

A New Adaptive Elastic Net Method for Cluster Analysis

J. Yi, P. Zhao, L. Zhang, G. Yang

Junyan Yi

Beijing Key Laboratory of Robot Bionics and Function Research,
Department of Computer Science and Technology
Beijing University of Civil Engineering and Architecture
1 Zhanlanguan Road
Beijing, 100044, China
yijunyan@bucea.edu.cn

Peixin Zhao*

School of Management,
Shandong University, Jinan, China
27 Shanda Nan Road
Jinan, 250100, China
*Corresponding author: pxzhao@sdu.edu.cn

Lei Zhang

Department of Computer Science and Technology
Beijing University of Civil Engineering and Architecture
1 Zhanlanguan Road
Beijing, 100044, China
lei.zhang@bucea.edu.cn

Gang Yang

School of Information
Renmin University of China
59 Zhongguancun Street
Beijing, 100872, China
yanggang@ruc.edu.cn

Abstract: Clustering is inherently a highly challenging research problem. The elastic net algorithm is designed to solve the traveling salesman problem initially, now is verified to be an efficient tool for data clustering in n -dimensional space. In this paper, by introducing a nearest neighbor learning method and a local search preferred strategy, we proposed a new Self-Organizing NN approach, called the Adaptive Clustering Elastic Net (ACEN) to solve the cluster analysis problems. ACEN consists of the adaptive clustering elastic net phase and a local search preferred phase. The first phase is used to find a cyclic permutation of the points as to minimize the total distances of the adjacent points, and adopts the Euclidean distance as the criteria to assign each point. The local search preferred phase aims to minimize the total dissimilarity within each clusters. Simulations were made on a large number of homogeneous and nonhomogeneous artificial clusters in n dimensions and a set of publicly standard problems available from UCI. Simulation results show that compared with classical partitional clustering methods, ACEN can provide better clustering solutions and do more efficiently.

Keywords: self-organizing neural network, elastic net, adaptive, cluster analysis.

1 Introduction

Cluster analysis is useful in exploratory data analysis. It is an important but challenging task in unsupervised learning. The essence of the task is to partition a set of objects into a number of

clusters, minimizing the within-cluster variability and maximizing the between-cluster variability [10]. Data clustering is a common technique for statistical data analysis. It can be widely used in a variety of scientific and engineering disciplines such as biology [2, 19, 22], computer vision [6, 8, 9, 17] and optimizations [18, 24]. Many clustering algorithms have been proposed. Generally they can be divided into the following categories: hierarchical clustering method, partitional clustering method, density based clustering method, grid based clustering method, and neural network based clustering method and so on [23].

In the last 20 years also, neural networks have been shown to be powerful tools for solving clustering problems. At present, neural networks-based clustering has been dominated by SOFM (Self-Organization Feature Map) and ART (Adaptive Resonance Theory). SOFM is a powerful tool for data analysis, and used widely in data mining, pattern recognition, clustering analysis [1, 11, 12, 16]. But when applied in real practice a number of user-dependent parameters cause problems. Like the K-means algorithm, SOMF need to predefine the size of the lattice, i.e., the number of the clusters, which is unknown for most circumstances. In addition, trained SOMF may be suffering from input space density misrepresentation, where areas of low pattern density may be over-represented, and areas of high density under-represented. Some researchers have proposed modifications of SOMF algorithm [4, 14].

In this work we propose a new Self-Organizing NN clustering method based on the elastic net, and attempt to alleviate the problems mentioned above. Elastic net is evolved from SOFM, proposed by R. Durbin and D. Willshaw [5] as a heuristic method. Like SOFM, the elastic net tries to find a topology preserving map between two spaces. It was originally proposed to solve the Traveling Salesman Problem (TSP) [20]. Later it was given a mechanical statistic foundation [15]. Then elastic net can be used as a tool for data clustering in n -dimensional space and for diverse of problems in pattern research [3, 7, 13, 21].

The advantages of the elastic net method for cluster analysis are as follows: firstly, elastic net is an unsupervised learning network, and the manual training is not required. Secondly, elastic net does not need the prior knowledge about the number of clusters, and its effectiveness is not effected by the input sequence of data and isolated data. Thirdly, elastic net's geometry corresponds well to the problem definition, which makes it suitable for solving high-dimensional clustering problems. And its convergence is governed by well-established physics theories leading to a sensible solution. Lastly, when solving the clustering problems, the number of neurons needed in the elastic net is linearly related to the number of spatial data points. That is, the complexity of the elastic network algorithm is $O(N)$, where N is the number of spatial data points [20]. The performance of the proposed method is evaluated by simulations on a number of homogeneous and nonhomogeneous synthetic clustering problems in n dimensions and some sets of publicly standard test problems available from UCI (accessible via the web at <http://archive.ics.uci.edu/ml/datasets.html>). Simulation results show that the proposed method can consistently and efficiently provide better clustering solutions compared with the classical partitional clustering algorithms.

2 The new adaptive clustering elastic net method

In the section, a new Self-Organizing NN approach for clustering ACEN is presented. ACEN consists of two phases: the adaptive clustering elastic net phase and local search preferred phase. Let $x_i (i = 1, 2, \dots, n)$ be a data set X with n objects; $k \leq n$ the number of clusters; $m_j \leq n$ the centroid of cluster $C_j (j = 1, 2, \dots, K)$. Then ACEN tries to minimize the cost function - Sum of

Euclidean Distance (SED):

$$SED = \sum_{i=1}^n \sum_{j=1, x_i \in C_j}^k d(x_i, m_j) \quad (1)$$

The details of two phases are described below.

2.1 The adaptive clustering elastic net phase

Given m unit coordinates $Y = \{y_1, y_2, \dots, y_m\}$ defining the rubber band of the elastic net, n data points $X = \{x_1, x_2, \dots, x_n\}$ defining the clustering data set. The new energy function of the adaptive clustering elastic net is defined as below:

$$E = -\alpha(t)K \sum_{i=1}^n \ln \sum_{j=1}^m e^{-|x_i - y_j|/2K^2} + \beta(t) \sum_{j=1}^m |y_j - y_{j+1}| \quad (2)$$

The positions of the points defining the rubber band are updated according to the following formula:

$$\Delta y_j = \alpha(t) \sum_{i=1}^n w_{ij}(x_i - y_j) + \beta(t)K(y_{j+1} - 2y_j + y_{j-1}) \quad (3)$$

$$\begin{cases} \alpha(t+1) = (1 - \Psi)\alpha(t) & \text{if } \alpha(t) > \alpha_{min}, \\ \alpha(t+1) = \alpha_{min} & \text{otherwise.} \end{cases} \quad (4)$$

$$\begin{cases} \beta(t+1) = (1 + \vartheta)\beta(t) & \text{if } \beta(t) < \beta_{max}, \\ \beta(t+1) = \beta_{max} & \text{otherwise.} \end{cases} \quad (5)$$

where

Ψ = damping factor of $\alpha(t)$ ($0 \leq \Psi < 1$).

ϑ = promoting factor of $\beta(t)$ ($0 \leq \vartheta < 1$).

Here w_{ij} is the same as the original elastic net. The parameters Ψ and ϑ in Eqn.(4) and (5) are small positive parameters, which are selected empirically.

This phase helps the elastic net to find a cyclic permutation of the points so as to minimize the total distance of adjacent points in shorter time. After the computation of the adaptive clustering elastic net phase, each element of the data set is matched with the points of the elastic band. Therefore the permutation of data points can be obtained from the sequence of the elastic net points. Then the distances of adjacent elements of the data set can be used to divide the clusters.

Each point dotted in Fig.1 represents the distances between the adjacent elements of the permutation. The value of the x -axis is the sequence of elements of the data set. The value of the y -axis is the distance between the adjacent elements of the data set. Therefore, each pick in Fig.1 represents a larger distance between the adjacent elements, meaning a cluster change.

For large data sets, some elements may be lost in the final permutation. These elements are regarded as noise elements in our algorithm. We proposed the nearest neighbor learning method to handle the noise elements. The learning method is described as below. For each noise element, find its nearest data x_i in all of the matched elements of the data set, and calculate the distances between the noise element and the two neighbors x_{i-1} and x_{i+1} in the path. Compare the two distances and insert the noise element between x_i and the nearer neighbor element x_{i-1} or x_{i+1} . Then update the sequence of the data set.

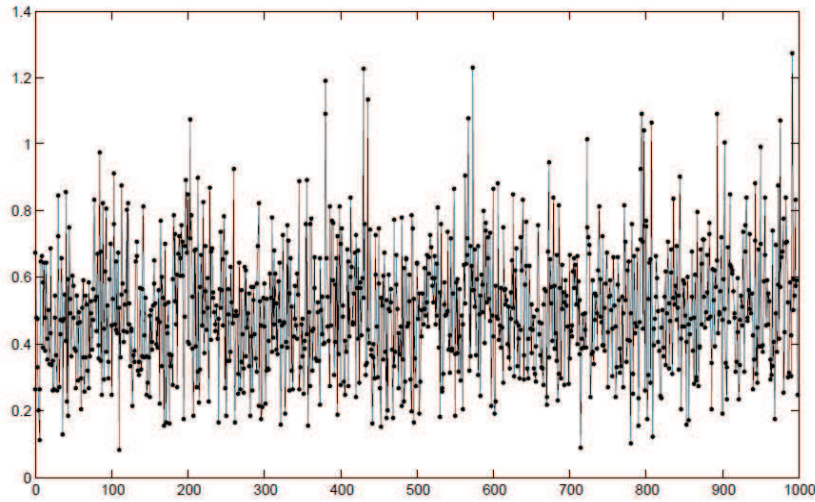


Figure 1: The distances between the adjacent elements of the data set.

2.2 The local search preferred phase

In the first stage, we could achieve initial clusters according to the distances between the adjacent points. For most of the clustering problems, the first phase could efficiently identify high quality solutions. According to the requirements of different practical applications, the clustering solutions sometimes still need to be further optimized. Then the local search preferred phase is needed to help the network to search for the better solutions. Mathematically the steps of the second phase can be stated as below.

Step 1. Select the number of the nearest neighbors p and get the centroids m_j of each cluster C_j in the first phase.

Step 2. Assign each element in the data set X to the clusters C_j with the closest centroid m_j using the Euclidean distance as the criteria.

Step 3. Update k medoids. For $j = 1$ to the number of clusters k do

(a) Find a subset C_{subset} in the cluster C_j . C_{subset} corresponds to m_j and its p nearest neighbors that have not been evaluated before current iteration.

(b) Calculate the new medoid

$$q = \arg \min_{x_i \in C_{subset}} \sum_{x_i \in C_j} d(x_k, x_i) \quad (6)$$

after that the old medoid m_j is replaced by q if it is different from m_j .

(c) Repeat steps (a) and (b) until the medoid does not change any more.

Step 4. Repeat steps 2 and 3 until the k medoid does not change any more.

The local search preferred phase uses the Euclidean distance as the criteria to assign each point like many other partitional clustering algorithms. The algorithm chooses the objects in the data set as the medoids of the clusters. The selection of the initial centroids in the first phase helps the algorithm to overcome the disadvantage of the instability of the other partitional algorithms, and could achieve better solutions.

3 Algorithm

For the clustering problem, the proposed algorithm ACEN is presented as below:

Step 1. Given the initial number m of dynamic nodes $Y = \{y_1, y_2, \dots, y_m\}$ and establish the nodes in a small circle with the center of the circle at the centroid of the clustered points. $m = 2.5n$.

Step 2. Decide the sequence of the nodes.

Step 3. For each data points $x_i (i = 1, 2, \dots, n)$, choose a node and move the node and its neighbors on the circle towards the points.

Step 4. Repeat until all the points $x_i (i = 1, 2, \dots, n)$ have been matched with the nodes.

Step 5. Get the sequence of the $x_i (i = 1, 2, \dots, n)$ according to the sequence of the nodes.

Step 6. Use the nearest neighbor learning method to update the sequence of data set.

Step 7. $d_i (i = 1, 2, \dots, n)$ (the distance of the adjacent data points) is used to divide the groups $C_j (j = 1, 2, \dots, k)$, where larger distances d_i mean cluster changes.

Step 8. Assign each point in the data set X to the clusters C_j with the closest centroid m_j under Euclidean distance metric.

Step 9. Update k medoids. For $j = 1$ to the number of clusters k do

(a) Find a subset C_{subset} in the cluster C_j . C_{subset} corresponds to m_j and its p nearest neighbors that have not been evaluated before current iteration.

(b) Calculate the new medoid

$$q = \arg \min_{x_i \in C_{subset}} \sum_{x_i \in C_j} d(x_k, x_i) \quad (7)$$

after that the old medoid m_j is replaced by q if it is different from m_j .

(c) Repeat steps (a) and (b) until the medoid does not change any more.

Step 10. Repeat steps 8 and 9 until k medoid does not change any more.

Our algorithm improves the original elastic net method by introducing the adaptive clustering parameter strategy and the nearest neighbor learning method. Combining the new elastic net method with the local search preferred strategy, the proposed algorithm could obtain better solutions for cluster analysis problems.

The advantages of ACEN mainly reflected in the following aspects:

- ACEN does not need the prior knowledge about the number of clusters.
- Algorithm's effectiveness is not effected by the input sequence of data and isolated data.
- Given n values of the number of clusters k , ACEN could obtain the different solutions for different numbers of clusters just running the network one time, which effectively improve the computation speed. On the contrary, given n values of the number of clusters k , the traditional partitional clustering algorithms such as K-means and K-medoids need to run n times.
- Due to the geometrical and physical characteristics of the elastic net, ACEN could express and calculate high-dimensional data set simply and efficiently, and display the data intuitively.
- By introducing the adaptive parameter strategy and the nearest neighbor learning method, combined with the local search preferred method, ACEN could consistently and efficiently identify high quality clustering solutions.
- The algorithm is geometric in nature, and the evolution of the algorithm can be tracked visually.

It has been pointed out that rearrangement clustering is equivalent to the TSP. And it can be solved to optimality by solving the TSP [3]. But there is a serious defect in previous approaches when applied to data which falls into natural clusters as shown in Fig.2. Objects in Fig.2 only have two features: horizontal and vertical coordinates. When these objects are clustered by the method of rearrangement clustering, although x and y are very close the objects x and y will be separated by 16 objects in two different clusters. It is clear that generally clusters may be broken into pieces so as to minimize the dissimilarity to adjacent clusters. Using the ACEN method these problems could be solved effectively. In ACEN the partition is taken based on the larger distances between the adjacent elements in the permutation, which is more suitable for the definition of the natural cluster problem, which could help the algorithm to achieve high quality clustering solutions.

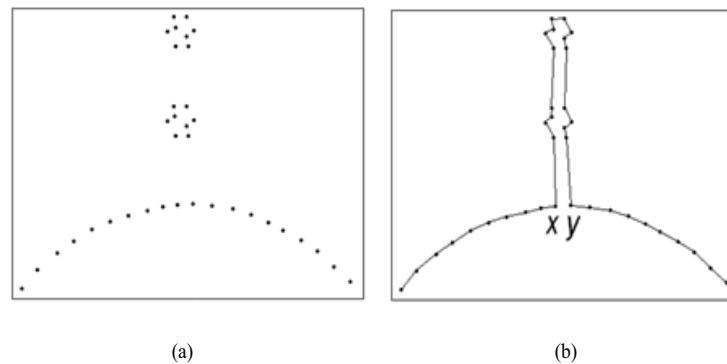


Figure 2: (a) Three clusters; (b) The TSP path specifying the optimal rearrangement. Although x and y are very close, their placement in the rearrangement is far apart.

4 Simulation results

Extensive simulations were implemented on a large number of homogeneous and nonhomogeneous synthetic clusters in n dimensions. And the value of SED shown in Eqn. (1) was adopted to judge the quality of the different methods. Parameters were set as follows: $\alpha(0) = 0.6$; $\beta(0) = 1.5$; $\Psi = 0.001$; $\vartheta = 0.002$; $k = 0.5$.

We tested the ACEN on some nonhomogeneous clusters with different density regions. First we compared the ACEN with K-means and K-medoids methods to solve the problem shown in Fig.2. These data falls into natural clusters. ACEN identified high quality clustering solution for this simple problem. On the contrary the other two partitioning clustering methods did not cluster the data well as predicted. The typical results produced by three methods were shown in Fig.3. In the following figures, the different shapes and colors of points represent different clusters.

The second of the comparative computational tests we report used the nonhomogeneous cluster points in 2D and 3D respectively. The problems both involved 200 points. To compare the performance of the three methods, we recorded the average SED values obtained from 20 simulations with different initial cluster centroids for k-means and k-medoids. Because the results of ACEN were not effected by the initial cluster centroids, we only took one trial for ACEN. The comparisons of ACEN with K-means and K-medoids were shown in Fig.4 and Fig.5.

From Fig.4 and Fig.5 we can see clearly:

- ACEN could identify high-quality solutions consistently for all of the tested problems. The

other two partitional clustering methods both get poor solutions.

- The values of SED produced by ACEN were 29.58% and 41.74% less than the average values of SED produced by K-means over 20 trials, and 28.18% and 45.54% less than the average values of SED produced by K-medoids over 20 trials respectively for the tested 2D and 3D clustering problems.

The proposed method consists of two phases: the adaptive clustering elastic net phase and the local search preferred phase. The first phase can identify clusters according to the distance between the adjacent points. The aim of the second phase is to minimize the cost function - Sum of Euclidean Distance (SED) in Eqn.(1). For most of the cases, solutions produced by the first phase are good enough, then the second phase is not necessary.

In the following, we clustered 500 points into 4 clusters in 2D using the proposed algorithm. The result of the first phase and the optimized result produced by the second phase were shown in Fig.6. From Fig.6 we can see clearly that the solution produced by the first phase was the same as the optimized result produced by the second phase, and ACEN could achieve high quality solution for the clustering problem without using the local search preferred phase.

To meet the requirements of different practical applications, the local search preferred phase is used to improve the solution quality. We tested the proposed method on a homogeneous artificial clustering problem involve 150 data points in 2D. Fig.7 shows the evolutions of clustering process using the proposed approach for this homogeneous artificial problem. Fig.7(a) shows the result of the adaptive clustering elastic net phase. Fig.7(b) and Fig.7(c) display the intermediate processes of clustering. Fig.7(d) shows the final solution produced by the local search preferred phase of the ACEN for the clustering problem. From Fig.7 we can see clearly that the proposed algorithm could identify high quality solutions of giving clustering problem, and the decrease of the value of SED showed that the local search preferred phase could optimize the result effectively.

In the third of our experimental tests we selected a large number of instances of 50, 100, 300, 500 and 1000 homogeneous and nonhomogeneous simulated cluster points in 2D, 3D, 4D, 7D 10D and 13D to show the solution qualities of ACEN, K-means and K-medoids. In the following experiments, to compare performance of the three methods, we recorded the average SED values obtained from 20 simulations with different initial cluster centroids for k-means and k-medoids. As described before, ACEN was performed only once. In the proposed method we use another set of parameter values as follows: $\alpha(0) = 0.6$; $\beta(0) = 1.5$; $\Psi = 0.0025$; $\vartheta = 0.0025$. Comparisons of the proposed method with the K-means and K-medoids methods were shown in Table 1 and Table 2.

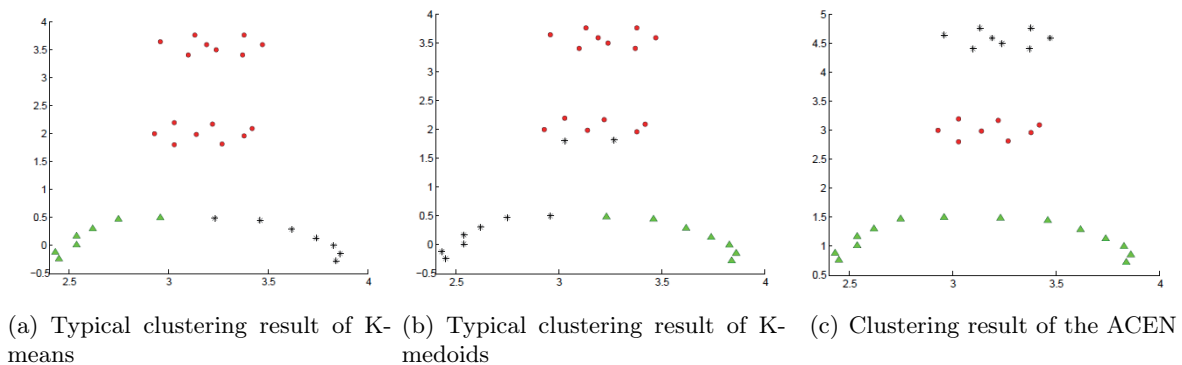


Figure 3: The typical clustering results of three methods for the natural cluster problem described in Fig. 2: (a) K-means method; (b) K-medoids method; (c) The proposed method.

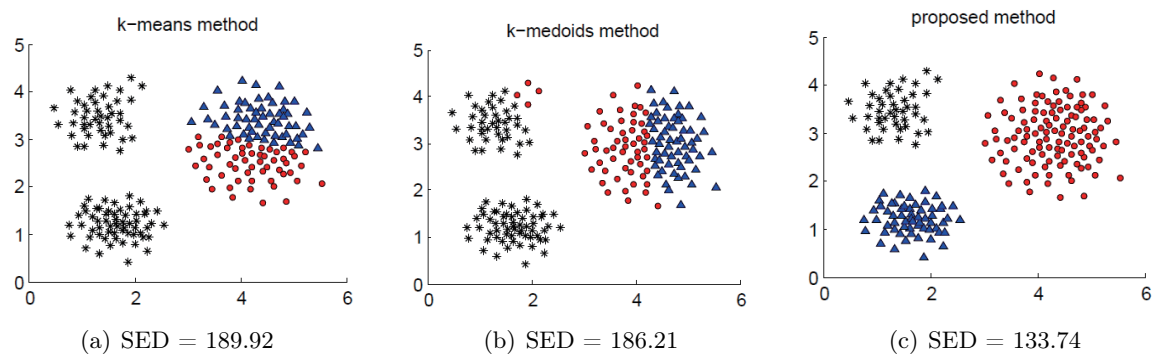


Figure 4: The typical solutions and the average values of SED obtained from 20 simulations produced by three methods for the nonhomogeneous data in 2D: (a) K-means method; (b) K-medoids method; (c) The proposed method.

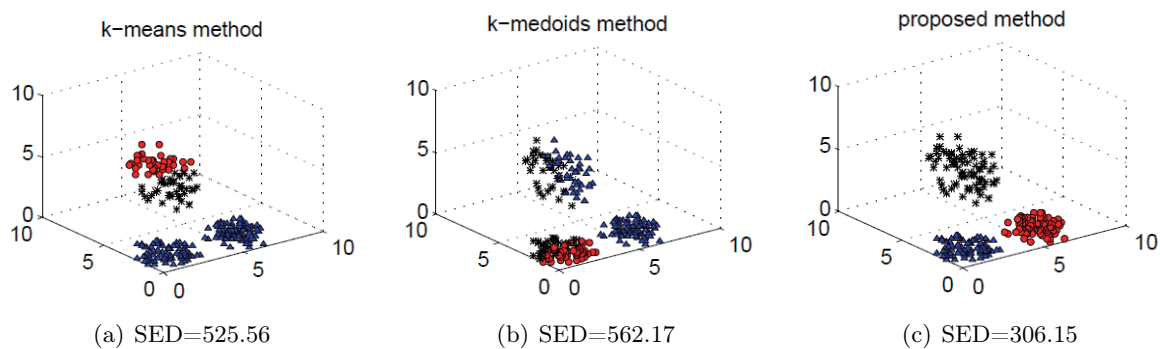


Figure 5: The typical solutions and the average values of SED obtained from 20 simulations produced by three methods for the nonhomogeneous data in 3D: (a) K-means method; (b) K-medoids method; (c) The proposed method.

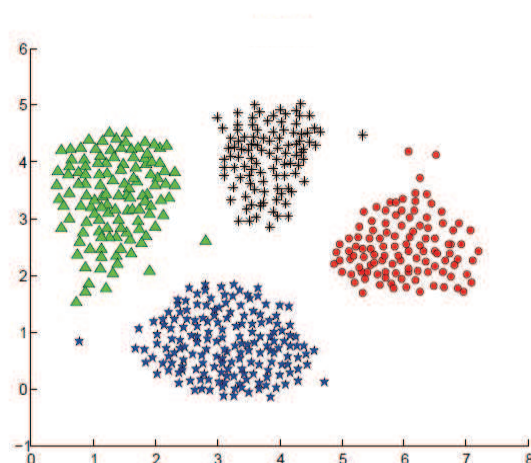
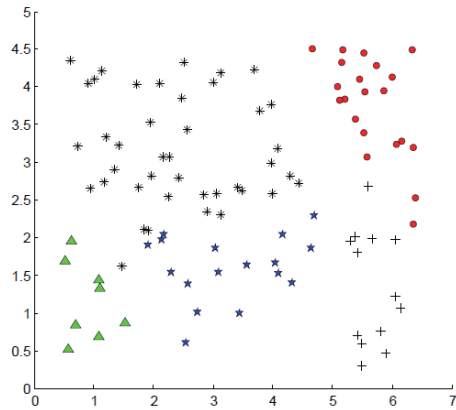
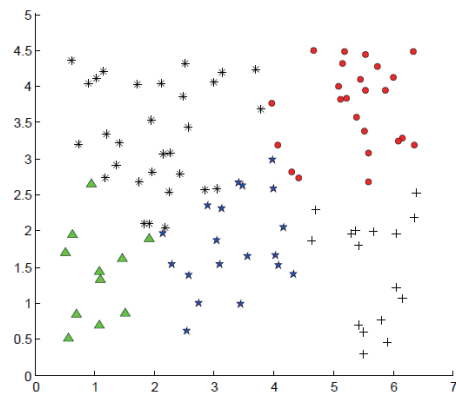


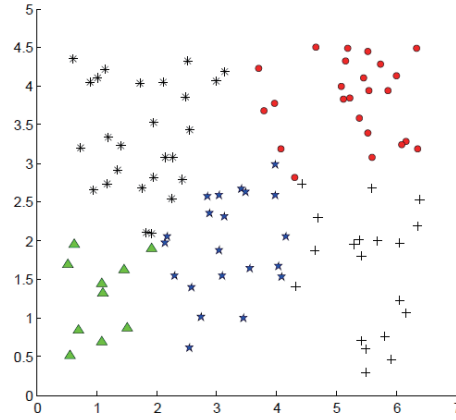
Figure 6: The result produced by the first phase and the optimized result produced by the second phase of ACEN. The two results were the same.



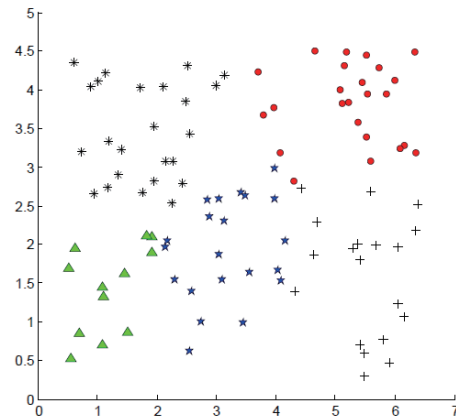
(a) The result of the adaptive clustering elastic net phase: SED=88.16



(b) The intermediate processes of clustering: SED=84.46



(c) The intermediate processes of clustering: SED=82.98



(d) The final result of the local search preferred phase: SED=82.86

Figure 7: Evolutions of ACEN and the values of SED of the clustering process.

Table 1: The average values of SED over 20 trials produced by K-means and K-medoids and the value of SED produced by ACEN on sets of 50, 100, 300, 500 and 1000 synthetic data in 2D, 3D and 4D.

Data Size	Number of Clusters	Dimension 2			Dimension 3			Dimension 4		
		Kmeans SED	Kmedoids SED	ACEN SED	Kmeans SED	Kmedoids SED	ACEN SED	Kmeans SED	Kmedoids SED	ACEN SED
50	3	32.33	29.8	20.85	46.82	42.19	29.66	44.28	44.27	36.83
	5	31.87	29.37	15.59	26.17	25.72	19.47	37.48	38.04	29
	10	19.35	16.85	7.29	23.16	21.4	11.39	24.54	24.7	17.4
100	3	76.85	73.98	45.21	75.41	74.14	67.7	109.27	109.38	109.27
	5	68.76	64.38	39.71	104.37	96.42	83.05	69.8	68.77	54.09
	10	37.72	31.45	20.88	67.02	55.91	35.95	62.36	59.54	44.91
300	3	316.14	288.72	156.74	252.86	253.44	252.32	381.8	351.98	247.73
	5	186.23	178.6	99.63	249.96	234.22	191.74	271.27	269.27	233.84
	10	114.38	109.69	80.17	146	14.25	127.76	184.75	183.44	173.1
500	3	645.76	505.98	221.87	564.6	595.15	388.7	654.73	698.28	448.91
	5	473.01	346.35	208.93	433.16	410.04	336.76	474.99	488.12	426.57
	10	263.33	225.86	158.97	301.64	288.96	261.16	346.28	365.4	309.63
1000	3	739.43	714.95	430.77	984.59	809.36	797.51	1216.3	1331.27	916.36
	5	639.8	616.95	396.37	781.41	724.79	693.75	969.57	926.82	869.2
	10	419.75	405.18	345.37	646.31	655.78	628.33	806.21	811.8	692.55

An examination of Table 1 and Table 2 yields the following observations:

- For all of the 30 homogeneous and nonhomogeneous clustering challenges, ACEN could obtain better clustering solutions than K-means and K-medoids.
- The values of SED calculated by the proposed method around 50, 100, 300, 500 and 1000 points for different dimensions were 35.03%, 31.2%, 30.25%, 30.19% and 29.53% less than the average values of SED calculated by the K-means over 20 trials; and 33.89%, 29.33%, 27.34%, 27.23% and 25.87% less than the average values of SED calculated by the K-medoids over 20 trials respectively.
- Like many partitional clustering algorithms, the results of K-means and K-medoids depended on the selection of the initial cluster centroids. On the contrary, the results of ACEN were stable and were not influenced by the selection of initial centroids.
- ACEN's effectiveness was not effected by the input sequence of data and isolated data, and could produce stable clustering solutions .

For the last of our experiments we took some sets of standard test problems available from UCI (accessible via the web at <http://archive.ics.uci.edu/ml/datasets.html>). The standard problems we solved include Iris data set and Zoo data set. Comparisons of ACEN with K-means and K-medoids on standard problems were summarized in Table 3. For this comparison, the value of SED and the accuracy were adopted to judge the quality of the clustering solutions. In Table 3, we also performed 20 simulations for K-means and K-medoids, and the average values of SED and accuracy were calculated for each standard problem. Because the results of ACEN were stable and were not effected by the initial centroids, it was performed only once.

The Iris Plants Database is the best known database found in the literature of cluster analysis and referred frequently in cluster analysis texts. The data set contains 3 classes of iris plant with

Table 2: The average values of SED over 20 trials produced by K-means and K-medoids and the value of SED produced by ACEN on sets of 50, 100, 300, 500 and 1000 synthetic data in 7D, 10D and 13D.

Data Size	Number of Clusters	Dimension 7			Dimension 10			Dimension 13		
		Kmeans SED	Kmedoids SED	ACEN SED	Kmeans SED	Kmedoids SED	ACEN SED	Kmeans SED	Kmedoids SED	ACEN SED
50	3	60.17	55.51	41.5	67.22	69.71	55.42	78.22	78.66	61.4
	5	57.21	67.12	37.61	58.7	60.28	45.58	72.84	69.75	52.91
	10	38.55	30.03	27.15	44.34	47.28	29.03	57.13	56.53	30.42
100	3	129.37	110.65	107.62	145.55	138.62	123.93	150.66	137.54	141.61
	5	119.03	118.4	98.99	147.81	133.84	128.27	157.99	135.34	133.59
	10	88	8740	63.89	108.39	108.49	75.16	121.19	116.77	91.71
300	3	468.58	492.8	376.95	594.06	539.17	444.53	707.93	621.61	513.57
	5	386.18	377.22	331.98	498.24	452.06	443.21	504.93	472.72	423.52
	10	335.11	332.33	292.95	420.15	383.47	352.41	421.29	419.52	376.32
500	3	925.75	887.49	726.5	918.19	864.07	759.74	1071.6	1035.17	945.52
	5	691.72	676.57	614.96	865.06	859.41	716.71	862.08	876.33	707.35
	10	618.28	586.41	530.67	774.71	660.53	615.03	834.16	761.27	694.91
1000	3	1603.92	1547.7	1327.82	2082.18	1812.84	1563.24	1945.86	1822.22	1713.09
	5	1374.88	1306.2	1166.27	1638.95	1599.03	1391.72	1981.69	1743.76	1507.95
	10	1289.32	1287.17	956.41	1565.89	1404.95	1135.8	1657.39	1520	1371.93

Table 3: Computational results- standard UCI clustering problems.

Algorithm	Iris		Zoo	
	SED	Accuracy	SED	Accuracy
K-means	125.74	69%	133.63	66%
K-medoids	123.66	71%	127.52	68%
ACEN	98.95	91%	102.38	90%

50 instances each. The Zoo data set contains 7 classes and each animal is made up of 16 properties, 15 of them as boolean and 1 numeric property. In the tests we set the parameter values as described before.

Considering Table 3 we can make the point: ACEN was much more efficient than the other two partitional clustering methods. The accuracies of the ACEN algorithm for standard UCI clustering problems Iris and Zoo were 22% and 24% higher than K-means, and 20% and 22% higher than K-medoids respectively.

Our experiments clearly illustrate the improvement of the clustering solution quality using the ACEN compared with the traditional partitional clustering methods. Moreover the results produced by ACEN are not effected by the input sequence of data and isolated data, which ensure it consistently and efficiently identify high quality solutions for different problems.

5 Conclusions

In this paper, we proposed a new Self-Organizing neural network approach-the Adaptive Clustering Elastic Net (ACEN) for cluster analysis. Our approach has two phases, the adaptive clustering elastic net phase and the local search preferred phase. The first phase is proposed by

introducing an adaptive clustering parameter strategy and the nearest neighbor learning method into the original elastic net. This phase is used to find a cyclic permutation of the points so as to minimize the total distances of the adjacent points, and assign each point to different clusters using the Euclidean distance as the criteria. The local search preferred phase aims to minimize the total dissimilarity within each clusters.

We tested the ACEN method and some classical partitional clustering algorithms on the same data sets include homogeneous and nonhomogeneous synthetic clusters in n dimensions and some sets of publicly standard problems available from UCI. Experiments show that compared with the other traditional partitional clustering algorithms, ACEN could provide better and stable clustering solutions and do more efficiently for all of the tested problems. ACEN algorithm's effectiveness is not effected by the input sequence of data and isolated data. In addition ACEN does not need the prior knowledge about the number of clusters, which makes it more suitable for solving clustering problems.

Acknowledgment

This work is funded by National Natural Science Foundation of China (No. 61402032), Beijing Municipal Natural Science Foundation (No. 4144072), Project Supported by the Beijing Excellent Talent (No. 2013D005017000017), Scientific Research Foundation of Beijing University of Civil Engineering and Architecture (No. 00331613002 and 00331616023, Y13-23), Open Research Fund Program of Beijing Key Laboratory of Robot Bionics and Function Research, the Independent Innovation Foundation of Shandong University, IIFSDU(IFW12109), the Key Project of Philosophy and Social Science Plan of Jinan(2014), Shandong Province Natural Science Foundation(ZR2015GM012), the China Postdoctoral Science Foundation (No.2014T70624), the National Statistical Science Research Programm(2013LZ38) and Beijing Municipal Commission of Education (No. KM201410016007 and 21271413117).

Bibliography

- [1] Alahakoon D., Halgamuge S.K. (2000); Dynamic Self-organizing Maps with Controlled Growth for Knowledge Discovery, *IEEE Trans. on Neural Networks*, 11(3), 601-614, 2000.
- [2] Baldi P., Hatfield G.W.(2002); *DNA Microarrays and Gene Expression*, Cambridge University Press, 2002.
- [3] Climer S., Weixiong Zhang W. (2006); Rearrangement Clustering: Pitfalls, Remedies, and Applications, *Journal of Machiner Learning Research*, 7, 919-943, 2006.
- [4] Dittenbach M., Rauber A. (2002); Uncovering Hierarchical Structure in Data Using the Growing Hierarchical Self-organizing Map, *Neurocomputing*, 48, 199-216, 2002.
- [5] Durbin R., Willshaw D. (1987); An Analogue Approach to the Traveling Salesman Problem Using an Elastic Net Approach, *Nature*, 326, 689-691, 1987.
- [6] Frigui H., Krishnapuram R. (1999); A Robust Competitive Clustering Algorithm with Applications in Computer Vision, *IEEE Trans. Pattern Analysis and Machine Intelligence*, 21(5), 450-465, 1999.
- [7] Gorbunov S., Kisel I. (2006); Elastic Net for Standalone RICH Ring Finding, *Proceedings-published in NIM*, 559, 139-142, 2006.

-
- [8] Jain A.K., Flynn P. (1996); Image Segmentation Using Clustering, *Advances in Image Understanding*, 65-83, 1996.
- [9] Jain K. (2010); Data Clustering: 50 Years Beyond K-Means, *Pattern Recognition Letters*, 31(8), 651-666, 2010.
- [10] Kantardzic M. (2011); *Data Mining: Concepts, Models, Methods, and Algorithms*, Wiley-IEEE Press, 2011.
- [11] Kohonen T. (1982); Self-organized Formation of the Topologically Correct Feature Maps, *Biological Cybernetics*, 43, 59-69, 1982.
- [12] Kohonen T. (2001); *Self-Organizing Maps*, 3rd Ed. New York: Springer-Verlag, 2001.
- [13] Levano M.; Hans Nowak H. (2011); New Aspects of the Elastic Net Algorithm for Cluster Analysis, *Neural Comput and Applic*, 20, 835-850, 2011.
- [14] Raube A., Merkl D. (2002); The Growing Hierarchical Self-Organizing Map: Exploratory Analysis of High-Dimensional Data, *IEEE Transactions on Neural Networks*, 13(6), 1331-1340, 2002.
- [15] Rose K., Gurewitz E., Fox G. (1990); Statistical Mechanics and Phase Transitions in Clustering, *Phys Rev Lett*, 65, 945-948, 1990.
- [16] Saric T.; Simunovic, G. (2016); Estimation of Machining Time for CNC Manufacturing Using Neural Computing, *International Journal of Simulation Modelling*, 15(4), 663-675, 2016.
- [17] Shi J., Malik J. (2000); Normalized Cuts and Image Segmentation, *IEEE Trans. Pattern Analysis and Machine Intelligence*, 22(8), 888-905, 2000.
- [18] Tang M., Gong D., Liu S., Zhang H. (2016); Applying Multi-phase Particle Swarm Optimization to Solve Bulk Cargo Port Scheduling Problem, *Advances in Production Engineering and Management*, 11(4), 299-310, 2016.
- [19] Tavazoie S., Hughes D.; Campbell M.J., Cho R.J., Church G.M. (1999); Systematic Determination of Genetic Network Architecture, *Nature Genetic*, 22, 281-285, 1999.
- [20] Vakhutinsky A.I., Golden B.L. (2003); The Co-adaptive Neural Network Approach to the Euclidean Traveling Salesman Problem, *Neural Networks*, 16(10), 1499-1525, 2003.
- [21] Wang J., Tang Z., Qiping Cao, Xinshun Xu (2003); An Elastic Net Learning Algorithm for Edge Linking of Images, *IEICE Trans. Fundamentals*, E86-A(11), 2879-2886, 2003.
- [22] Wu S., Liew A.W.C., Yan, H., Yang M. (2014); Cluster Analysis of Gene Expression Database on Self-Splitting and Merging Competitive Learning, *IEEE Trans. Information Technology in Biomedicine*, doi: 10.1109/TITB.2004.824724, 8(1), 5-15, 2014.
- [23] Xu R. (2005); Survey of Clustering Algorithm, *IEEE Transaction On Neural Networks*, 16, 645-678, 2005.
- [24] Yang K. W. (2015); A Variables Clustering Based Differential Evolution Algorithm to Solve Production Planning Problem, *International Journal of Simulation Modelling*, 14(3), 525-538, 2015.