AntMeans: A New Hybrid Algorithm based on Ant Colonies for Complex Data Mining

Messaoud Hameurlaine
Assistant professor
Institute of economics, University
center of Tissemsilt, Algeria

Abdelouahab Moussaoui
Dr,Professor
faculty of sciences, Ferhat Abbas
university, Setif, Algeria

Hadda Cherroun
Dr, Master conference
faculty of sciences, Amar Thledji
university, Laghouat, Algeria

ABSTRACT

The artificial learning is a scientific discipline that covers many aspects of study in mathematics, statistics and algorithms. The techniques of artificial learning have been widely used in extracting knowledge from data, but we are now facing new challenges "complex data." The improvement of data mining techniques has become demanded. After a synthetic study of the different techniques of artificial learning and meta heuristics, this paper propose a hybrid solution "AntMeans" based on tools from data mining: kmeans and AntClust (algorithm based on artificial ants) then this solution is applied firstly in spatial data mining (the extraction of areas in a map of forest fires), and secondly in image data mining (characterization of image regions).

General Terms

Artificial learning, Datamining, Algorithms.

Keywords

Artificial learning, clustering, ant colony optimization, Kmeans, complex data, metaheuristics, spatial datamining, image datamining.

1. INTRODUCTION

Extracting knowledge from complex data often raises more problems than simple data. In general, there are either adaptations in the upstream phase of data preparation, and / or extensions of many mining algorithms by incorporating the specific complex data analyzed [1]. The methods and techniques of machine learning continue to grow and provide better results in various areas in particular data mining.

Many clustering algorithms exist in literature, the most famous is the Kmeans and its variants. These algorithms are very simple to implement and converge rapidly with a locally optimal solution, it is also used when the data is from different sources [2], or in complex data clustering [3]. However their major drawback is that they require an initial partition and the number of classes. Metaheuristics arrive to contribute effectively to solve some problems of clustering [4] ; [5]; [6]; [7]. Recently ants manage to resolve naturally a distributed classification problems such as organization of the brood and the collection of bodies [8]. These behaviors have inspired computer scientists to introduce this new heuristics to solve clustering problem. These algorithms, artificial ants were able to improve the quality of clustering by diversifying the search for solution to obtain a more optimal [9]. The use of Ant colony optimization became an appropriate strategy to solve a lot of problems in deferent domains [10]; [11]; [12]; [13]; [14].

AntClust is a new algorithm for unsupervised partitioning based on the collective behavior of aggregation / segregation of ants to the brood, and based on the work modeling such behavior. AntClust uses the principles of stochastic and distributed exploration in a population of artificial ants to provide an image segmentation of the relevant classes without a partition and start without knowing the number of classes that will be needed. According to the phenomenon of stigmergy, optimal partition emerges from the collective activity of all ants and local interactions between ants and their environment, this algorithm has been developed in [15]. However the major drawback of this algorithm is determining the parameters of the algorithm and the execution time.

To take advantage of Kmeans and AntClust and avoid their drawbacks, we propose a new hybrid algorithm "AntMeans" , the algorithm AntClust is executed with a limited number of iteration to determine the number of initial cluster and a partition of departure for Kmeans.

Firsty this paper present the motivations to propose AntMeans, existing work inspired by the behavior of ants, then it describe the choice of AntClust compared to classifiers based ants exist in literature. Following, it describe in detail the various steps of the algorithm and an experimental study on real databases to show the effectiveness of the method. At the end we validate our algorithm by its application in detection of forest fires and the segmentation of images.

2. MOTIVATIONS

To classify data, many deterministic algorithms exists, the mobile centers is the best known for its fast convergence and simplicity, but requires an initial partition [9]. This is the major drawback of this method: the partition obtained from this initialization may be locally optimal, the only way around this problem is to restart the method with a different initial partition. The same remark can be made regarding the number of classes: the mobile centers require that the number of classes is initialized, which reduces the interest method. The introduction of a stochastic search provides diversification of the search for more optimal results. [4] ; [5])

The use of artificial ants in classification is relevant in the sense that real ants have such problems to solve. In a sense, the model of artificial ants for classification is certainly closer to the problem of classification as the genetic model. In this sense, the model proposed by AntClass, Monmarché has been achieved good results [16], this new algorithm is based on AntClust proposed by Ouadfel for image segmentation [17], this algorithm has shown its quality of classification and ease of choice of parameters compared to all other algorithms in the literature of artificial ants. But since its time complexity is worse compared to Kmeans, we wanted to use just AntClust to determine an initial partition and the number of initial class to start Kmeans. Hopefully our AntMeans will perform better in terms of quality, speed and simplicity of choice of parameters.

3. EXISTING WORK

Deneubourg and his team were the first to model the collective sorting behavior observed in ants. In simulation experiments, the objects to collect are also randomly placed on a grid. Ants are modeled by simple agents that are placed too randomly on the grid representing the environment in which they operate. Each ant agent has only local perception of its environment and has the task of moving objects based on the concentration of similar objects in their immediate environment called "neighborhood." [18].

The algorithm proposed by Deneubourg was repeated and extended by Lumer and Faita in the classification of digital data. The extensions introduced, relate in particular points of data representation, similarity and the principle of neighborliness. [19]

The work of Lumer and Faieta has inspired other authors to solve classification problem with ants. In 1999, Langham has proposed a classification algorithm based ants for minimizing communication between processors in a simulation system where the treatments are spread over multiple processors. [20]

A lot of clustering algorithms use ant model exist in literature [21]. Monmarché introduces AntClass: a classification algorithm using ant populations. AntClass based on the algorithm of Lumer and Faieta with basic modifications. AntClass uses a toroidal grid and each ant has the ability to carry multiple objects at once and bring in a bunch of objects on the same grid cell. [16]; [17].

In AntClust, Ouedfel abandoned the grid because it is not easy to find the correct setting of his several parameters. The environment of ants is an array of N cells each connected to a location representing the colony nest to facilitate the movement of ants from one cell to another. [17]

4. ANTMEANS ALGORITHM

AntMeans is the successive application of AntClust and Kmeans, the first provides diverse research of the solution and the second provides more research. AntClust performs a limited number of iterations which provides an initial partition and a number of favorable Kmeans classes that improves the score.

4.1 Antclust algorithm

In AntClust, the environment of ants is a table (list) of N cells, each connected to a location representing the colony nest, to facilitate the movement of ants from one cell to another. No bodies were initially placed on the table, so that a cell contains only one instance at a time, during the process of classifying a cell contains zero or more instances. At the end of the algorithm the number of nonblank cells represents the number of possible classes.

AntClust uses a population A of X ants $\{F_1, F_2, ..., F_x\}$ that cooperate together and communicate by stigmergy to provide an optimal classification of a set of N instances. The ants move virtual instances from their cells and are grouped according to a criterion of similarity between instances to form homogeneous classes.

It is now necessary to elaborate a criterion of similarity between an instance and other instances of a table cell. For this, we need to evaluate a distance measured between two instances. In our work, we adopted the Euclidean distance with a normalization of distances. Normalization is that all attributes distances are between 0 and 1, so that these distances have the same relevance. To do this we divided the distance between values of an attribute by the maximum distance.

Given a set of N instances described by M attributes $\{a_1,a_2,...a_m\}$, p_i and p_j two instances of this set, the square of the distance d (p_i, p_j) between two instances is given by the sum of squared distances between attributes:

$$d^2(p_i, p_j) = \sum_{l=1}^{l=M} d^2(a_{il}, a_{jl}) \qquad (1)$$

Where $d(a_{il},a_{jl})$ is the distance between the attribute l of the instance i and that of j, which is given by:

$$d(a_{il}, a_{jl}) = \frac{a_{il} - a_{jl}}{d\max_l} \qquad (2)$$

where $dmax_l$ is the maximum distance between the values of the attribute a_l given by: $dmax_l = max_l - min_l$, where :

max_l: maximum value of attribute l of all instances .

min_l: minimum value of the attribute l of all instances .

The main steps of the algorithm are:

Initialization cells: Put each instance in a table cell
Initializing the memory of the ant F_i vacuum
each ant F_i carries a random instance: state [F_i] = carrier;
Main loop
For i = 0 to nb_iterations make
   For each ant F_i do
    Begin
     if state [F_i] = carrier then look in memory of the ants
      If found = true then file(F_i)
      else file(F_i) with probability Pdepot
       update memory of ant
      End if
     else move the ant to another cell
      carry(F_i) with probability Pporter
    Endif
   End
Late nonblanks cells represent classes

The probability to file an instance p_i in a cell C_k is given by:

$$Pdepot(p_i, C_k) = 1 - \cos^2(\frac{\Pi}{2} \cdot F(p_i, C_k)) \qquad (3)$$

The probability to carry an instance p_i from its cell C_k containing n_k instances is given by:

$$Pporter(p_i, C_k) = \begin{cases} 1 & si \quad n_k = 1 \\ q & si \quad n_k = 2 \\ \cos^2(\frac{\Pi}{2} \cdot F(p_i, C_k)) & si \quad n_k > 2 \end{cases} \qquad (4)$$

Where $F(p_i,C_k)$ is the similarity function between an instance p_i and the cell C_k containing n_k instances is as follows:

$$F(p_i, C_k) = \frac{1}{n_k} \sum_{p_i \in C_k} \frac{\alpha^2}{\alpha^2 + d(p_i, p_j)} \qquad (5)$$

$$\alpha = \frac{1}{N(N-1)} \sum d(p_i, p_j) \qquad (6)$$

α is a constant representing the average distance between all instances.

We chose AntClust developed by Ouedfel because it contains fewer parameters compared to AntClass, and we tried to make small changes to AntClust to minimize further the number of parameters and hybridized with Kmeans. The main changes with the algorithm Ouedfel concern:

- The similarity criterion is the Euclidian distance.
- The parameter m (memory size) set to 2 so that the memory of the ant does not store travel exceeded.
- The parameter q (the probability to carry an instance out of two) is maximum.
- The minimum probability pmin=0.5
- The ants file all instances at the end of the algorithm.

4.2 Hybridization with Kmeans

The results found by AntClust represents a partitioning with average quality within the meaning of the number of cluster, the intra-class inertia and interclass inertia, therefore the partitioning built by ants is not necessarily optimal. This point can be avoided by giving more iterations to ants which increases the computing time. The algorithm of mobile centers (or more commonly Kmeans) is then used to correct this type of failure.

5. EXPERIMENTAL RESULT

5.1 Test Data

To evaluate the results obtained by AntMeans, real digital databases are used, this data is downloaded from http://archive.ics.uci.edu/ml/ a recognized site for the test data. These databases are supervised (class is known for each object) in order to assess the quality of partitioning. This class is of course not given to AntMeans.

Table 1. Test data

Databases	Nber of attributs	Nber of classes	Nber of instances	Type of data
Iris Plants	4	3	150	Real
Wine recognition	13	3	178	Integer, Real
Glass Identification	9	7	214	Real
Wisconsin Breast Cancer	9	2	699	Integer
Pima Indians Diabetes	8	2	768	Integer, Real

5.2 Measures of clustering quality

The real databases were already labeled in reference segmentation, and we compare the results to desired results in the reference segmentation. If, for every instance i, $I_{ref}(i)$ is the knownts initial class in reference segmentation and $I_{seg}(i)$ is the class obtained by the clustering algorithm which we want to measure its quality. The commonly used measures for the quality of clustering are: the number of classes, run time, the rand index [17] and classification error [16]; [22]; [23] and the numerical criteria such inertia interclass and intra-class inertia (the Radius).

- Rand index: calculating the frequency of instances classified correctly by the total number of instances. The index of rand defined as follows:

$$R = \frac{a+d}{a+b+c+d} \tag{7}$$

Where :

$$a = \left|\left\{i, j \,/\, I_{ref}(i) = I_{ref}(j) \wedge I_{seg}(i) = I_{seg}(j)\right\}\right|$$

$$b = \left|\left\{i, j \,/\, I_{ref}(i) = I_{ref}(j) \wedge I_{seg}(i) \neq I_{seg}(j)\right\}\right|$$

$$c = \left|\left\{i, j \,/\, I_{ref}(i) \neq I_{ref}(j) \wedge I_{seg}(i) = I_{seg}(j)\right\}\right|$$

$$d = \left|\left\{i, j \,/\, I_{ref}(i) \neq I_{ref}(j) \wedge I_{seg}(i) \neq I_{seg}(j)\right\}\right|$$

- Classification error: This error considers all possible pairs of objects and increases whenever two objects were not classified together by AntMeans while they were in the same class at home and vice versa . This measure has the advantage to take into account the number of classes found.

$$E = \frac{b+c}{a+b+c+d} \tag{8}$$

- The time complexity: it is also considered as a measure of quality of the algorithm, for this we compare the execution times of algorithms Kmeans, AntClust and AntMeans after their implementation with the same programming language and run on the same machine.

5.3 Results

We ran the algorithms Kmeans, AntClust and AntMeans with different parameters, these algorithms were implemented in JAVA language on a machine Pentium dual core 3.40 GHz under a Windows XP operating system. We present the results found:

- The execution time of AntMeans is better than AntClust, this is justified by: AntMeans use Kmeans to improve results but AntClust improves results by the number of iterations. Here AntMeans solved the problem of time complexity of AntClust (see Figure 1):
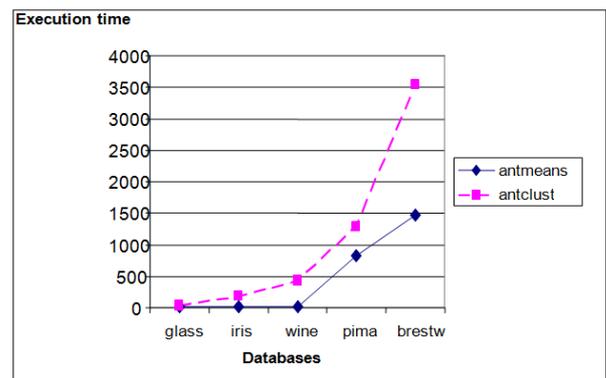

Fig. 1: Variation of execution time

- The number of classes found by AntMeans is close to that of the reference relative to that of KMeans and similar to AntClust except data Glass. Indeed, this is justified by the application of AntClust by AntMeans to find himself a good starting partition needed to Kmeans(see Figure 2):
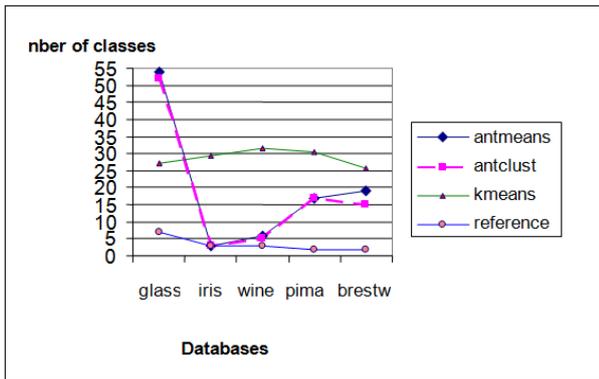
Fig. 2: variation of number classes

- AntMeans is similar to Kmeans and better than AntClust viewpoint classification error, this is due to the use of Kmeans for more research. The Kmeans advantage of rapid convergence is operated by AntMeans for increased research and improved results (see Figure 3):
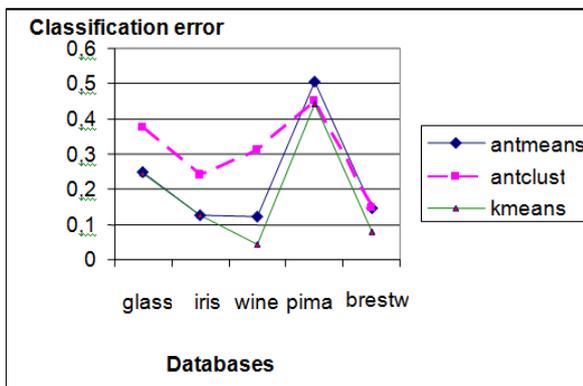

Fig. 3: Variation of classification error

- AntMeans is similar to AntClust but Kmeans is slightly better than AntClust and AntMeans viewpoint intra-class inertia. The classification quality is better if intra-class inertia is minimal (see Figure 4):
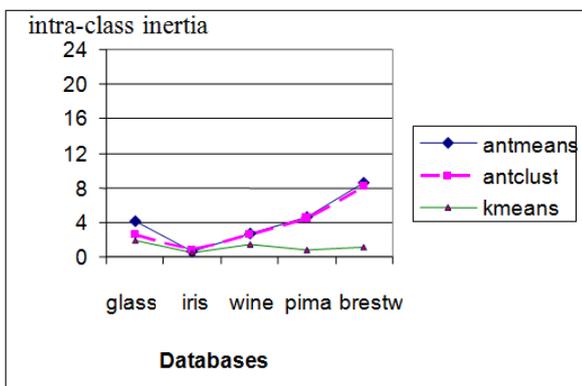

Fig. 4: Variation of intra-class inertia

- AntMeans quality is better than AntClust and KMeans viewpoint interclass inertia. The colaboration between AntClust and Kmeans improves interclass inertia. The quality of classification is better if the interclass inertia is maximized (see Figure 5) :
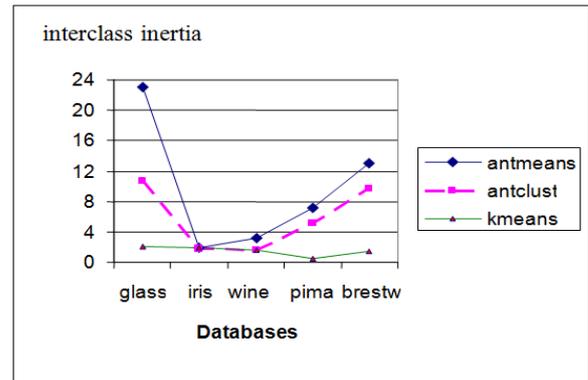

Fig. 5: Variation of interclass inertia

The results tables and graphs that illustrate comparative obtained:

1. The algorithm AntMeans is more efficient than AntClust in:

- Quality of clustering (classification error, Rand index, inertia)
- Ease of setup (number of parameters): the only AntMeans parameters are the number of ant (set at 10) and the number of iterations   easily determined by linear equation: number of iterations = (1 / 3) * number of instances + 10.
- Runtime: AntMeans improves quality by Kmeans algorithm which converges quickly, but AntClust assumed it by a large number of iterations.

2. The algorithm AntMeans is more efficient than KMeans because it looks for his own score for departure without a priori knowledge.

The results obtained show that the quality of clustering found by applying Kmeans is very sensitive to the partition of departure and the number of classes as input. As the number of iterations needed to AntClust to get good results affects the time complexity.

AntMeans is a new approach for unsupervised classification that hybrid AntClust and Kmeans. AntClust is inspired by brood sorting behavior observed in real ants. The main interest of our algorithm is its ability to automatically extract the classes without knowing the number of classes without a priori partition start and improves the quality of classification respecting the constraints of time factor.

After tests on real data, this algorithm was able to show its effectiveness compared to k-means and AntClust. The results of tests and experiments and graph above shows that antmeans is more efficient in terms of quality, speed and simplicity of setting because antmeans has solved a major problem of KMeans (the partition of departure) and the problem of speed and complexity of choice of parameters for ant. This is demonstrated also after comparing our results to the results obtained by the hybrid algorithm ACO-SA (hybrid evolutionary algorithm based on combining Ant Colony Optimization and simulated Annealing) [24].

To show that our algorithm "antmeans" is more powerful algorithm of clustering than others, and he is fit to be adopted for application clustering, particularly complexes data mining. We will apply this hybrid solution to the extraction and characterization of complex data.

6. APPLICATIONS

6.1 Detection of areas of forest fires

Forest fires are an environmental, economic and ecological problem, which can endanger human lives. Their early detection is key to control this phenomenon. To do this, one solution is to use automated tools to extract useful knowledge from spatiotemporal data such as coordinates of the fire, the time of the fire, weather conditions (temperature, wind) and other indices system FWI (Fire Weather Index) [25]. In this work, we will explore our AntMeans hybrid algorithm to predict the category of the burned area of forest fires.

In case of detecting areas of forest fires, any prior knowledge is known (the numbers of zones for example), the zones are represented by complex data like form of maps. The algorithm is applied on tabular data representative of the spatiotemporal data. The reel number of zones given by an expert is 8 zones.

These data recently collected from the northern region of Portugal are available for public research. This data set is created by Paulo Cortez and Anibal Morais (Univ. Minho) in 2007, downloaded from: ftp://ftp.ics.uci.edu/pub/machine-learning-databases/ . These data consist of 517 instances and 12 attributes plus the attribute area (the class).

Tab. 2 : The best results obtained by applying k-means, and AntClust AntMeans on data fireforest

	time	Nber of clas.	Intra-class inert.	Inter-class inert.	error
k-means	1078	8	3,72	4,64	0,37
AntClust	1124094	12	5,28	5,18	0,52
AntMeans	905015	13	4.70	7.34	0,36

After a series of execution Kmeans on fireforest dataset, the best result Kmeans discovers the number of 8 zones and a good quality of classification, but it requires a good prior knowledge of many areas of origin. But if we give a random initial partition (with a random number of classes) the result is incomparable.

After a series of execution AntClust on fire forest dataset, the best result of AntClust discovered 12 areas without prior knowledge, but with bad time complexity and worse quality of clustering.

While, after an execution of AntMeans with simple, known values and unique parameters, AntMeans could detect 13 fire zones with time complexity better than AntClust and clustering with quality better than Kmeans and without knowledge of priori number of areas. This shows the effectiveness of our algorithm in terms of quality, speed and simplicity.

Therefore, the results obtained by AntMeans proves that its exploitation in extracting knowledge from complex data (spatiotemporal) is promising and provides a workable solution to the problems of clustering, especially those that require knowledge of the number of classes.

6.2 Image segmentation

Image segmentation is an image processing operation which aims to bring together the pixels according to predefined criteria. The pixels are grouped together into regions, which constitute a partition of the image. Develop segmentation algorithms high level (each region is a semantic object) is still a topic of research the most common image processing. [26]; [27] ; [28].

In this application, the objective is the extraction and characterization of objects from images. The objects (brickface, sky, foliage, cement, window, path, grass) for which a cluster is defined. This knowledge of objects (type of regions) image can be used for another task or data mining to make decisions. The algorithm will applied for segmentation of images taken at random from the outside to identify the types of areas of these images.

The data collected in November 1990 by Vision Group, University of Massachusetts are available for public research. This data set was randomly drawn from a database of 7 outdoor images. The images were segmented by hand to create a classification for each region. Each instance describes the characteristics of a region of 3x3 pixels of the image must belong to a class of seven (brickface, sky, foliage, cement, window, path, grass).

These data in tabular form were downloaded from: ftp://ftp.ics.uci.edu/pub/machine-learning-databases/ and used for image segmentation. These data consist of 210 instances and 19 attributes plus the class attribute of the image area. Attributes are descriptors of a region consist of 3X3 pixels.

Tab. 3: The best results obtained by applying k-means, and AntClust AntMeans on image data

	time	Nber of clas.	Intra class inert.	Inter class inert.	error
k-means	719	7	3.19	4.36	0,08
AntClust	1124094	7	2.85	2.77	0,51
AntMeans	130891	7	3.10	4.70	0.24

After a series of execution of Kmeans on the image data, the best result of Kmeans discovers the number of 7 items and good quality of classification, but it requires a good prior knowledge of initial partition.

After a series of execution of AntClust on the image data, the best result discovered 7 objects without a priori knowledge, but with bad time complexity and worse quality of clustering.

While, after an execution of AntMeans with simple, known values and unique parameters, AntMeans could detect objects with a time complexity better than AntClust and clustering with quality better than Kmeans and without a priori knowledge of objects number.

Therefore, the results obtained by AntMeans prove that its operation in extracting knowledge from complex data (image segmentation) is promising and provides a workable solution to the problems of clustering, especially those that require knowledge of class number.

In the second example of application of AntMeans for extracting knowledge from complex data, our algorithm has been applied sucked for image segmentation to characterize the types of image regions chosen at random on the basis Data collected from specific regions descriptors 3X3 pixels. The algorithm has achieved the extraction and characterization of objects from images, the seven classes or clusters represent the results of clustering objects (brickface, sky, foliage, cement, window, path, grass) for which a cluster is defined. This knowledge of objects (type of regions) image can be used for another task or data mining to make decisions.

We also deduced by comparison of results obtained by the conventional method Kmeans, the result obtained by AntMeans is slightly better than KMeans even with a priori

knowledge of ideal number of classes. In addition the AntMeans results are highly better than Antclust.

7. Conclusion

After a synthetic study on different learning techniques, this paper proposes a hybrid algorithm based on tools from data mining and apply this algorithm to the extraction and characterization of complex data. This algorithm is able to improve certain criteria based on the collective behavior and self-organized ant. Ants have self-organized behavior, through which simple interactions at the local level allow the emergence of a complex overall behavior.

In AntMeans algorithm, the ability of globalized searching of the antclust algorithm and the fast convergence of the K-means algorithm are combined. The antclust algorithm is used at the initial stage to help discovering the vicinity of the optimal solution by a global search. The result from antclust is used as the initial seed of the K-means algorithm, which is applied for refining and generating the final result.

The proposed hybrid algorithm has been applied to real data, it generates the highest clustering compact result in the experiments. It has shown its effectiveness as regards Kmeans especially a priori knowledge of class number and the partition of departure, because our algorithm requires no knowledge priori. In addition to this, our algorithm improves the time complexity of AntClust and simplifies its implementation because it requires no special configuration.

The hybridization technique can improve the quality of the stain of unsupervised classification, it was clear in the first application where the algorithm was able to detect various types of area fire drills without knowing the number of areas and in the second application where it was determined the objects in an image without knowing the number of objects.

So, under the prospects of this modest work, Future work on the use of ants in classification should address the case of symbolic data. This algorithm based on ant colonies can be parallelized to speed up processing for large data. Secondly, the feasibility of the proposed algorithms, their simplicity of implementation and the quality of solutions compared to conventional methods, it is important to generalize this work by modeling other collective behavior of social insects to propose other segmentation methods using the concept of self-organization and emergence.

8. REFERENCES

[1] Boussaïd O., Gançarshi P., Florent M.and Brigitte T. "Complex data mining" Review of New Technologies of Information, Series E, No. 4, 2005 Cépadues-Editions, 286 p.

[2] Meskine F. And Naitbahloul S "Privacy preserving K-means clustering: A survey research" the International Arab Journal of Information Technology, (2012) Vol.9, No.2, p 194-200.

[3] Pitchandi P. And Raju N. "MLK-Means – A hybrid Machine Learning Based k-Means Clustering Algorithm for Document Clustering" International Journal of Computer Science Issues, Vol. 9, Issue 5, No 2, September 2012.

[4] Dréo J., Alain P., Patrick S. and Taillard É. " Metaheuristics for optimization difficult " book coordinated by Patrick Siarry, Editions Eyrolles(2003).

[5] Jourdan Laetitia "Metaheuristics for knowledge discovery: Application to genomic" doctoral thesis (2003).

[6] Nguyen, Q. and Rayward-Smith, V. J. « CLAM: Clustering Large Applications using Metaheuristic » Journal of Mathematical Modelling and Algorithms, (2010) V.10, N1, pp. 57-78(22).

[7] Swagatam Das, Ajith Abraham and Amit Konar « Metaheuristic Clustering » springer edition(2009).

[8] Dorigo M. and Stutzle T. « Ant colony optimization» A Bradford Book, The MIT Press (2004), Cambridge, Massachusetts London, England.

[9] Labroche Nicolas "Modeling the chemical recognition system of ants to the problem of unsupervised classification: Application to Internet audience measurement "PhD thesis(2003).

[10] Leandro P. and Guilherme E. « Ant Colony optimisation for Backward Production Scheduling » Advances in Artificial Intelligence Volume 2012, Article ID 312132.

[11] Ostfeld A. « Ant Colony Optimization - Methods and Applications » Edited by InTech(2011).

[12] Geetha R. and Umarani Srikanth G. « Ant Colony Optimization in Diverse Engineering Applications: an Overview » International Journal of Computer Applications (0975 – 8887) Volume 49– No.17, July 2012, p 19-25.

[13] Binoy B. Nair; V.P. Mohandas and N.R. Sakthivel « Predicting stock market trends using hybrid ant-colony-based data mining algorithms: an empirical validation on the Bombay Stock Exchange » Int. J. of Business Intelligence and Data Mining (2011) Vol.6, No.4, pp.362 – 381

[14] Kannan S., Kalaikumaran T., Karthik S. And Arunachalam V.P "Ant colony optimization for routing in mobile Ad-Hoc Networks" International Journal of Soft Computing (2010) 5(6): 223-228.

[15] Ouadfel S. and M. Batouche " Ants for image segmentation", Art Workshops in Computer Graphics JIG05 (2005).

[16] Monmarché Nicolas " Artificial ant algorithms: applications to classification and optimization "PhD thesis(2000).

[17] Ouadfel S. "Contributions to the segmentation of images based on solving collective colony of artificial ants "PhD thesis (2006) .

[18] Deneubourg, J.-L., Goss, S., Franks, N., Sendova-Franks, A., Detrain, C., and Chretien, L. " The dynamics of collective sorting: robot-like ant and ant-like robots". In (Meyer and Wilson, 1990), pages 356–365.

[19] Lumer, E. and Faieta, B. "Diversity and Adaptation in Populations of Clustering Ants" In (Cliff et al., 1994), pages 501–508.

[20] Langham, A. and Grant, P. "Using Competing Ant Colonies to Solve k-way Partitioning Problems with Foraging and raiding strategies". In (Floreano et al., 1999), pages 621–625. 1999.

[21] Yu-Chiun Chiou and Shih-Ta Chou. « Ant Custering Algorithms » International Journal of Applied Evolutionary Computation 2010,1(1), 1-15

[22] Monmarché Nicolas & Alain Lenoir " Real ants to artificial ants: Two examples of biomimetic approaches to solving problems in computer science "article CASCIMODOT Round 6 (2007).

[23] Candillier L. () "Contextualization, visualization and evaluation unsupervised learning" doctoral thesis 2006.

[24] Niknam T., Olamaei J. and Amiri B. « A Hybrid Evolutionary Algorithm Based on ACO and SA for Cluster analysis » Journal of Applied Sciences(2008) 8(15) : 2695-2702.

[25] 25Cortez P. and A. Morais "A Data Mining Approach to Predict Forest Fires using Meteorological Data" article Portuguese Conference on Artificial Intelligence 2007 .

[26] Bignon J.C., Alin G., Nakapan W. And Wagner M., "Extraction and indexing image applied to the field of architectural and engineering design" Revue UMR MAP CNRS N°694 vol. 1, 1998.

[27] Stanchev Peter (2003) "Using image mining for image retreival" Kettering University Flint, Michigan, USA article IASTEAD conf mexico, mai 2003 p 214-218.

[28] Marie-Anne Germaine and Anne Puissant " Extraction of landscape indices and quantitative landscapes of "valleys ordinary" from image data: The example of Seulles (Calvados, France) (2008)".