Finding and Evaluating Fuzzy Clusters in Networks

Jian Liu

LMAM and School of Mathematical Sciences, Peking University,
Beijing 100871, P.R. China
dugujian@pku.edu.cn

Abstract. Fuzzy cluster validity criterion tends to evaluate the quality of fuzzy partitions produced by fuzzy clustering algorithms. In this paper, an effective validity index for network fuzzy clustering is proposed, which involves the compactness and separation measures for each cluster. The simulated annealing strategy is used to minimize this validity index, associating with a dissimilarity-index-based fuzzy c-means iterative procedure, under the framework of a random walker Markovian dynamics on the network. The proposed algorithm (SADIF) can efficiently identify the probabilities of each node belonging to different clusters during the cooling process. An appropriate number of clusters can be automatically determined without any prior knowledge about the network structure. The computational results on several artificial and real-world networks confirm the capability of the algorithm.

Keywords: Fuzzy clustering, Validity index, Dissimilarity index, Fuzzy c-means, Simulated annealing.

1 Introduction

Recently, the structure and dynamics of networks have been frequently concerned in physics and other fields as a foundation for the mathematical representation of various complex systems [1,2,3]. Network models have also become popular tools in social science, economics, the design of transportation and communication systems, banking systems, etc, due to our increased capability of analyzing these models [4,5]. Modular organization of networks, closely related to the ideas of graph partitioning, has attracted considerable attention, and many real-world networks appear to be organized into clusters that are densely connected within themselves but sparsely connected with the rest of the networks. A huge variety of cluster detection techniques have been developed into partitioning the network into a small number of clusters [6,7,8,9,10,11], which are based variously on centrality measures, flow models, random walks, optimization and many other approaches. On a related but different front, recent advances in computer vision and data mining have also relied heavily on the idea of viewing a data set or an image as a graph or a network, in order to extract information about the important features of the images or more generally, the data sets [12,13].

Y. Tan, Y. Shi, and K.C. Tan (Eds.): ICSI 2010, Part II, LNCS 6146, pp. 17–26, 2010.

The dissimilarity index for each pair of nodes and the corresponding hierarchical algorithm to partition the networks are proposed in [9]. The basic idea is to associate the network with the random walker Markovian dynamics [14]. In traditional clustering literature, a function called validity index [15] is often used to evaluate the quality of clustering results, which has smaller values indicating stronger cluster structure. This can motivate us to solve the fuzzy clustering problem by an analogy to the fuzzy c-means algorithm [16] and construct an extended formulation of Xie-Beni index under this measure. Then simulated annealing strategy [17,18] is utilized to obtain the minimum value of such index, associating with a dissimilarity-index-based fuzzy c-means iterative procedure. The fuzzy clustering contains more detailed information and has more predictive power than the old way of doing network partition.

We will construct our algorithm — simulated annealing with a dissimilarity-index-based fuzzy c-means (SADIF) for fuzzy partition of networks. From the numerical performance to three model problems: the ad hoc network with 128 nodes, the karate club network and and sample network generated from Gaussian mixture model, we can see that our algorithm can efficiently and automatically determine the optimal number of clusters and identify the probabilities of each node belonging to different clusters during the cooling process.

The rest of the paper is organized as follows. In Section 2, we briefly introduce the dissimilarity index [9] which signifies to what extent two nodes would like to be in the same cluster, then proposed the extended fuzzy c-means and validity index for network partition. After reviewing the idea of simulated annealing, we describe our algorithm (SADIF) and the corresponding strategies in Section 3. In Section 4, we apply the algorithm to three representative examples mentioned before. Finally we make the conclusion in Section 5.

2 The Framework for Fuzzy Clustering of Networks

In [9], a dissimilarity index between pairs of nodes is defined, which one can measure the extent of proximity between nodes of a network. Let $G(S, E)$ be a network with n nodes and m edges, where S is the nodes set, $E = \{e(x, y)\}_{x,y \in S}$ is the weight matrix and $e(x, y)$ is the weight for the edge connecting the nodes x and y. We can relate this network to a discrete-time Markov chain with stochastic matrix $P = (p(x, y))$ whose entries are given by

$$p(x, y) = \frac{e(x, y)}{d(x)}, \qquad d(x) = \sum_{z \in S} e(x, z), \tag{1}$$

where $d(x)$ is the degree of the node x [10,11,14]. Suppose the random walker is located at node x. The mean first passage time $t(x, y)$ is the average number of steps it takes before it reaches node y for the first time, which is given by

$$t(x, y) = p(x, y) + \sum_{j=1}^{+\infty} (j + 1) \cdot \sum_{z_1, \cdots, z_j \neq y} p(x, z_1) p(z_1, z_2) \cdots p(z_j, y). \tag{2}$$

It has been shown that $t(x,y)$ is the solution of the linear equation

$$[I - B(y)] \begin{pmatrix} t(1,y) \\ \vdots \\ t(n,y) \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}, \tag{3}$$

where $B(y)$ is the matrix formed by replacing the y-th column of matrix P with a column of zeros [9]. The difference in the perspectives of nodes x and y about the network can be quantitatively measured. The dissimilarity index is defined by the following expression

$$\Lambda(x,y) = \frac{1}{n-2} \left(\sum_{z \in S, z \neq x,y} \Big(t(x,z) - t(y,z) \Big)^2 \right)^{\frac{1}{2}}. \tag{4}$$

We take a partition of S as $S = \bigcup_{k=1}^{N} S_k$ with $S_k \bigcap S_l = \emptyset$ if $k \neq l$. If two nodes x and y belong to the same cluster, then the average distance $t(x,z)$ will be quite similar to $t(y,z)$, therefore the network's two perspectives will be quite similar. Consequently, $\Lambda(x,y)$ will be small if x and y belong to the same cluster and large if they belong to different clusters.

However, this is often too restrictive for the reason that nodes at the boundary among clusters share commonalities with more than one cluster and play a role of transition in many diffusive networks. This motivates the extension to the fuzzy clustering concept where each node may belong to different clusters with nonzero probabilities. Let $\rho_k(x)$ represent the probability of the node x belonging to the k-th cluster. An extended form of fuzzy c-means is considered to address the optimization issue

$$\min_{\rho_k(x),\ m(S_k)} J_{\mathrm{DI}}(\rho, m) = \sum_{k=1}^{N} \sum_{x \in S} \rho_k^2(x) \Lambda^2(m(S_k), x), \tag{5}$$

which guarantees convergence towards a local minimum [16]. The Euler-Lagrange equation for (5) with constraints $\sum_{k=1}^{N} \rho_k(x) = 1$ is given by the following

$$\rho_k(x) = \frac{1/\Lambda^2(m(S_k), x)}{\sum_{l=1}^{N} 1/\Lambda^2(m(S_l), x)}, \quad x \in S, \quad k = 1, \cdots, N, \tag{6a}$$

$$m(S_k) = \arg\min_{x \in S_k} \frac{1}{|S_k|} \sum_{y \in S_k, y \neq x} \Lambda(x,y), \quad k = 1, \cdots, N, \tag{6b}$$

where $|S_k|$ is the number of nodes in S_k and we set $x \in S_k$ if $k = \arg\max_l \rho_l(x)$.

A well known validity index for fuzzy clustering called Xie-Beni index [15] are widely used to classify samples overlap in Euclidean space, which is based on the fuzzy c-means algorithm [16]. We extend the idea of considering both compactness and separateness to our formulation, and propose a new dissimilarity-index-based validity index for network partition as following

$$V_{\mathrm{DI}} = \frac{J_{\mathrm{DI}}}{K(m)} = \frac{\sum_{k=1}^{N} \sum_{x \in S} \rho_k^2(x) \Lambda^2(m(S_k), x)}{\min_{k \neq l} \Lambda^2(m(S_k), m(S_l))}, \tag{7}$$

where J_{DI} is the objective function constructed for the dissimilarity-index-based c-means which reflects compactness of the data set S and $K(m)$ plays the role of separation. The more separate the clusters, the larger $K(m)$ and the smaller V_{DI}. An ideal partition requires a more stable state in space $\mathbb{S} = \{S_1, \ldots, S_N\}$, which has smaller J_{DI} and larger $K_(m)$. Thus, an optimal partition can be find by solving

$$\min_N \left\{ \min_{\{S_1,\cdots,S_N\}} V_{\mathrm{DI}} \right\}. \tag{8}$$

3 The Algorithm

The simulated annealing strategy is utilized here to address (8), which is motivated by simulating the physical process of annealing solids [17]. Firstly, a solid is heated from a high temperature and then cooled slowly so that the system at any time is approximately in thermodynamic equilibrium. At equilibrium, there may be many configurations with each one corresponding to a specific energy level. The chance of accepting a change from the current configuration to a new configuration is related to the difference in energy between the two states. The simulated annealing strategy is widely used to optimization problems [18].

Let $E = V_{\mathrm{DI}}$. $E^{(n)}$and $E^{(n+1)}$ represent the current energy and new energy, respectively. $E^{(n+1)}$is always accepted if it satisfies $E^{(n+1)} < E^{(n)}$, but if $E^{(n+1)} > E^{(n)}$ the new energy level is only accepted with a probability as specified by $\exp(-\frac{1}{T}\triangle E^{(n)})$, where $\triangle E^{(n)} = E^{(n+1)} - E^{(n)}$ is the difference of energy and T is the current temperature. The initial state is generated by randomly N clusters, here $N \in [N_{\min}, N_{\max}]$, and the initial temperature T is set to a high temperature $T_{\max}$. A neighbor of the current state is produced by randomly flipping one spin, then the energy of the new state is calculated. The new state is kept if the acceptance requirement is satisfied. This process will be repeated for R times at the given temperature. A cooling rate $0 < \alpha < 1$ decreased the current temperature until reached the bound $T_{\min}$. The whole procedure of the Simulated Annealing with a Dissimilarity-Index-based Fuzzy c-means algorithm (SADIF) is summarized as follows

(1) Set parameters $T_{\max}$, $T_{\min}$, $N_{\min}$, $N_{\max}$, α and R. Choose N randomly within range $[N_{\min}, N_{\max}]$ and initialize the memberships $\{\rho_k^{(0)}\}_{k=1}^N$ randomly; Set the current temperature $T = T_{\max}$.

(2) Compute the centers $\{m(S_k^{(0)})\}_{k=1}^N$ according to (6b), then calculate the initial energy $E^{(0)}$ using the definition of V_{DI} (7); Set $n^* = 0$.

(3) For $n = 0, 1, \cdots, R$, do the following

  (3.1) Generate a set of centers $\{m(S_k^{(n)})\}_{k=1}^{N'}$ according to our proposal below and set $N = N'$;

  (3.2) Update the memberships $\{S_k^{(n+1)}\}_{k=1}^N$ and the corresponding centers $\{m(S_k^{(n+1)})\}_{k=1}^N$ according to (6a) and (6b), respectively, then calculate the new energy $E^{(n+1)}$ using (7);

(3.3) Accept or reject the new state. If $E^{(n+1)} < E^{(n)}$ or $E^{(n+1)} > E^{(n)}$ with $u \sim \mathcal{U}[0,1]$, $u < \exp\{-\frac{1}{T}\triangle E^{(n)}\}$, then accept the new solution by setting $n = n+1$; Else,reject it;

(3.4) Update the optimal state, i.e. if $E^{(n)} < E^{(n^*)}$, set $n^* = n$.

(4) Cooling temperature $T = \alpha \cdot T$. If $T < T_{\min}$, go to Step (5); Else, set $n = n^*$, repeat Step (3).

(5) Output the optimal solution $\{\rho_k^{(n^*)}\}_{k=1}^N$ and the minimum energy $E^{(n^*)}$ of the whole procedure; Classify the nodes according to the majority rule, i.e. $x \in S_k$ if $k = \arg\max_l \rho_l(x)$, gives the deterministic partition.

Our proposal to the process of generating a set of new partitions in Step (3.1) comprises three functions, which are deleting a current center, splitting a current center and remaining a current center. At each iteration, one of the three functions can be randomly chosen and the size of a cluster

$$M(S_k) = \sum_{x \in S_k} \rho_k(x), \quad k = 1, \cdots, N, \tag{9}$$

is used to select a center. Obviously, the size of a cluster is larger indicates the structure of the cluster is stronger. The three functions are described below

(i) Delete Center. The cluster with the minimal size S_d is identified using (9) and its center should be deleted from $\{m(S_k)\}_{k=1}^N$.

(ii) Split Center. The cluster with the maximal size S_s is identified using (9) and should be split into two clusters. The new center $m(S_{N+1})$ is randomly chosen in S_s but $m(S_{N+1}) \neq m(S_s)$.

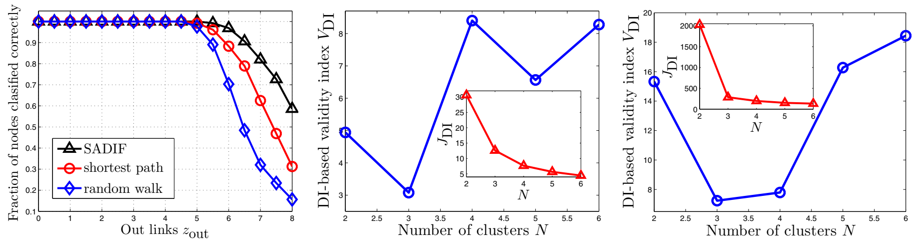
(iii) Remain Center. We remain the center set $\{m(S_k)\}_{k=1}^N$.

The number of the iterative steps depends on the initial and terminal temperature, the cooling rate and the repeating times at the given temperature. The advantages of our algorithm are the initial memberships $\{\rho_k^{(0)}\}$ can be randomly chosen and the whole annealing process does not cost so much as in the traditional cases according to the selected model parameters. The global minimum of (8) can be also obtained by searching over the all possible N using the fuzzy c-means algorithms (6). This will cost extremely much since for each fixed N, the fuzzy c-means procedure should be operated 1000 to 5000 trials due to its local minima. However, the simulated annealing strategy can avoid repeating ineffectively and lead to a high degree of efficiency and accuracy.

4 Experimental Results

4.1 Ad Hoc Network with 128 Nodes

We apply our method to the ad hoc network with 128 nodes. The ad hoc network is a typical benchmark problem considered in many papers [7,9,10,11]. Suppose we choose $n = 128$ nodes, split into four clusters containing 32 nodes each.



(a) (b) (c)

Fig. 1. (a)The fraction of nodes classified correctly of ad hoc network by SADIF comparing with the methods used in [7]. (b)V_{DI} and J_{DI} changed with N for the karate club network. The optimal V_{DI} is reached at $N = 3$ with the value $V_{\mathrm{DI}} = 3.0776$. (c)V_{DI} and J_{DI} changed with N for the 3-Gaussian mixture network. The optimal V_{DI} is reached at $N = 3$ with the value $V_{\mathrm{DI}} = 7.1404$.

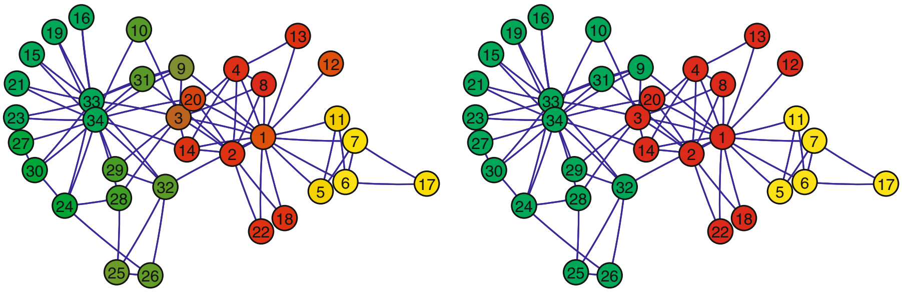
Assume pairs of nodes belonging to the same clusters are linked with probability p_{in}, and pairs belonging to different clusters with probability p_{out}. These values are chosen so that the average node degree, d, is fixed at $d = 16$. In other words p_{in} and p_{out} are related as

$$31p_{\mathrm{in}} + 96p_{\mathrm{out}} = 16. \tag{10}$$

Here we naturally choose the nodes group $S_1 = \{1 : 32\}, S_2 = \{33 : 64\}, S_3 = \{65 : 96\}, S_4 = \{97 : 128\}$. We change z_{out} from 0.5 to 8 and look into the fraction of nodes which correctly classified. The model parameters are set by $T_{\mathrm{max}} = 3$, $T_{\mathrm{min}} = 10^{-2}$, $\alpha = 0.9$ and $R = 50$. The fraction of correctly identified nodes is shown in Figure 1(a), comparing with the two methods described in [7]. It seems that our algorithm performs noticeably better than the two previous methods, especially for the more diffusive cases when z_{out} is large. This verifies the accuracy of our method, but our method gives more detailed information for each node.

4.2 The Karate Club Network

This network was constructed by Wayne Zachary after he observed social interactions between members of a karate club at an American university [19]. Soon after, a dispute arose between the clubs administrator and main teacher and the club split into two smaller clubs. It has been used in several papers to test the algorithms for finding clusters in networks [6,7,8,9,10,11]. The validity index function V_{DI} changed with N using (6) is shown in Figure 1(b). Our method is operated with the model parameters $T_{\mathrm{max}} = 3$, $T_{\mathrm{min}} = 10^{-2}$, $\alpha = 0.9$, $R = 20$ and the numerical and partitioning results are shown in Table 1 and Figure 2. We can see clearly that nodes nodes in the transition region have diffusive weights of belonging to the different clusters.



(a) (b)

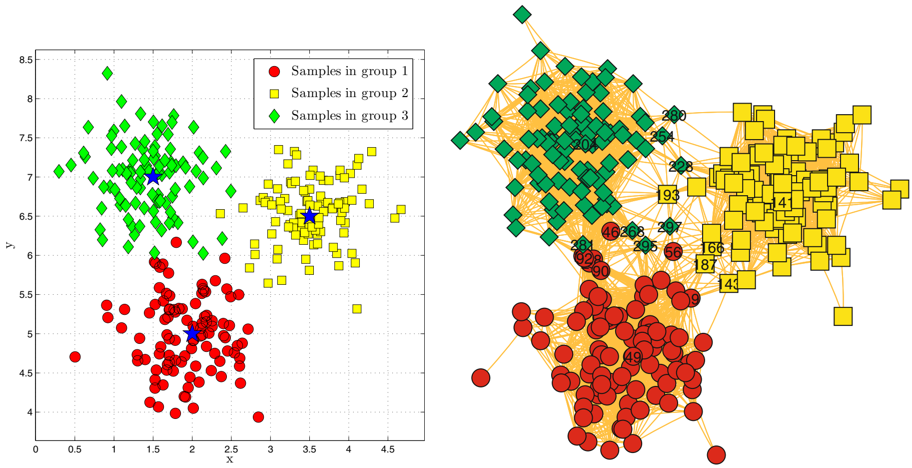
Fig. 2. (a)The fuzzy partition for the karate club network. The optimal validity index achieved is $V_{\mathrm{DI}} = 3.0776$ and corresponds to the 3 clusters represented by the weighted colors which is done as in [11]. (b)The hard partition for the karate club network obtained by the majority rule.

Table 1. The probabilities of each node belonging to different clusters of the karate club network. ρ_R, ρ_Y or ρ_G means the probability belonging to red, yellow or green colored cluster, respectively.

Nodes	1	2	3	4	5	6	7	8	9	10	11	12
ρ_R	0.7889	0.9347	0.6766	0.9528	0.1088	0	0.0166	1.0000	0.3566	0.2355	0.1405	0.7991
ρ_Y	0.1211	0.0214	0.0468	0.0173	0.8485	1.0000	0.9756	0	0.0445	0.0344	0.8044	0.1146
ρ_G	0.0900	0.0439	0.2766	0.0299	0.0427	0	0.0078	0	0.5989	0.7301	0.0551	0.0863
Nodes	13	14	15	16	17	18	19	20	21	22	23	24
ρ_R	0.9284	0.9298	0	0	0.0306	0.9285	0	0.8473	0	0.9285	0	0.1144
ρ_Y	0.0332	0.0166	0	0	0.9541	0.0316	0	0.0337	0	0.0316	0	0.0328
ρ_G	0.0384	0.0536	1.0000	1.0000	0.0153	0.0399	1.0000	0.1190	1.0000	0.0399	1.0000	0.8528
Nodes	25	26	27	28	29	30	31	32	33	34		
ρ_R	0.2543	0.2355	0.0966	0.1807	0.2035	0.1028	0.2449	0.2310	0.0129	0.0253		
ρ_Y	0.0792	0.0746	0.0280	0.0443	0.0354	0.0310	0.0353	0.0534	0.0033	0.0059		
ρ_G	0.6665	0.6899	0.8754	0.7750	0.7611	0.8662	0.7198	0.7156	0.9838	0.9688		

4.3 Sample Network Generated from Gaussian Mixture Model

To further test the validity of the algorithm, we apply it to a sample network generated from a Gaussian mixture model. This model is quite related the concept random geometric graph [20]. We generate n sample points $\{\boldsymbol{x}_i\}$ in two dimensional Euclidean space subject to a K-Gaussian mixture distribution $\sum_{k=1}^{K} q_k G\left(\boldsymbol{\mu}_k, \Sigma_k\right)$, where $\{q_k\}$ are mixture proportions satisfying $0 < q_k < 1$, $\sum_{k=1}^{K} q_k = 1$. $\boldsymbol{\mu}_k$ and Σ_k are the mean positions and covariance matrices for each component, respectively. Then we generate the network as following: if $|\boldsymbol{x}_i - \boldsymbol{x}_j| \leq dist$, we set an edge between the i-th and j-th nodes; otherwise they



(a) (b)

Fig. 3. (a)300 sample points generated from the given 3-Gaussian mixture distribution. The star symbols represent the centers of each Gaussian component. The circle, square and diamond shaped symbols represent the position of sample points in each component respectively. (b)The network generated from the sample points in Figure 3(a) with the parameter $dist = 0.7$.

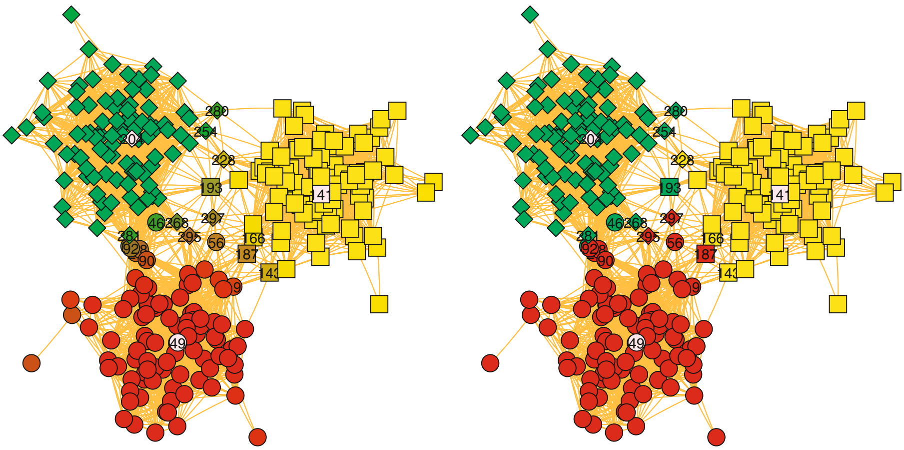
Table 2. The probabilities of the nodes which have intermediate weights belonging to different clusters for the Gaussian mixture network. For other nodes, though they have not 0-1weights, one dominate component have strength weight more that 0.85.

Nodes	19	25	29	46	56	58	90	92	143	166
ρ_R	0.8499	0.6937	0.4728	0.1906	0.5276	0.6312	0.7813	0.5170	0.2891	0.1021
ρ_Y	0.0400	0.0336	0.0389	0.0295	0.1765	0.0364	0.0280	0.0388	0.4971	0.8050
ρ_G	0.1101	0.2727	0.4883	0.7799	0.2959	0.3324	0.1907	0.4442	0.2138	0.0929
Nodes	187	193	228	254	268	280	281	295	297	
ρ_R	0.4691	0.3007	0.1939	0.1180	0.3087	0.1502	0.1714	0.6067	0.4223	
ρ_Y	0.2556	0.2260	0.5646	0.0477	0.0398	0.0689	0.0255	0.0552	0.1893	
ρ_G	0.2753	0.4733	0.2415	0.8343	0.6515	0.7809	0.8031	0.3381	0.3884	

are not connected. We take $n = 400$ and $K = 3$, then generate the sample points with the means and the covariance matrices

$$\boldsymbol{\mu}_1 = (2.0, 5.0)^T, \boldsymbol{\mu}_2 = (3.5, 6.5)^T, \boldsymbol{\mu}_3 = (1.5, 7.0)^T, \tag{11a}$$

$$\Sigma_1 = \Sigma_2 = \Sigma_3 = \begin{pmatrix} 0.18 & 0 \\ 0 & 0.18 \end{pmatrix}. \tag{11b}$$



(a) (b)

Fig. 4. (a)The fuzzy partition for the Gaussian mixture network. The optimal validity index achieved is $V_{\mathrm{DI}} = 7.1404$ and corresponds to the 3 clusters represented by the weighted colors. The positions of m ={49, 141, 204} which are colored pink has a mean L^2-error at 0.06 with respect to $\boldsymbol{\mu}$. (b)The hard partition obtained by the majority rule.

Here we pick nodes 1:100 in group 1, nodes 101:200 in group 2 and nodes 201:300 in group 3 for simplicity. With this choice, approximately $q_1 = q_2 = q_3 = 100/300$. The thresholding is chosen as $dist = 0.7$ in this example. The sample points are shown in Figure 3(a) and the corresponding network is shown in Figure 3(b). The validity index function V_{DI} changed with N using (6) is shown in Figure 1(c). Our method is operated with the model parameters $T_{\max} = 3$, $T_{\min} = 10^{-2}$, $\alpha = 0.9$, $R = 20$ and the numerical and partitioning results are shown in Table 2 and Figure 4. The results are reasonable to indicate that our algorithm can go smoothly with several hundreds of nodes.

5 Conclusions

In this paper, we have proposed an effective validity index for fuzzy clustering in networks and used the simulated annealing strategy to minimize this index associating with a dissimilarity-index-based fuzzy c-means procedure. The algorithm (SADIF) can not only identify the probabilities of each node belonging to different clusters but also determine the optimal number of clusters automatically without any prior knowledge about the network structure. Successful applications to three representative examples, including the ad hoc network, the karate club network and the sample networks generated from Gaussian mixture model, indicate that our method can always lead to a high degree of efficiency and accuracy.

Acknowledgements. This work is supported by the National Natural Science Foundation of China under Grant 10871010 and the National Basic Research Program of China under Grant 2005CB321704.

References

1. Albert, R., Barabási, A.L.: Statistical Mechanics of Complex Networks. Rev. Mod. Phys. 74(1), 47–97 (2002)
2. Newman, M.: The Structure and Function of Networks. Comput. Phys. Commun. 147(1), 40–45 (2002)
3. Newman, M., Barabási, A.L., Watts, D.J.: The Structure and Dynamics of Networks. Princeton University Press, Princeton (2005)
4. Barabási, A., Jeong, H., Neda, Z., Ravasz, E., Schubert, A., Vicsek, T.: Evolution of the Social Network of Scientific Collaborations. Physica A 311, 590–614 (2002)
5. Ravasz, E., Somera, A., Mongru, D., Oltvai, Z., Barabási, A.: Hierarchical Organization of Modularity in Metabolic Networks. Science 297(5586), 1551–1555 (2002)
6. Girvan, M., Newman, M.: Community Structure in Social and Biological Networks. Proc. Natl. Acad. Sci. USA 99(12), 7821–7826 (2002)
7. Newman, M., Girvan, M.: Finding and Evaluating Community Structure in Networks. Phys. Rev. E 69(2), 026113 (2004)
8. Newman, M.: Modularity and Community Structure in Networks. Proc. Natl. Acad. Sci. USA 103(23), 8577–8582 (2006)
9. Zhou, H.: Distance, Dissimilarity Index and Network Community Structure. Phys. Rev. E 67(6), 061901 (2003)
10. Weinan, E., Li, T., Vanden-Eijnden, E.: Optimal Partition and Effective Dynamics of Complex Networks. Proc. Natl. Acad. Sci. USA 105(23), 7907–7912 (2008)
11. Li, T., Liu, J., Weinan, E.: Probabilistic Framework for Network Partition. Phys. Rev. E 80, 26106 (2009)
12. Shi, J., Malik, J.: Normalized Cuts and Image Segmentation. IEEE Trans. Pattern Anal. Mach. Intel. 22(8), 888–905 (2000)
13. Meilă, M., Shi, J.: A Random Walks View of Spectral Segmentation. In: Proceedings of the Eighth International Workshop on Artificial Intelligence and Statistics, pp. 92–97 (2001)
14. Lovasz, L.: Random Walks on Graphs: A Survey. Combinatorics, Paul Erdos is Eighty 2, 1–46 (1993)
15. Xie, X.L., Beni, G.: A Validity Measure for Fuzzy Clustering. IEEE Tran. Pattern Anal. Mach. Intel. 13(8), 841–847 (1991)
16. Hastie, T., Tibshirani, R., Friedman, J.: The Elements of Statistical Learning: Data Mining, Inference and Prediction. Springer, New York (2001)
17. Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E.: Equation of State Calculations by Fast Computing Machines. J. Chem. Phys. 21(6), 1087 (1953)
18. Kirkpatrick, S., Gelatt, C.D., Vecchi, M.P.: Optimization by Simulated Annealing. Science 220(4598), 671–680 (1983)
19. Zachary, W.: An Information Flow Model for Conflict and Fission in Small Groups. J. Anthrop. Res. 33(4), 452–473 (1977)
20. Penrose, M.: Random Geometric Graphs. Oxford University Press, Oxford (2003)