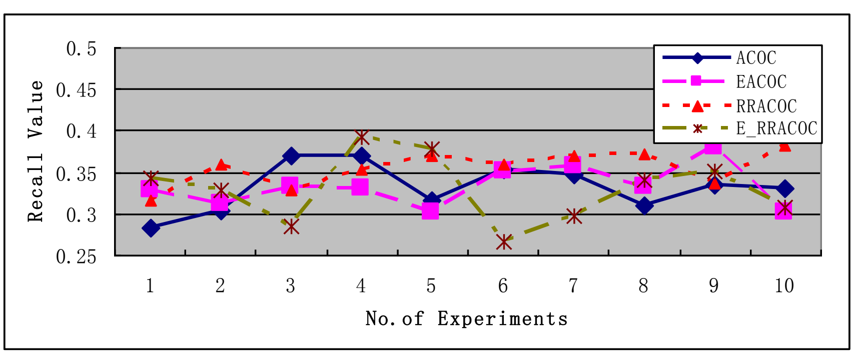

Figure 7. The value of Recall in Wine dataset

 Copyright ⓒ 2014 SERSC

3.3. Discussion

It is seen obviously that the E_RRACOC, which uses strategies of dynamical pheromones and best solutions kept, has larger value of Precision and Recall than the other algorithms. RRACOC, which only uses strategy of dynamical pheromones, is larger than the EACOC. The traditional ACOC has lowest value of Precision and Recall in four algorithms. All in all, the simulation experiment shows that the strategy of dynamical pheromones can improve the performance of ACOC.

4. Conclusion and Future Work

Applying ant colony algorithm to solve some problems, pheromone is used as medium to transfer message, so the variation of pheromone has important influence to solve problems. Two main parameters, ρ and Q, have key effect to the results of algorithms. This paper combines two dynamical adjustment strategies of pheromones with clustering algorithm with ant colony. Numeric experiment illustrates that the dynamical adjustment strategies have better performance. The future research direction is to continue to study the influence of parameters in algorithms, and explore an effective way to improve the efficiency of the algorithm.

Acknowledgements

This work has been supported by grants from Program for Excellent and Creative Young Talents in Universities of Guangdong Province (Yq2013108).

References

[1] J. L. Deneubourg, S. Goss, N. Franks, *et al.*, “The dynamics of collective sorting: Robot- like ants and ant-like robots”, Proceedings of the First international Conference on Simulation of Adaptive Behavior, **(1991)**, pp. 356-366, ISBN: 0-2626-3138-5.

[2] E. Lumer and B. Faieta, “Diversity and Adaptation in Populations of Clustering Ants”, Proceedings of the 3rd International Conference on Simulation of Adaptive Behavior, **(1994)**, pp. 501, ISBN: 0-2625-3122-4.

[3] N. Labroche, N. Monmarche and G. Venturini, “A new clustering algorithm based on the chemical recognition system of ants”, Proceedings of the 13th European Conference on Artificial Intelligence, **(2002)**, pp. 345, ISBN: 3-5404-4036-4.

[4] S. Gao, K. Tang and J. Yang, “A New Clustering Algorithm Based on Hybrid Ant Colony Algorithm”, Microelectronics and Computer, vol. 23, no. 12, **(2006)**, pp. 38-40.

[5] J. Xing, Q. Zhu and P. Guo, “Research on ant colony clustering combination method”, Computer Engineering and Applications, vol. 45, no. 18, (**2009**), pp. 146-148.

[6] X. Xu and Q. Chen, “An Adaptive Ant Clustering Algorithm”, Journal of Software, vol. 17, no.1, (**2006**), pp. 1884-1889.

[7] Y. Meng, X. Liu and Y. Liu, “Application of K-means Algorithm in Macroscopic Planning of Highway Transportation Hub Based on ant Clustering Algorithm”, Computer Engineering and Applications, vol. 44, no. 1, (**2008**), pp. 179-182.

[8] K. Deng and J. Lin, “Research on module partition method for mass customization based on ant colony clustering algorithm”, Computer Engineering and Applications, vol. 44, no. 2, (**2008**), pp. 130-132.

[9] H. Xie, X. Wu and Z. Zhang, “Research of intrusion detection technology based on ant colony clustering”, Application Research of Computers, vol. 27, no. 8, (**2010**), pp. 3050-3052.

[10] P. S. Shelokar, V. K. Jayaraman and B. D. Kulkani, “Ant colony approach for clustering”, Analytica Chimica Acta, vol. 509, no. 2, (**2004**), pp. 187-195.

[11] G. Kekec and N. Yumusak, “Data Mining and Clustering With Ant Colony”, Proceedings of 5th International Symposium on Intelligent Manufacturing Systems, **(2006)**, pp. 1178.

[12] X. Liu, “Text Clustering Algorithm with Ant Colony Based on the Best Solution Kept”, Computer engineering and Science, vol. 32, no. 5, (**2010)**, pp. 79-81.

[13] Y. Wang and J. Xie, “An Adaptive Ant Colony Optimization Algorithm and Simulation”, Journal of System Simulation, vol. 14 no. 1, (**2012)**, pp. 31-33.

Copyright ⓒ 2014 SERSC

[14] G. Qin and J. Yang, "An improved ant colony algorithm based on adaptively adjusting pheromone", International Journal of Information and Control, vol. 31, no. 3, (**2012)**, pp. 198-201.
[15] C. Chen and Z. Luo, "Rapid Algorithm Based on Ant Colony Algorithm", Computer Engineering, vol. 33, no. 6, (**2007)**, pp. 206-207, 210.

Author

Xiaoyong Liu

Currently, he is an Associate Professor who joined Department of Computer Science, Guangdong Polytechnic Normal University in 2007. He obtained the Ph.D. degree in Graduate University of Chinese Academy of Sciences in 2011. His research interests include text mining, data mining, ant colony optimization and genetic algorithm.

 Copyright ⓒ 2014 SERSC