Fuzzy Algorithm Based on Diffusion Maps for Network Partition

Jian Liu

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

Abstract. To find the best partition of a large and complex network into a small number of communities has been addressed in many different ways. The method conducted in k-means form under the framework of diffusion maps and coarse-grained random walk is implemented for graph partitioning, dimensionality reduction and data set parameterization. In this paper we extend this framework to a probabilistic setting, in which each node has a certain probability of belonging to a certain community. The algorithm (FDM) for such a fuzzy network partition is presented and tested, which can be considered as an extension of the fuzzy c-means algorithm in statistics to network partitioning. Application to three representative examples is discussed.

Keywords: Complex networks, Fuzzy community structure, Diffusion maps, K-means, Fuzzy c-means.

1 Introduction

There has been an explosive growth of interest and activity on the structure and dynamics of complex networks [1,2,3] during recent years. This is partly due to the influx of new ideas, particularly ideas from statistical mechanics, to the subject, and partly due to the emergence of interesting and challenging new examples of complex networks such as the internet and wireless communication networks. Network models have also become popular tools in social science, economics, the design of transportation and communication systems, banking systems, powergrid, etc, due to our increased capability of analyzing these models. Since these networks are typically very complex, it is of great interest to see whether they can be reduced to much simpler systems. In particular, much effort has gone into partitioning the network into a small number of communities [4,5,6,7,8,9,10,11,12], which are constructed from different viewing angles comparing different proposals in the literature.

In a previous paper [10], a k-means approach is proposed to partition the networks based on diffusion maps theory [13]. The basic idea is to associate the network with the random walker Markovian dynamics [14,15], then the diffusion distance is induced by a non-linear embedding coordinates that reflects the connectivity of the data set, and the time parameter in the Markov chain determines

D.-S. Huang et al. (Eds.): ICIC 2010, LNAI 6216, pp. 163–172, 2010.

the dimensionality reduction in the embedding. The final minimization problem is solved by an analogy to the traditional k-means algorithm [16].

The current paper extends the work [10] to a probabilistic setting. In statistical literature, a widely used generalization of k-means algorithm is the fuzzy c-means (FCM) algorithm [16]. In this framework, each node has a certain probability of belonging to a certain cluster, instead of assigning nodes to specific clusters. This idea is quite valuable since usually it is not well separated for most of networks. For the nodes lying in the transition domain between different communities, the fuzzy partition will be more acceptable. To obtain the hard clustering result, one only needs to threshold the weights. But the fuzzy partition presents more detailed information than the hard one, and it gives more reasonable explanations in some cases.

We constructed our algorithm — fuzzy algorithm based on diffusion maps (FDM) for network partition. From the numerical performance to three model problems: the ad hoc network with 128 nodes, sample networks generated from Gaussian mixture model and the karate club network, we can see that our algorithm can automatically determine the association probability of each node belonging to a certain community and lead to a high degree of efficiency and accuracy.

The rest of the paper is organized as follows. In Section 2, we briefly introduce the framework of diffusion maps and coarse-graining [10], and then extend it to a probabilistic setting. The algorithm (FDM) and corresponding computational cost are described in Section 3. In Section 4, we apply our algorithm to three examples mentioned before. The numerical results and performance are typically compared. Finally, we conclude the paper in Section 5.

2 Framework for Fuzzy Partition of Networks

The main idea of [10] is to define a system of coordinates with an explicit metric that reflects the connectivity of nodes in a given network and the construction is based on a Markov random walk on networks. 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 with entries $p_1(x, y)$ given by

$$p_1(x, y) = \frac{e(x, y)}{d(x)}, \qquad d(x) = \sum_{z\in S} e(x, z), \tag{1}$$

where $d(x)$ is the degree of the node x [14,15]. This Markov chain has stationary distribution $\mu(x) = \frac{d(x)}{\sum_{z\in S} d(z)}$ and it satisfies the detailed balance condition $\mu(x)p_1(x, y) = \mu(y)p_1(y, x)$. The diffusion distance $D_t(x, y)$ between x and y is defined as the weighted L^2 distance

$$D_t^2(x, y) = \sum_{z\in S} \frac{\big(p_t(x, z) - p_t(y, z)\big)^2}{\mu(z)}, \tag{2}$$

where the weight $\mu(z)^{-1}$ penalize discrepancies on domains of low density more than those of high density.

The transition matrix P has a set of left and right eigenvectors and a set of eigenvalues $1 = \lambda_0 \geq |\lambda_1| \geq \cdots \geq |\lambda_{n-1}| \geq 0$

$$P\varphi_i = \lambda_i \varphi_i, \quad \psi_i^T P = \lambda_i \psi_i^T, \quad i = 0, 1, \cdots, n-1. \tag{3}$$

Note that $\psi_0 = \mu$ and $\varphi_0 \equiv 1$. We also have $\psi_i(x) = \varphi_i(x)\mu(x)$. Let $q(t)$ be the largest index i such that $|\lambda_i|^t > \delta|\lambda_1|^t$ and if we introduce the diffusion map

$$\Psi_t : x \longmapsto \begin{pmatrix} \lambda_1 \varphi_1^t(x) \\ \vdots \\ \lambda_{q(t)} \varphi_{q(t)}^t(x) \end{pmatrix}, \tag{4}$$

then the diffusion distance $D_t(x, y)$ can be approximated to relative precision δ using the first $q(t)$ non-trivial eigenvectors and eigenvalues

$$D_t^2(x, y) \simeq \sum_{i=1}^{q(t)} \lambda_i^{2t} \Big(\varphi_i(x) - \varphi_i(y)\Big)^2 = \|\Psi_t(x) - \Psi_t(y)\|^2. \tag{5}$$

We take a partition of S as $S = \bigcup_{k=1}^{N} S_k$ with $S_k \bigcap S_l = \emptyset$ if $k \neq l$, then the geometric centroid $c(S_k)$ of community S_k is defined as

$$c(S_k) = \sum_{x \in S_k} \frac{\mu(x)}{\hat{\mu}(S_k)} \Psi_t(x), \quad k = 1, \cdots, N, \tag{6}$$

where $\hat{\mu}(S_k) = \sum_{x \in S_k} \mu(x)$. The optimal partition is found by minimizing the following distortion in diffusion space

$$J_h = \sum_{k=1}^{N} \sum_{x \in S_k} \mu(x) \|\Psi_t(x) - c(S_k)\|^2 = \sum_{k=1}^{N} \sum_{x \in S} \mathbf{1}_{S_k}(x) \mu(x) \|\Psi_t(x) - c(S_k)\|^2. \tag{7}$$

The procedure of k-means algorithm is considered to address this optimization issue which guarantees convergence towards a local minimum [10,16].

In the formulation given above, each node belongs to only one community after the partition. This is often too restrictive since in many diffusive networks, nodes at the boundary among communities share commonalities with more than one community and play a role of transition. This motivates the extension of the diffusion maps theory to a probabilistic setting. Here we use the terminology hard clustering since we take indicator function $\mathbf{1}_{S_k}(x)$ in (7) when the node x belongs to the k-th community. Now we extend it to the fuzzy clustering concept where each node may belong to different communities with nonzero probabilities at the same time. We will denote it as $\rho_k(x)$ with the probability of the node x belonging to the k-th community. Similar as before, given the number of the

communities N, we optimally reduce the random walker dynamics by considering the following minimization problem

$$\min_{\rho_k(x),\ c(S_k)} J = \sum_{k=1}^{N} \sum_{x \in S} \rho_k^b(x) \mu(x) \|\Psi_t(x) - c(S_k)\|^2 \tag{8}$$

subject to the constraints $\sum_{k=1}^{N} \rho_k(x) = 1, x \in S$. Like most of the traditional fuzzy c-means, $b > 1$ is a known constant. To minimize the objective function J in (8), we define

$$\hat{\mu}_k = \sum_{z \in S} \rho_k^b(z) \mu(z). \tag{9}$$

The Euler-Lagrange equation according to our derivation for the minimization problem (8) with constraints $\sum_{k=1}^{N} \rho_k(x) = 1, x \in S$ is given by the following

$$c = I_{\hat{\mu}}^{-1} \rho^b I_\mu \Psi_t, \tag{10a}$$

$$\rho = W I_{\mathbf{1} \cdot W}^{-1}, \tag{10b}$$

where $\rho^b = (\rho_k^b(x))_{k=1,\ldots,N, x \in S}$ is a $N \times n$ matrix and W is also $N \times n$ with entries

$$W_k(x) = \frac{1}{\|\Psi_t(x) - c(S_k)\|^{\frac{2}{b-1}}}. \tag{11}$$

The diagonal matrices $I_{\hat{\mu}}$, I_μ and $I_{\mathbf{1} \cdot W}$ are $N \times N$, $n \times n$ and $n \times n$ respectively, with entries

$$(I_{\hat{\mu}})_{kl} = \hat{\mu}_k \delta_{kl}, \quad k, l = 1, \ldots, N, \tag{12a}$$

$$I_\mu(x, y) = \mu(x) \delta(x, y), \quad x, y \in S, \tag{12b}$$

$$I_{\mathbf{1} \cdot W}(x, y) = \sum_{k=1}^{N} W_k(x) \delta(x, y), \quad x, y \in S, \tag{12c}$$

where δ_{kl} and $\delta(x, y)$ are both Kronecker delta symbols.

Note that each geometric centroid in the set $\{c(S_k)\}$ may not belong to the set $\{\Psi_t(x)\}_{x \in S}$ itself. This can be a problem in some applications where such combinations have no meaning. In order to obtain representatives $\{c_k\}$ of the communities that belong to the original set S, the following definition of diffusion centers should be computed

$$c_k = \arg\min_{x \in S} \|\Psi_t(x) - c(S_k)\|^2, \quad k = 1, \cdots, N. \tag{13}$$

3 The Algorithm

A strategy suggested immediately by the Euler-Lagrange equations (10) is to iterate alternatively between the equations for $\hat{p}$ and ρ. This leads to the following Fuzzy algorithm based on Diffusion Maps (FDM):

(1) Set up the initial state $\rho^{(0)}$ at random, $n = 0$;
(2) Perform the following iteration according to (10) until $\|\rho^{(n+1)} - \rho^{(n)}\| \leq TOL$:

$$c(S_k)^{(n)} = \frac{\sum_{x \in S} \rho_k^b(x)^{(n-1)} \mu(x)}{\hat{\mu}_k^{(n-1)}} \Psi_t(x), \tag{14a}$$

$$\rho_k(x)^{(n)} = \frac{\frac{1}{\|\Psi_t(x) - c(S_k)^{(n)}\|^{\frac{2}{b-1}}}}{\sum_{l=1}^{N} \frac{1}{\|\Psi_t(x) - c(S_l)^{(n)}\|^{\frac{2}{b-1}}}}, \tag{14b}$$

Here $\hat{\mu}_k^{(n)} = \sum_{z \in S} \rho_k^b(z)^{(n)} \mu(z)$ and TOL is a prescribed tolerance;
(3) The final $\rho^{(n)}$ gives the fuzzy partition for each node and $\{c_k\}$ defined in (13) gives the node play the central role in diffusion process of each community;
(4) Classifying the nodes according to the majority rule, i.e. $x \in S_m$ if $m = \arg\max_k \rho_k(x)$, gives the deterministic partition.

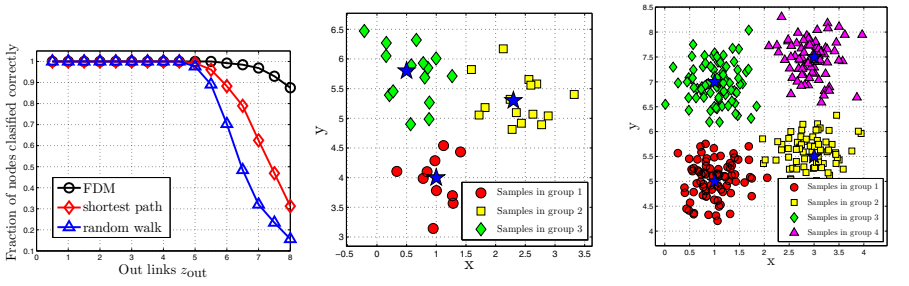
We have found that the convergence rate depends on the structure of the network. For a complex network with well-clustered community structure, the convergence is usually fast. But for a very diffusive network, convergence may be very slow. Now let us estimate the computational cost in each iteration. It is easy to see that the computational cost for Ψ_t is $O(q(t)n)$ and the computation of $\hat{\mu}$ costs $O(Nn)$, therefore the total cost in in the step of computing c is $O((N + q(t))n)$. The cost for $\|\Psi_t(x) - c(S_k)\|^2$ is $O(q(t)Nn)$. So the cost for ρ is also $O(q(t)Nn)$.

The advantages of FDM algorithm are the initial values $\{\rho_k^{(0)}\}$ can be randomly chosen, and each process does not cost much. Though the algorithm can only find the local minimum of the objective function, we still can improve it to a global optimum by operating it for several times. The fuzzy community structure contains more detailed information and has more predictive power than the old way of doing network partition.

4 Numerical Examples

4.1 Ad Hoc Network with 128 Nodes

The first example is the ad hoc network with 128 nodes. The ad hoc network is a benchmark problem used in many papers [6,7,8,11]. It has a known partition and is constructed as follows. Suppose we choose $n = 128$ nodes, split them into 4 communities with 32 nodes each. Assume that pairs of nodes belonging to the same communities are linked with probability p_{in}, and pairs belonging to different communities 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 $31p_{\text{in}} + 96p_{\text{out}} = 16$. We denote $S_1 = \{1 : 32\}, S_2 = \{33 : 64\}, S_3 = \{65 : 96\}, S_4 = \{97 : 128\}$. Typically, we define z_{out} as the average number of links a node has to nodes belonging to any other communities, i.e. $z_{\text{out}} = 96p_{\text{out}}$, and we use this quantity as a control parameter. We consider several value of



(a) (b) (c)

Fig. 1. (a)The fraction of nodes classified correctly of ad hoc network by our method compared with the methods used in [6]. (b)40 sample points generated from the given 3-Gaussian mixture distribution. (c)320 sample points generated from the given 4-Gaussian mixture distribution. The star symbols represent the centers of each Gaussian component. The circle, square, diamond and triangle shaped symbols represent the position of sample points in each component, respectively.

z_{out} between 0 and 8 and calculated the fraction of correctly identified nodes by our procedure if according to the majority rule. The time parameter is set by $t = 3$ and the tolerance is $TOL = 10^{-6}$. The result is shown in Figure 1(a). It seems that FDM performs noticeably better compared to the techniques listed in [6], especially for the more diffusive cases when z_{out} is large. This verifies the accuracy of FDM, but fuzzy method give more detailed information for each node.

4.2 Sample Network Generated from the Gaussian Mixture Model

To further test the validity of the algorithms, we apply them to a sample network generated from a Gaussian mixture model, which is quite related the concept random geometric graph [17]. We generate n sample points $\{\boldsymbol{x}_i\}$ in two dimensional Euclidean space subject to a N-Gaussian mixture distribution $\sum_{k=1}^{N} q_k G(\boldsymbol{\mu}_k, \Sigma_k)$, where $\{q_k\}$ are mixture proportions satisfying $0 < q_k < 1$, $\sum_{k=1}^{N} q_k = 1$ at first. $\boldsymbol{\mu}_k$ and Σ_k are the mean positions and covariance matrices for each component, respectively. Then we generate the network with a thresholding strategy. That is, if $|\boldsymbol{x}_i - \boldsymbol{x}_j| \leq dist$, we set an edge between the i-th and j-th nodes; otherwise they are not connected. With this strategy, the topology of the network is induced by the metric.

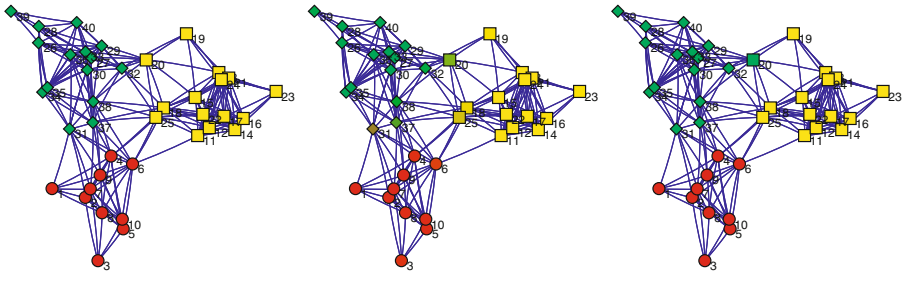
Firstly We take $n = 40$ and $N = 3$, then generate the sample points with the means and the covariance matrices as follows

$$\boldsymbol{\mu}_1 = (1.0, 4.0)^T, \boldsymbol{\mu}_2 = (2.3, 5.3)^T, \boldsymbol{\mu}_3 = (0.5, 5.8)^T, \tag{15a}$$

$$\Sigma_1 = \Sigma_2 = \Sigma_3 = \begin{pmatrix} 0.15 & 0 \\ 0 & 0.15 \end{pmatrix}. \tag{15b}$$

Table 1. The association probability of each node belonging to different communities. ρ_R, ρ_Y or ρ_G means the probability belonging to red, yellow or green colored community in Figure 2(c), respectively.

Nodes	1	2	3	4	5	6	7	8	9	10	11	12	13	14
ρ_R	0.9935	0.9829	0.9262	0.9469	0.9668	0.8628	0.9953	0.9666	0.9924	0.9668	0.0021	0.0014	0.0010	0.0058
ρ_Y	0.0025	0.0069	0.0313	0.0201	0.0138	0.0550	0.0018	0.0138	0.0028	0.0138	0.9954	0.9969	0.9976	0.9874
ρ_G	0.0041	0.0103	0.0425	0.0330	0.0194	0.0822	0.0029	0.0195	0.0048	0.0194	0.0025	0.0017	0.0013	0.0069
Nodes	15	16	17	18	19	20	21	22	23	24	25	26	27	28
ρ_R	0.0004	0.0095	0.0024	0.0751	0.0273	0.1293	0.0041	0.0014	0.0132	0.0041	0.1552	0.0245	0.0006	0.0245
ρ_Y	0.9992	0.9793	0.9948	0.8250	0.9289	0.2708	0.9909	0.9969	0.9712	0.9909	0.6277	0.0204	0.0004	0.0204
ρ_G	0.0004	0.0113	0.0028	0.0999	0.0437	0.5999	0.0050	0.0017	0.0156	0.0050	0.2171	0.9551	0.9990	0.9551
Nodes	29	30	31	32	33	34	35	36	37	38	39	40		
ρ_R	0.0063	0.0006	0.4525	0.0483	0.0029	0.0039	0.0049	0.0048	0.1864	0.0980	0.0391	0.0199		
ρ_Y	0.0054	0.0004	0.0921	0.0543	0.0024	0.0028	0.0036	0.0037	0.0869	0.0612	0.0332	0.0170		
ρ_G	0.9884	0.9990	0.4554	0.8974	0.9947	0.9933	0.9914	0.9916	0.7267	0.8408	0.9276	0.9631		



(a) (b) (c)

Fig. 2. (a)The network generated form the sample points in Figure 1(b) with the parameter $dist = 1.0$. (b)The fuzzy community structure obtained by the weighted color average in [12]. (c)Partition the network with the majority rule.

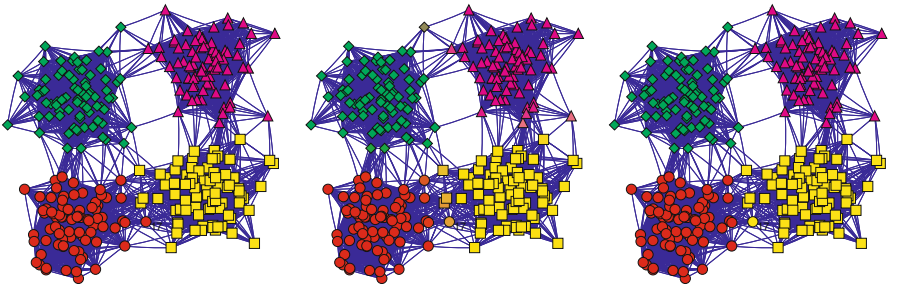
Here we pick nodes 1:10 in group 1, nodes 11:25 in group 2 and nodes 26:40 in group 3 for simplicity. With this choice, approximately $q_1 = 10/40, q_2 = q_3 = 15/40$. We take $dist = 1.0$ in this example. The sample points are shown in Figure 1(b) and the corresponding network is shown in Figure 2(a). We operate our method with 1000 trials by setting $t = 3$ and $TOL = 10^{-6}$. The association probabilities ρ are listed in Table 1. We can see that FDM algorithm achieve reasonable result that fits the intuition from the network topology visualization. The weights$\{\rho_k(x)\}$ are shown in Figure 2(b), which shows more clearly the transition between different communities [12]. Particularly the nodes {6,18,20,25,31,32,37,38} show clearly transitional behavior. If we further partition by the majority rule, we obtain Figure 2(c).

Secondly we take $n = 320$ and $N = 4$, where nodes 1:80 are in group 1, nodes 81:160 in group 2, nodes 161:240 in group 3 and nodes 241:320 in group 4. This means approximately $q_1 = q_2 = q_3 = q_4 = 80/320$. The other model parameters are chosen as

$$\begin{aligned} \boldsymbol{\mu}_1 &= (1.0, 5.0)^T, \boldsymbol{\mu}_2 = (3.0, 5.5)^T, \\ \boldsymbol{\mu}_3 &= (1.0, 7.0)^T, \boldsymbol{\mu}_4 = (3.0, 7.5)^T, \end{aligned} \tag{16a}$$

$$\Sigma_1 = \Sigma_2 = \Sigma_3 = \Sigma_4 = \begin{pmatrix} 0.15 & 0 \\ 0 & 0.15 \end{pmatrix}. \tag{16b}$$

Here we take $dist = 0.7$. The sample points are shown in Figure 1(c) and the corresponding network is shown in Figure 3(a). The fuzzy and hard partitioning results are shown in Figure 3(b) and Figure 3(c). The partition obtained by the majority rule coincide with the original sample points except node 17.



(a) (b) (c)

Fig. 3. (a)The network generated form the sample points in Figure 1(c) with the parameter $dist = 0.7$. (b)The fuzzy community structure obtained by the weighted color average in [12]. (c)Partition the network with the majority rule.

4.3 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 [18]. Soon after, a dispute arose between the club administrator and main teacher and the club split into two smaller clubs. It has been used in several papers to test the

Table 2. The association probability of each node belonging to different communities for the karate club network. ρ_R or ρ_Y means the probability belonging to red or yellow colored community in Figure 4(c), respectively.

Nodes	1	2	3	4	5	6	7	8	9	10	11	12
ρ_R	0.1766	0.2510	0.3330	0.1942	0.3935	0.4356	0.4356	0.2095	0.5455	0.5318	0.3935	0.1962
ρ_Y	0.8234	0.7490	0.6670	0.8058	0.6065	0.5644	0.5644	0.7905	0.4545	0.4682	0.6065	0.8038
Nodes	13	14	15	16	17	18	19	20	21	22	23	24
ρ_R	0.1864	0.2426	0.6029	0.6029	0.4674	0.2227	0.6029	0.3191	0.6029	0.2227	0.6029	0.6054
ρ_Y	0.8136	0.7574	0.3971	0.3971	0.5326	0.7773	0.3971	0.6809	0.3971	0.7773	0.3971	0.3946
Nodes	25	26	27	28	29	30	31	32	33	34		
ρ_R	0.5329	0.5472	0.7456	0.5569	0.6391	0.7353	0.5517	0.5734	0.7270	0.7287		
ρ_Y	0.4671	0.4528	0.2544	0.4431	0.3609	0.2647	0.4483	0.4266	0.2730	0.2713		

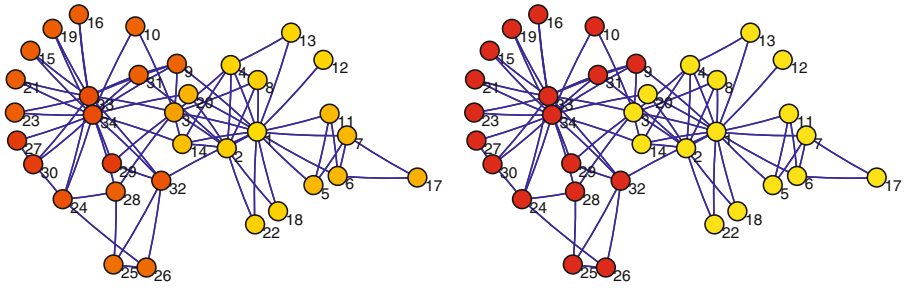


Fig. 4. (a)The fuzzy colored plot of ρ_R and ρ_Y for each node in karate club network. The transition nodes or intermediate nodes are clearly shown. (b)The partition obtained by the majority rule. The result is the same as Zachary's [18].

algorithms for finding community structure in networks [6,7,11,12]. By operating FDM with $t = 3$ and $TOL = 10^{-6}$ to this example, we obtain the mean iterative steps is 25 and the minimal value of the objective function $J_{\min} = 0.009$ during 1000 trials. The numerical result is presented in Table 2. It shows that the nodes $\{3, 9, 10, 14, 20, 31, 32\}$ have more diffusive probabilities and they play the role of transition between the red and yellow colored groups. We can visualize the data ρ more transparently with the the color vector average for each node shown in Figure 4(a). If we classify the nodes according to the majority rule, we obtain the same partition as Zachary's [18], which is shown in Figure 4(b).

5 Conclusions

We extend the fuzzy clustering in statistics to the network partitioning problem in this paper. It is a generalization of the previous k-means algorithm based on diffusion maps of a random walker Markovian dynamics on the network [10]. The hard clustering concept, a node belongs to only one community, is extended to the fuzzy clustering concept where each node may belong to different communities with nonzero probabilities. This is extremely meaningful in many diffusive cases, nodes at the boundary among communities share commonalities with more than one community and play a role of transition, and such probabilities can give people more detailed information. We have proposed the fuzzy algorithm based on diffusion maps (FDM), which is derived to search for the local minimum of the objective function (8) under the fuzzy clustering framework. The algorithm can be considered as a transform of fuzzy c-means for network partition. Moreover, it succeeds in three examples, including the ad hoc network, sample networks generated from Gaussian mixture model and the karate club network.

Acknowledgements. This work is supported by the 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., Barabasi, A.L., Watts, D.: The Structure and Dynamics of Networks. Princeton University Press, Princeton (2005)
3. National Research Council: Network Science. National Academy of Sciences, Washington DC (2005)
4. Shi, J., Malik, J.: Normalized Cuts and Image Segmentation. IEEE Trans. Pattern Anal. Mach. Intel. 22(8), 888–905 (2000)
5. 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)
6. Newman, M., Girvan, M.: Finding and Evaluating Community Structure in Networks. Phys. Rev. E 69(2), 26113 (2004)
7. Newman, M.: Detecting Community Structure in Networks. Eur. Phys. J. B 38(2), 321–330 (2004)
8. Danon, L., Diaz-Guilera, A., Duch, J., Arenas, A.: Comparing Community Structure Identification. J. Stat. Mech. 9, P09008 (2005)
9. Newman, M.: Modularity and Community Structure in Networks. Proc. Natl. Acad. Sci. USA 103(23), 8577–8582 (2006)
10. Lafon, S., Lee, A.: Diffusion Maps and Coarse-Graining: A Unified Framework for Dimensionality Reduction, Graph Partitioning, and Data Set Parameterization. IEEE Trans. Pattern. Anal. Mach. Intel. 28, 1393–1403 (2006)
11. 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)
12. Li, T., Liu, J., Weinan, E.: Probabilistic Framework for Network Partition. Phys. Rev. E 80, 26106 (2009)
13. Coifman, R., Lafon, S.: Diffusion Maps. Applied and Computational Harmonic Analysis 21(1), 5–30 (2006)
14. Lovasz, L.: Random Walks on Graphs: A Survey. Combinatorics, Paul Erdös is Eighty 2, 1–46 (1993)
15. Chung, F.: Spectral Graph Theory. American Mathematical Society, Rhode Island (1997)
16. Hastie, T., Tibshirani, R., Friedman, J.: The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer, New York (2001)
17. Penrose, M.: Random Geometric Graphs. Oxford University Press, Oxford (2003)
18. Zachary, W.: An Information Flow Model for Conflict and Fission in Small Groups. J. Anthrop. Res. 33(4), 452–473 (1977)