# 人工免疫系統應用於資料分類上的探討
# Artificial Immune System Application on Data Clustering

邱正毅(Jeng-Yih Chiou)
崑山科技大學資訊工程系
jychiou@mail.ksu.edu.tw

**論文摘要**

本論文提出一個運用人工免疫系統的方法對資料進行自主分群的技術。透過人工免疫系統中的免疫網路定理，經過編碼後的資料將可以自動被分到某一類別內。除此之外，有限資源問題也被考慮並且有提出解決方案。系統中用於分類的參數也會於處理資料的過程中被修正進而影響原先的分類結果以便於最終產生最適當的類群。Fisher's Iris 資料集被用於驗證此分群技術，實驗結果也同時與數種分群技術做比較。結果顯示，所提出的方法，其分類結果並不輸於其他有名的分群技術，甚至比其中的某方法獲得更好的結果。

**關鍵詞：**共振適應理論、人工免疫系統、資料分群、免疫網路理論、分割

## Abstract

A clustering method based on the artificial immune system is proposed in this paper. This method uses the immune network theory to data clustering problem. Architecture of this method is shown clearly to help hardware implement. In addition, limited resource constrains question have been considered. All parameters needed by this method have been studied in order to find the best appropriate set of parameters. Moreover, the time complex and space complex of the proposed method are also discussed. Some well-known sets, such as the Fisher's Iris data set, have been used for examination of clustering techniques. The ART1 clustering method, the general fuzzy min-max method, and the fuzzy min-max method are used to compare with this proposed method. The experimental results show this method has better results than those compared methods. In addition, even though considering the limited resources, the error rate still keeps in the same level.

**Key words:** Adaptive resonance theory, artificial Immune System, clustering, data analysis, Fisher Iris data, immune network theory, segmentation.

## Introduction

The human immune system recognizes and attacks foreign substance, called pathogens, to provide protection for human physical health. It has the ability to recognize unknown pathogens results from primary responses, secondary responses, and cross-reactive responses. Researchers have proposed the clonal selection theory and immune network theory to describe how these responses work [1-3]. In addition, a complement mechanism has been proposed to model the recognition process.

Since, a complement mechanism is a kind of similarity measurement mechanism. Because clustering is also based on the similarity to recognize and cluster input data, artificial immune systems have the potential to be as clustering or self-organizing classifiers.

There are many kinds of clustering methods have been proposed and applied to many fields, such as pattern classifications [4,5], the problem of incremental exploration of unknown environment [6], finding decision boundary [7], segmentation [8-11], and so on. Basically, a classifier receives a feature and outputs what category this feature belongs to, and clustering groups data according to the similarity measured on features. Therefore, how to describe a category/cluster and a feature in an artificial immune system is the first problem. Obviously, the second problem is how to determine what category/cluster a feature belongs to in an artificial immune system. In addition, a self-organizing classifier or clustering must have an ability to learn recognizing new categories/cluster so how artificial immune systems learn recognizing a new category/cluster is the third problem.

Timmis et al. have proposed an artificial immune system for cluster identifying [12]. All training data are pathogens in their system. Each matching B-cell will clone and mutate. They use a concept of network affinity threshold (NAT) to decide whether a link between two B cells exists or not. The value of NAT is evaluated by all links that connect B-cells to from a network. These connected B-cells are regarded as clusters, each of which represents a category. Timmis et al. have also proposed a improved version, called the resource limited artificial immune network (RLAIN), to control the reportoire size [13], They use a concept of artificial recognition balls (ARBs), which are based on an recognition ball concept, to represent a group of identical B-cells. Thus, they can use a finite number of ARBs to represent an infinite number of epitopes.

Since the immune system has the clustering ability, we propose alternative artificial immune system that categorizes clusters based on immune network. In this system, clusters are made up grouped B-cells, which are connected together to form an artificial immune network.

The immune network theory considers all antibodies and antigens as a network relationship. N. K. Jerne formally proposed it in 1974 [1]. Jerne thought that each antibody has two portions on the binding site. One is the paratope, which is responsible for recognizing idiotopes of other antibodies, or epitopes of antigens. The other is the idiotope, which can be recognized by other antibodies. In addition, an antigen has at least one epitope, each of which represents a characteristic of its own, and which appropriate antibodies could recognize. In case of no antigens, the density of each antibody population, called concentration, will keep at a balance situation because of stimulation and suppression interactions between antibodies. Since, an antibody suffers suppression and results in reducing concentration while being recognized by other antibodies; and, oppositely, it will be stimulated and results in increasing concentration while recognizing other antibodies or antigens. In addition, an antibody is going to die since the life period is limit. This phenomenon has been represented by many equations proposed by Jerne, Farmer, Varela, Coutinho, and so on [2,3,14-16]. Perelson has suggested that all of these network models could be expressed as the following form [17]:

$$\begin{aligned}&\frac{d}{dt}\left(Population_i(t)\right)\\&= Stimulation_i(t) + New_i(t)\\&- Suppression_i(t) - Died_i(t),\end{aligned} \quad (1)$$

where 'Populationi' represents population of the antibody, i; 'Stimulationi' represents increasing number of antibodies, i; 'Newi' represents number of influx newborn antibodies, i; 'Suppressioni' represents the reducing number of antibodies, i, resulting from suppression; 'Diedi' represents the reducing number of antibodies, i, resulting from death; t represents time.

After the introduction, the proposed algorithm for clustering is described in detail, including the concepts and discussions. In addition, an implement algorithm is listed by pseudo-codes. Then, experiments demonstrating the performance of this algorithm are shown. Finally, a conclusion is shown at last section.

## Artificial Immune Clustering System

The architecture of this system includes four major parts: an artificial marrowbone, a sensitive level evaluator, a pruning mechanism and an artificial immune network poor. The whole architecture is shown in Fig. 1.

We use idiotope to record and encode the feature of each data. In addition, we use paratope to represent the feature of a category/cluster. An idiotope is recognized by paratope. Thus, this process represents a measurement of the similarity relationship between a class/cluster and a data because idiotope represents data feature and paratope represents a category/cluster. Recognizing is achieved by the Complement Mechanisms, in which regions of complementarity are used, and the quantity of the complementarity is evaluated by affinity functions [2]. There are many affinity function has been proposed [2]. Each of them is regarded as a kind of similarity measurement in a space, such as shape-space. For examples, many researchers use a series of numbers to represent a region. Thus, the affinity function for measuring complementarity of two regions can be a function of distance, which is inverse proportion to distance between these two numeric strings. When the distance is small, the affinity is large. Accordingly, the feature of each data is regarded as a point in this n-dimension data feature space, and Euclidean distance can be used to evaluate distance between two features. In addition, because a category/cluster is a group of similar features, a category/cluster can be represented by using the average, called mean, of all feature vectors belonging to this category/cluster and a range value or vector, which can be the maximum distance of all features from the mean. Accordingly, in the proposed system, an idiotope of an antibody is a sample data feature, and paratope of the same antibody is a category/cluster feature that the data belong to. Each feature digit of a paratope is constructed as follows:

$$p_i = \frac{1}{n}\sum_{k=1}^{n} e_k, \quad (2)$$

where $p_i$ is an paratope digit of a category/cluster $i$; $n$ is total number of antibodies in the category/cluster $i$; $e_k$ is an epitope of an antibody that belongs to the category/cluster.

An input data is regarded as a newborn B-cell in the system. The idiotope of this newborn B-cell is the data feature. Because the category/cluster of this data has not been unknown yet, this newborn B-cell is regarded as an independent category/cluster until it is recognized by an existed category/cluster according to sensitive level. In the other words, the paratope is also given the same data feature. This sensitive level is evaluated as follows.

$$s_j(Ab_i) = \frac{k_1 \cdot \sum_{k=1}^{m_j} f(\mathbf{p}_i, \mathbf{e}_{j,k}) - k_2 \cdot \sum_{k=1}^{m_j} f(\mathbf{p}_{j,k}, \mathbf{e}_i)}{m_j}, \quad (3)$$

where $s_j$ is the sensitive level of an antibody $Ab_i$ to network$_j$, which represents a category/cluster $j$; $k_1$ and $k_2$ are constants and between 0 and 1; $m_j$ is the total number of Antibodies in the network$_j$, $\mathbf{p}_i$ is the paratope string of $Ab_i$; $\mathbf{e}_i$ is the idiotope string of $Ab_i$; $\mathbf{p}_{j,k}$ is the paratope string of $Ab_k$ that belongs to network$_j$; $\mathbf{e}_{j,k}$ is the idiotope string of Ab$_k$ that belongs to network$_j$; $f$ is an affinity function and defined as follows.

$$f(\mathbf{p}, \mathbf{e}) = e^{-\frac{d(\mathbf{p},\mathbf{e})}{\|\mathbf{p}\|}}, \quad (4)$$

where $d$ is a distance function in the feature space; $\|\mathbf{p}\|$ is the norm of the vector $\mathbf{p}$.

The numerical part of (3) has two terms. The first term represents stimulation, and the second represents suppression. According to (3), the relationship of an antibody with a set of antibodies can be measured. Because a distance function yields a positive real value, Eq. (4) must be a value between 1 and 0. In addition, when distance approaches 0, the value of *f* approaches 1; oppositely, when distance increases, the value of *f* decreases toward 0. The $\|\mathbf{p}\|$ plays a role that makes a relative value on the power of the equation in order to represent a rate of distance with respect to the magnitude of the vector **p**.

A network in the AIS is expanded when a new antibody (Ab) is added in. Thus, an Ab in a network may have more sensitive level to other networks in the same AIS. When the situation happens, the Ab must be moved to that network and both paratopes of these networks have to be modified, respectively. The process resembles pruning weight process in artificial neural networks. In addition, when a network only has one member after learning finishes, this class will be cancel and its member must be moved into another network where the member has maximum sensitive level.

Two constants, k1 and k2, in (3) determine the sensitive level. Basically, according to (3), the constant k1 affects the stimulation. In other words, it is a weight of stimulation. In addition, the constant k2 affects the suppression and is also a weight of the suppression.

*The limited resource problem*

A problem of the proposed algorithm is the storage space problem because the number of B-cells in a network and the number of clusters are unlimited. This means that the system needs unlimited storage space. The problem leads to hardware implement problem and unexpected calculating time. In order to consider these constrains, we add a mechanism in the B-cell/Net adder/remover block to examinant whether the number of clusters and B-cells has got to a given upper boundary. If the number of clusters and B-cells has got to the upper boundary, some clusters or B-cells must be destroyed and released to the raw data pool. This mechanism is inspired by the weak cells death procedure. Thus, the whole architecture has been modified and shown in Fig. 2.

This new mechanism will remain a desired number of clusters and the total number of members in the system. That means if the desired number of clusters is *m*, the first *m*-th cluster, which are sorted by the number of members, is kept, and the others else are destroyed and regarded as new raw data, whose features are the epitopes of their B-cells, respectively. For example, if a network, including 2 B-cells, is going to be destroyed. These epitopes of the B-cells are <0.2, 1.1, 2,1> and <0.5, 2.3, 1.2>, respectively; the raw data set will have 2 new data: <0.2, 1.1, 2.1> and <0.5, 2.3, 1.2> after this network is destroyed. According to this method, the number of networks/clusters in the network pool is kept under a limited desired number.

*The time complex problem*

This algorithm includes lots of measurement actions, each of which needs to measure the sensitive level between all existence clusters/networks. This process needs very large of computation time. For example, if there are *m* clusters and $n_j$ members in each cluster *j*, the computation magnitudes of stimulation and suppression are $\sum_j((m\text{-}1)\times n_j)$ for each member in a cluster. So, the time complex is $O(n^3)$ for all clusters and members. However, when a cell is added in a cluster or to make a new cluster, only one cluster is going to change. This means that only members in one cluster need to been examined. So, the time complex in this case become $O(n^2)$.

So far, the time complex has been reduced 1 order. If the time complex needs to be reduced more, the calculation operation must be reviewed. Moreover, the similar measurement includes a root square and power operators. In order to reduce the time complex more and still have the same concept on distance measurement, the distance equation form has to be replaced by following equation:

$$d(\mathbf{p},\mathbf{e}) = \sum_{k=1}^{n} |p_k - e_k|, \tag{5}$$

In addition, the norm equation form is replaced by following equation:

$$\|\mathbf{p}\| = \sum_{k=1}^{n} |p_k|. \tag{6}$$

Thus, these calculations only include summation and subtraction so that the calculation time will reduce more.

*Stop learning problem*

This new method still has a potential problem that the raw data in the raw data pool might not get to the empty situation since the raw data will have the chance to be recruited. In order to resolve this problem, we study in what situation the data pool will not go to empty by changing the threshold and constants. We also use the Firesh's Iris data set and keep the cluster number under a desired number, such as 20. We use a successful flag to indicate whether the raw data pool is empty or not. In addition we also use original distance measurement function and new distance measurement function (5) in order to know the effects on change this function. The experimental arguments, k1, k2, and threshold value, are changed in bounded ranges with the same step, 0.1. The boundary ranges of both arguments, k1 and k2, are between 0.1 to 1, and the threshold is changed from 0.1 to 0.9. Some results are shown in Fig. 3-6. Fig. 3 and 4 show the results using original distance measurement, and Fig. 5 and 6 show the results using the new distance function. Because there are 4 arguments, we show these experimental results by 3-D figures, each of which have the whole

k1 and k2 values as x-axis and y-axis but different measured events on z-axis under a certain threshold. These measured events on z-axis are the number of clusters, the error of assigned member, success flag, and the repeat time, which represents how long the raw data pool becomes empty.

Observing these results, we can find the success appear when k1>threshold only and despite what value the k2 is. We rearrange these data and make figures, such as Fig. 7 and Fig. 8, to describe error rate vs. k2 in order to know when success flag is one, how the k1 and threshold affect the error rate. According to these figures, the smallest error rate at different threshold values appears whilst k1 is closer to and greater than the threshold value. In these experimental results, the best result, in which the error rate is 2.66%, appears when threshold is 0.9, k1 is 1.0, and k2 is 0.7. When k1 and threshold value at appropriate values, where the error rate is less than 9%, the best results always appear at k2=0.7 or 0.8. According to these phenomena, a set of best parameters must depend on what the threshold is and whether it is closer to the threshold. In addition, the best threshold value is the value that is closer to 1. In addition, the most appropriate k2 is greater than 0.6. Thus, we can use these phenomena as information to chose the possible most appropriate parameters and a threshold value, such as threshold value is 0.9, k1 is 1.0, and k2 is 0.7.

Because this artificial system has been modified in order to possess the capability to keep the number of networks under a desired magnitude, using the artificial system under limited resources environment is possible. In addition, this clustering method is still a linear clustering because the similar measurement function is based on the distance measurement. This leads to clustering error increases when the number of clusters is decreased. This means that when a set of test data, which includes a known number of categories, is given to this clustering mechanism, more number of clusters including different type members increases when the maximum number of clusters decreases. However, according to the results under different parameters and thresholds, a minimum number clusters still exists to result in the best clustering results. The experimental results also show a phenomenon that the error rate has a little increasing, about 0.6% under changing the distance measurement.

## Experiments

All experiments in the paper use the Fisher's Iris data. The data downloaded from "http:// www.ii.fmph.uniba.sk/~benus/courses/ Iris.data.txt". This data set has 150 random samples of flowers from the Iris species: *setosa*, *versicolor*, and *virginica* , collected by Anderson [18]. From each species there are 50 observations for sepal length, sepal width, petal length, and petal width in mm as shown in Fig. 9. Fisher used this dataset on data analysis for his linear discriminating function technique [19]. The data set has become a well-know benchmark.

In order to understand the effects of these parameters, the constant k1 and k2 are respectively changed from 0 to 1, and the threshold is changed from 0.1 to 0.9 with step 0.1. Thus, there are (1/0.1+1) * (1/0.1+1) * ((0.9-0.1) / 0.1 + 1) = 1089 experimental results. The data set used on these experimental results is Fisher's Iris data set.

According to the experimental results, obviously, the error rates are mostly affected by the threshold value and constant k1 because when threshold value increase, the error rate always decrease under appropriate constant values of k2. As the results of experiments, obviously, the constant k1 and the threshold affect the number of clusters significantly. In addition, the constant, k2, seems to have a potential of turning the cluster number when the error rate keeps at a certain small value as the table 1 shown.

The table 1 shows that the best results appear when the threshold values are 0.6, 0.7, 0.8, and 0.9; the k1 values are 0.7, 0.8, 0.9 and 1.0; and the k2 values are 0.3, 0.5, 0.6, 0.7, 0.8, 0.9, and 1.0. The k2 values seems to be affected by k1 and threshold because the average of k2 values decreases while the threshold value and k1 values increase. In addition, the smallest error rate appears whilst k1 is closer to but greater than threshold value in these results. Accordingly, if a smaller error rate is desired, the k1 must be closer to and greater than threshold value. If the bigger number of clusters is desired, the threshold value needs to closer to 1.

In order to compare the performance of proposed method, a clustering method based on the adaptive resonance theory (ART) is used to clustering the Fisher's Iris data as well. The threshold values of the ART are changed from 0.65 to 0.95 with step 0.05 and from 0.95 to 0.99 with step 0.01.

The experimental results are shown in Fig. 10 and 11. In these figures, the x-axis represents the threshold of ART and the y-axis represents the error rate in percent in Fig. 10 and the number of clusters in Fig. 11. Obviously, the cluster number increases whilst the threshold value increases. In addition, changing the threshold form 0.95 to 0.99 with step 0.01 the smallest error rate still is almost 17.5%.

Comparing to these experimental results, the smallest error rate of our method with space complex problem solution and time complex problem solution is 3.3% when the threshold=0.9, k1=1.0 and k2=0.7. In addition, its number of clusters is 24 at the same parameters. Oppositely, the smallest error rate of the ART method is almost 18% when the threshold is 0.97, and its number of cluster is only 8. Although the ART can has lesser number of clusters but has worse error rate than ours. In other words, the proposed method has better clustering results than ART method.

In addition, we compare to a paper that is also proposed a better clustering method, the general fuzzy max-min method (GFMM), and compare to some other method such as fuzzy max-min method (FMM) [7] because they use the same Fisher's Iris data set. Their results show their GFMM method has the best recognition rate between 92%~100%, and Simperson's FMM method [20,21] has the recognition rate between 92% and 97.33%. In other words, the average error rate of both their methods are about 4% and 5.5%. However, our method has the error rates between 2.66% and 3.3%, In other words, the average error rate is about 3% and, obviously, is better than GFMM and FMM.

## Conclusion

A clustering method based on immune network theory has been proposed and examined by experimental ways. This method uses an artificial marrowbone to create newborn B-cells, each of which is an input data. Then, according to the immune network theory to cluster these B-cells. Since the natural features of immune system, these B-cells can form many sets of B-cells. Because the similar measurement method is based on the distance measurement, each set of B-cells represents a cluster and this method is a kind of linear discrimination. In other words, each cluster seems to be surround by a hyper-ball. Therefore, more number of clusters, higher of correct rate the system has. This has been confirmed by the parameters analysis procedure. In these analysis procedures, the threshold values and constants k1 and k2 are changed from 0.1 to 0.9 or 1.0 and the clustering error and the number of clusters are recorded in order to know the effects under different parameters. As the results, the proposed parameters to have the possible smallest error rate are threshold value is 0.9, k1=1.0, and k2=0.7. In addition, comparing to an ART based clustering and to other paper results has been done to know how better the proposed method has. According to these experimental results, our method has better clustering results than those methods. In addition, our method may have $O(n^2)$ at best situation in time complex, and the limited resource constrain has been consider into the algorithm as well. Thus, using the clustering method in real-time constrain and implement by hardware could be possible. Accordingly, in the future, real segmentation applications and hardware implements are going to be done.

## Reference


[1] N. K. Jerne, "Towards a Network Theory of the Immune System," *Annual Immunol 125C*, pp. 373-389, 1974.

[2] L. N. de Castro and J. Timmis, *Artificial Immune System: A New Computational Intelligence Approach*, Springer, 2002.

[3] N. K. Jerne, *Clonal Selection in a Lymphocyte Network," Cellular Selection and Regulation in the Immune Response*, G. M. Edelman (ed.), Raven Press, N.Y., pp. 39, 1974.

[4] G. Yen and Phayung Meesad, "An Effective Neuro-Fuzzy Paradigm for Machinery Condition Health Monitoring," *Proceedings of the 1999 IEEE International Conference on Control Applications*, pp. 1567-1572, Kohala Coast-Island of Hawaii, Hawaii, USA, August 22-27, 1999.

[5] S. Schupp, A. Elmoataz, J. Fadili, P. Herlin, and D. Bloyet, "Image segmentation via multiple active contour models and fuzzy clustering with biomedical applications," *International Conference on Pattern Recongnition*, vol. 1, pp. 622-625, 2000.

[6] R. Dellacasa, P. Morasso, S. Repetto, G. Vercelli, and R. Zaccaria, "Self-Organizing Navigation: from Neural Maps to Navigation Situations," *Proceedings of the 1993 IEEE International Conference on Tools with AI*, pp. 458-459, Boston, Massachusetts, Nov. 1993.

[7] B. Gabrys and A. Bargiela, "General Fuzzy Min-Max Neural Network for Clustering and Classification," *IEEE Transactions on Neural Networks*, vol. 11. no. 3, pp. 769-783, May 2000.

[8] Y. Li, H. Chen, M. Zhao, and P. Ou, "Self-adaptive cluster segmentation aircraft objects in aerial images," *Fifth World congress on Intelligent Control and Automation*, vol. 6, pp. 5415-5418, 2004.

[9] Q. Lin and X. Zhang, "Video object clustering segmentation," *International Conference on Machine Learning and Cybernetics*, vol. 5, pp. 2840-2843, 2003.

[10] G. I. Sanchez-Ortiz, G. J. T. Wright, N. Clarke, J. Declerck, A. P. Banning and J. A. Noble, "Automated 3-D echocardiography analysis compared with manual delineations and SPECT MUGA," *IEEE Transactions on Medical Imaging*, vol. 21, issue 9. pp. 1069-1076, 2002.

[11] J. Chen, T. N. Pappas, A. Mojsilovic, and B. E. Rogowitz, "Image segmentation by spatially adaptive color and texture features," *International Conference on Image Processing*, vol. 1, pp. I1005-1008, 2003.

[12] J. Timmis, M. Neal, J. Hunt, "An artificial immune system for data analysis," *Bio Systems*, vol. 55, pp. 143-150, 2000.

[13] J. Timmis and M. Neal, "A resource limited artificial immune system for data analysis," *Knowledge-Based Systems*, vol. 14, Issues 3-4, June 2001, pp. 121-130.

[14] J. D. Farmer, N. H. Packard, and A. S. Perelson, "The Immune System, Adaptation, and Machine Learning,", *Physical 22D*, pp. 187-204, 1986.

[15] J. D. Farmer, S. A. Kauffman, N. H. Packard, and A. S. Perelson, "Adaptive Dynamic

Networks as Models for the Immune System and Autocatalytic Sets," *Annual of the N. Y. Academics of Science*, vol. 504, pp. 118-131, 1987.
[16] F. J. Varela and A. Coutinho, "Second Generation Immune Networks," *Immune Today*, vol. 12, no.5, pp. 159-166, 1991.
[17] A. S. Perelson, "Immune Network Theory," *Immune revolution*, vol. 110, pp. 5-36, 1989.
[18] E. Anderson, "The irises of the Gaspé peninsula," *Bulletin of the American Iris Society*, vol. 59, pp. 2-5, 1935.
[19] R. A. Fisher, "The use of multiple measurements in taxonomic problems," *Annals of Eugenics*, vol. 7, pp. 179-188, 1936.
[20] P. K. Simpson, "Fuzzy min-max neural networks-Part 1: Classification," IEEE Transactions on Neural Networks, vol. 3, pp. 776-786, 1992.
[21] P. K. Simpson, "Fuzz min-max neural networks-Part 2: Clustering," IEEE Transactions on Fuzzy System, vol. 1, pp. 32-45, 1993.

Table 1. The relationship among k1, k2, the number of clusters, and error

| Threshold | k1 | k2 | Cluster number | Error number | Error rate (%) |
|---|---|---|---|---|---|
| 0.6 | 0.7 | 0.8 | 17 | 5 | 3.33 |
| 0.6 | 0.7 | 0.9 | 20 | 5 | 3.33 |
| 0.6 | 0.7 | 1.0 | 21 | 5 | 3.33 |
| 0.7 | 0.8 | 0.6 | 20 | 5 | 3.33 |
| 0.7 | 0.8 | 0.7 | 22 | 5 | 3.33 |
| 0.8 | 0.9 | 0.5 | 22 | 5 | 3.33 |
| 0.9 | 1.0 | 0.3 | 20 | 5 | 3.33 |

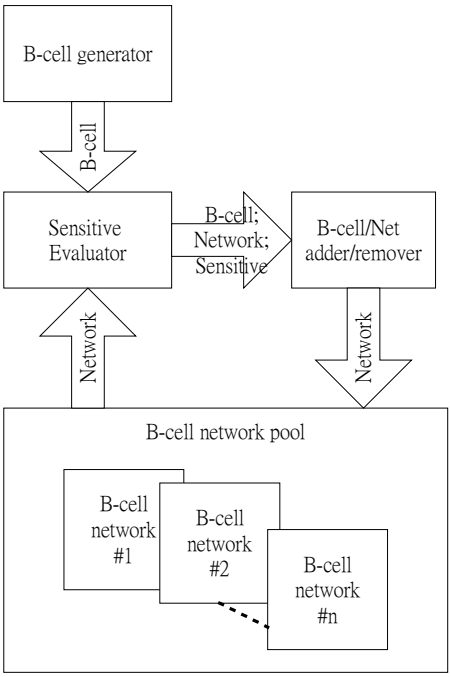

Fig. 1 System architecture

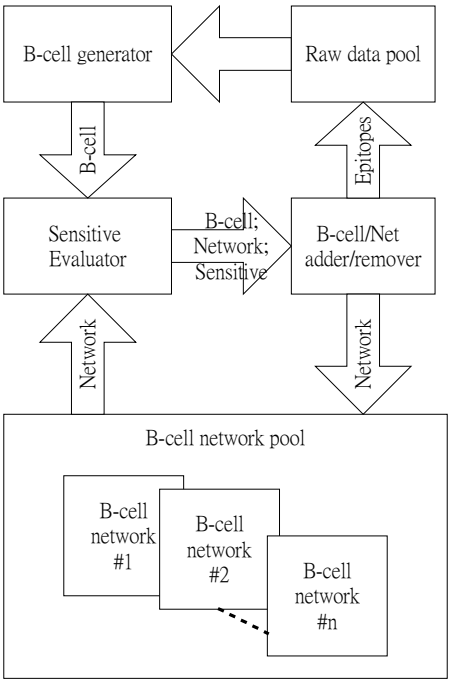

Fig. 2 The architecture of system considering limited resources

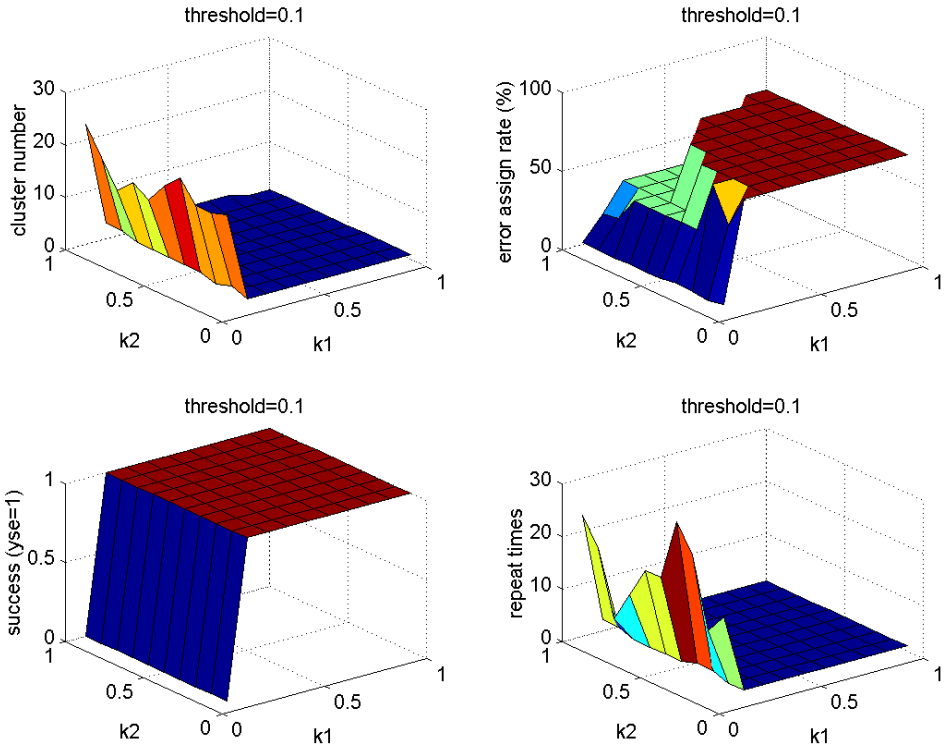


Fig. 3 The success, repeat times, error, and cluster number under different k1 and k2 while threshold = 0.1

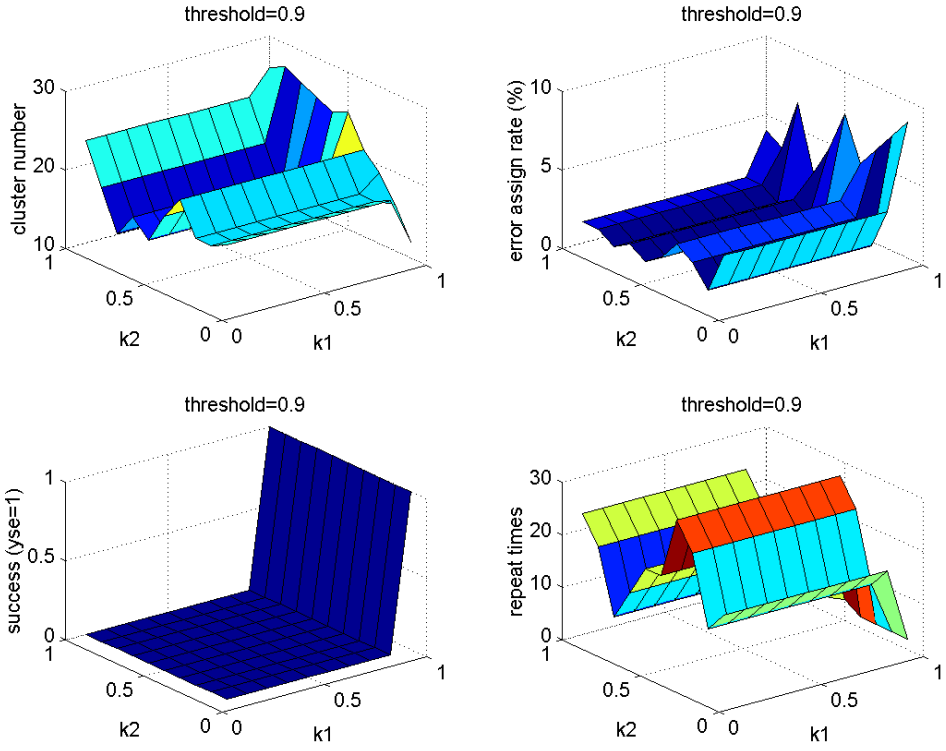


Fig. 4 The success, repeat times, error, and cluster number under different k1 and k2 while threshold = 0.9

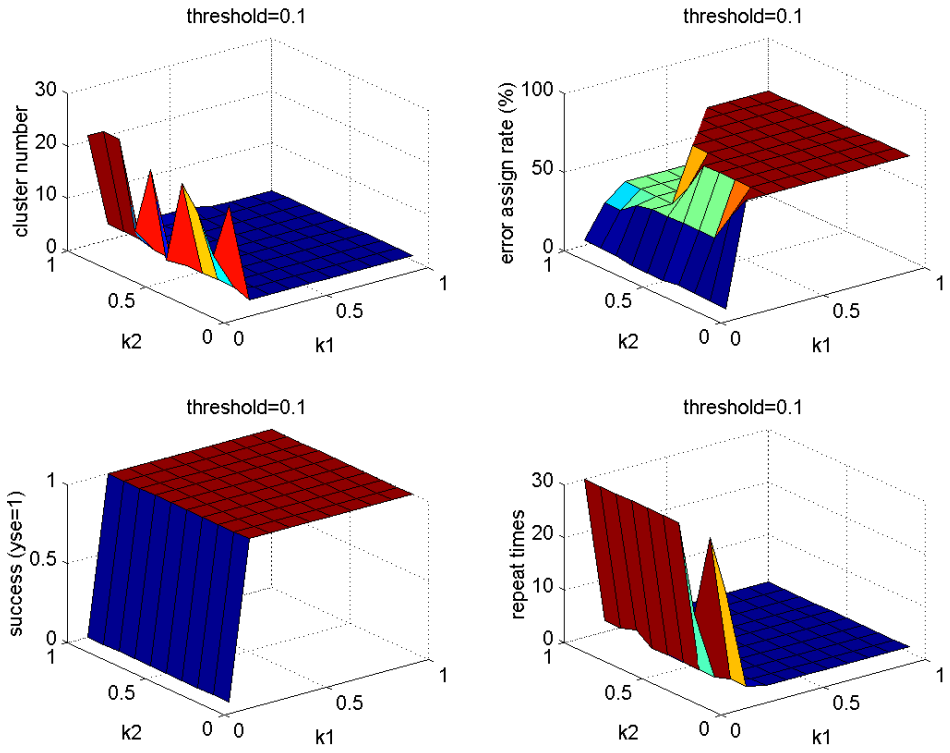


Fig. 5 The success, repeat times, error, and cluster number under different k1 and k2 while threshold = 0.1

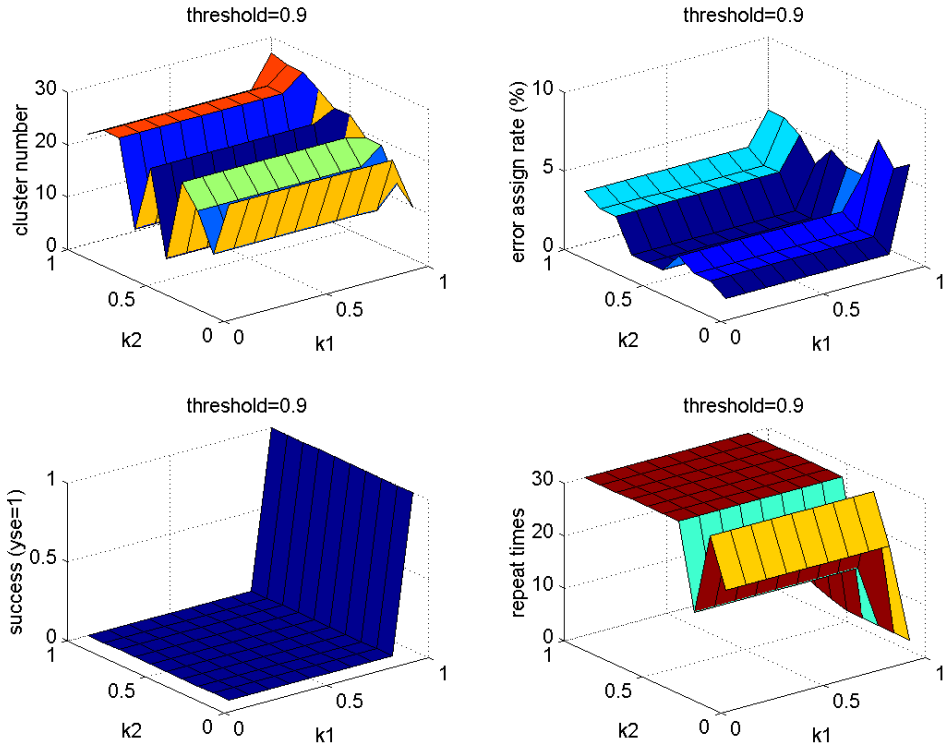


Fig. 6 The success, repeat times, error, and cluster number under different k1 and k2 while threshold = 1.0

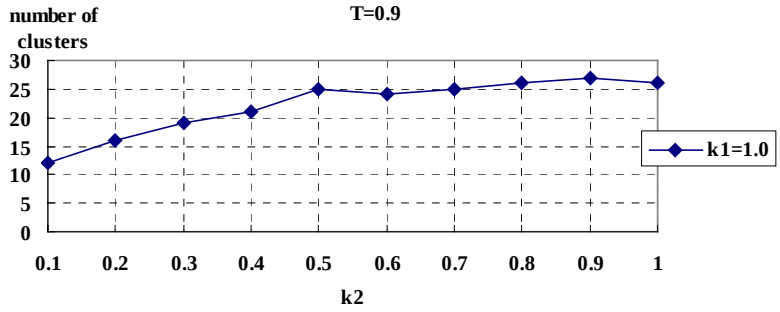


Fig. 7 The number of clusters under different k1 and k2 while successful flag = 1 and threshold = 0.9

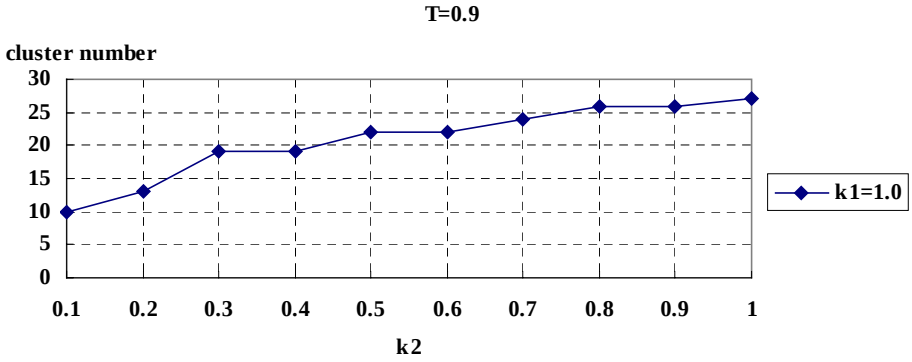


Fig. 8 Number of clusters under different k1 and k2 whilst success flag = 1 and threshold = 0.9

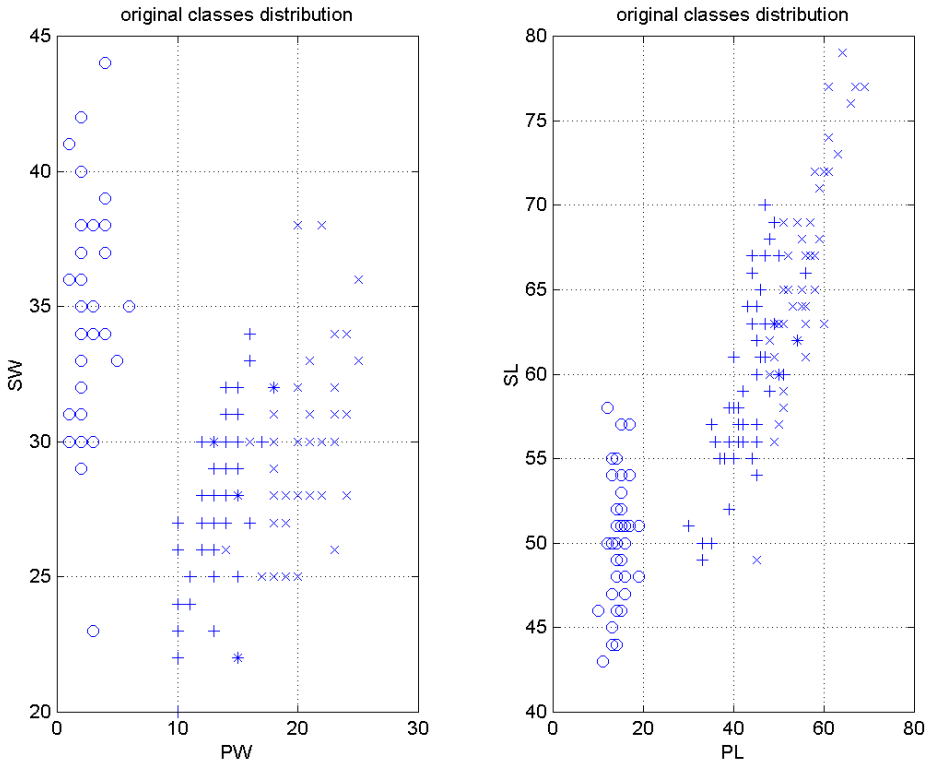


Fig. 9 Distribution of features (PW: petal width; SL: sepal length; SW: sepal width; PL: petal length)

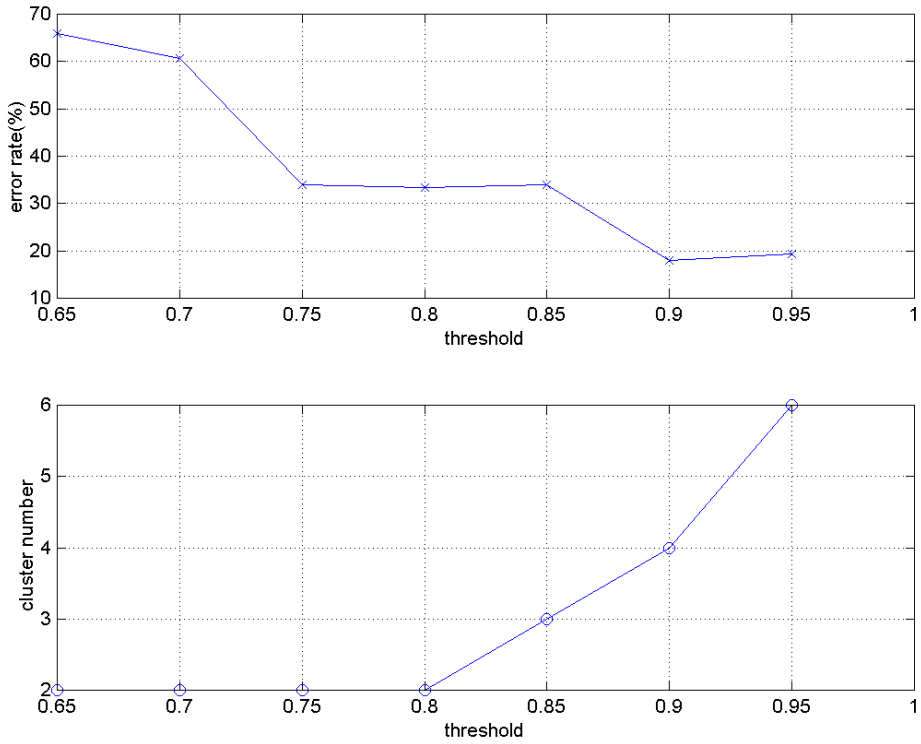


Fig. 10 The error rate and cluster number of ART at different threshold values

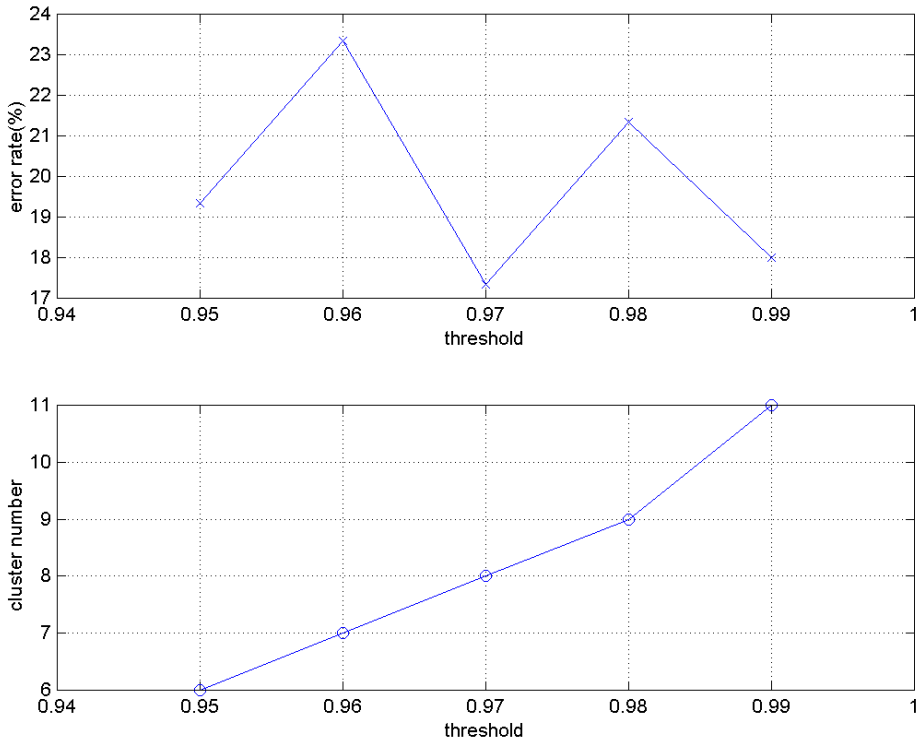


Fig. 11 The error rate and cluster number of ART under threshold values in 0.95~0.99 with step 0.01