



Research Article

ISSN : 0975-7384  
CODEN(USA) : JCPRC5

## Community detection model based on incremental EM clustering method

Qiu Li-qing, Liang Yong-quan and Chen Zhuo-yan

College of Information Science and Technology, Shandong University of Science and Technology,  
Qingdao, Shandong, China

---

### ABSTRACT

Networks are widely used in a variety of different fields and attract more and more researchers. Community detection, one of the research hotspots, can identify salient structure and relations among individuals from the networks. Many different solutions have been put forward to detect communities. EM as a model on statistical inference methods has received more attention because of its simple and efficient structure. Unlike many other statistical inference methods, no extra information is assumed except for the network itself and the number of groups for the EM approach. However, practical usefulness of the EM method is often limited by computational inefficiency. The EM method makes a pass through all of the available data in every iteration. Thus, if the size of the networks is large, every iteration can be computationally intensive. Therefore we put forward an incremental EM method-IEM for community detection. IEM uses the machinery of probabilistic mixture models and the incremental EM algorithm to generalize a feasible model fit the observed network without prior knowledge except the networks and the number of groups. Using only a subset rather than the entire networks allows for significant computational improvements since many fewer data points need to be evaluated in every iteration. We also argue that one can select the subsets intelligently by appealing to EM's highly-appreciated likelihood judgment condition and increment factor. We perform some experimental studies, on several datasets, to demonstrate that our IEM can detect communities correctly and prove to be efficient.

**Key words:** Community Detection, Expectation Maximization, Incremental Expectation Maximization

---

### INTRODUCTION

As a new emerging discipline, research on networks attracts researchers from a variety of different fields. In fact, studies that can qualitatively and quantitatively characterize networks will help to unveil the general laws regulating different real systems modeled by networks, and therefore will be relevant in a number of disciplines (biology, social sciences, et al). Community structure is one of the crucial structural characteristics of networks; therefore, accurately analyzing their community structure represents a very relevant topic [1-6].

Communities are groups of nodes with a high level of group inter-connection [1]. They can be seen as relative isolated subgroups with few contacts with the rest of the network. Community detection can identify salient structure and relations among individuals from the network. Researchers put forward many different methods, which are mainly used to detect the groups with dense connections within groups but sparser connections between them. To detect more latent structures in reality networks, various models on statistical inference have been proposed recently, which are on sound theoretical principles and have better performances identifying structures, and have become the state-of-the-art models [7-10]. The models 'aim is to define a generative process to fit the observed network, and transfer the community detecting problem to Bayesian inference or Maximum Likelihood methods [11-14]. The drawback, shared with many other methods, is that structure detection usually implies computational expensive exploration of the solutions maximizing the posterior probability of the likelihood. More recently, a maximum likelihood method that considers model clustering as missing information and deals with it using an iterative

Expectation Maximization (EM) method has been introduced by Newman and Leiche [2]. The EM method is a simple algorithm that is capable of detecting a broad range of structural signatures in networks, including conventional community structure, bipartite or disassortative structure, fuzzy or overlapping classifications, and many mixed or hybrid structural forms that have not been considered explicitly in the past.

Due to the simple structure of the EM method, there has been a growing body of work on the analysis of the EM algorithm [3-5]. Many improvements have been put forward to better the EM method since then. However, a common weakness in these studies, as we will discuss in detail in related work, is that the EM method will be low efficient when the networks are large-scaled. In fact, the EM method may make sense when the networks are small-scaled or medium-scaled. On the contrary, more often than not, real-world networks are large-scaled. Under such scenarios, if an algorithm like iterative EM method evaluates all samples at each step, it may result in high complicity and low efficiency. Therefore, we argue that a more appropriate approach is to improve the EM method in order to reduce samples at each step. Consequently, we propose an incremental EM algorithm on the sample subset that converges to optimal solutions using the proposed formulations. We prove the correctness and convergence of our algorithm and show that this algorithm has low time complexity when the data of the networks is large-scaled.

The rest of the paper is organized as follows: in Section 2 we discuss related work and the EM method is formally introduced in Section 3. Next, we describe in Section 4 our generalization of incremental EM method of community detection. In Section 5, we provide experimental studies. Finally in Section 6, we give the conclusion and future directions.

## RELATED WORKS

Community structure has been extensively studied in various research areas such as social network analysis, Web community analysis, computer vision, et al. In network analysis, an important research topic is to identify cohesive subgroups of individuals within a network where cohesive subgroups are defined as "community detection". Recently there exists a growing body of literature on community detection. Many approaches, such as clique-based, degree-based, and matrix-perturbation-based, have been proposed to extract cohesive subgroups from network. The approach of community detection can be characterized as heuristic measure methods and statistical inference methods according to the basis of object function. Heuristic measure methods such as modularity maximization [6] and extreme optimization [7] use a heuristic metric to measure community structure and lack of rigorous theoretical basis. Statistical inference methods such as planted partition model [8] and the EM method [2] can identify the structure of structural equivalence and regular equivalence, and classify the vertices of the networks using the observed networks fit by a generative process. Statistical inference methods have perfect theoretical basis which is different from heuristic measure methods, and have become the state-of-the-art methods. Statistical inference methods have the advantage that, depending on the statistical model used, they can be very general detecting both structural equivalent and regular equivalent set of individuals. Consequently, more literatures have been proposed on statistical inference methods. The EM approach as a model on statistical inference methods has received more attention because of its simple and efficient structure. Unlike many other statistical inference methods, no extra information is assumed except for the network itself and the number of groups for the EM approach. Contrast to traditional community detection methods, the EM approach is capable of detecting disassortative structure as well as overlapping classifications.

There are some recent studies on the EM method for community detection. The EM approach to community detection is first introduced by Newman et al [2]. We will denote it by the acronym NL-EM from now on. They use the machinery of probabilistic mixture models and the EM algorithm to generalize a feasible model fit the observed network without prior knowledge except the networks and the number of groups. They also give a number of examples demonstrating how the method can be used to shed light on the properties of real-world networks. In their model, their parameter definition implies that the classification must be such that each class has at least one member with non-zero out-degree. The constraint forces NL-EM to classify a simple bi-partite graph. Based on the idea, Ramasco et al. [3] generalize an extension of NL-EM, in which they extend the parameter  $\theta$ . The examples show both numerically and analytically that the new generalized EM method is able to recover the process that led to the generation of content-based networks. Muggan et al. [4] use NL-EM to yield a stability analysis the grouping that quantifies the extent to which each node influences its neighbor group membership. All these studies, however, have a common weak point-the EM method is usually low efficient and high complexity when the networks are large-scaled. That is, when the EM method is used to detect communities in the networks, it evaluates all samples in every iteration which may result in low convergence rate and bad clustering effect. In contrast, in our proposed method, we study an incremental EM method on the sample subset instead of whole samples which prove higher efficient.

**NL-EM METHOD**

NL-EM is capable of detecting networks structure relying on following basic assumptions: (1) The actual connectivity of the networks is related to a priori unknown grouping of the individuals; (2) The presence or absence of a link is independent from the other links of the networks.

We begin with a quick summary of NL-EM as applied to graphs. Given a graph  $G$  of  $N$  nodes and an adjacency matrix  $A_{ij}$ , NL-EM method searches for a partition of the nodes into  $K$  groups such that a certain log-likelihood function for the graph is maximized. Henceforth we will refer to the groups into which NL-EM divides the nodes, as classes. There are three variables as follows in NL-EM:  $\pi_r$ , the probability that a randomly selected node is in group  $r$ ;  $\theta_{rj}$ , the probability that an edge leaving group  $r$  connects to a certain node  $j$ ;  $q_{ir}$ , the probability that node  $i$  is assigned to group  $r$ . The parameters  $\pi_r$  and  $\theta_{rj}$  satisfy the normalization conditions:

$$\sum_{r=1}^K \pi_r = 1, \quad \sum_{i=1}^N \theta_{ri} = 1 \quad (1)$$

Assuming that the parameters  $\pi$  and  $\theta$  are given, the probability  $\Pr(A, g | \pi, \theta)$  of realizing the given graph under a node classification  $g$ , such that  $g_i$  is the group that node  $i$  has been assigned to, can be written as:

$$\Pr(A, g | \pi, \theta) = \prod_i \pi_{g_i} \left[ \prod_j \theta_{g_i, j}^{A_{ij}} \right], \quad (2)$$

$\Pr(A, g | \pi, \theta)$  is the likelihood to be maximized, but it turns out to be more convenient to consider its logarithm instead:

$$L(\pi, \theta) = \sum_i \left[ \ln \pi_{g_i} + \sum_j A_{ij} \ln \theta_{g_i, j} \right]. \quad (3)$$

Treating a prior unknown group assignment  $g_i$  of the nodes as statistical missing information, one considers the averaged log-likelihood constructed as:

$$\bar{L}(\pi, \theta) = \sum_{ir} q_{ir} \left[ \ln \pi_r + \sum_j A_{ij} \ln \theta_{rj} \right]. \quad (4)$$

The final results are

$$\pi_r = \frac{1}{N} \sum_i q_{ir}, \quad (5)$$

$$\theta_{rj} = \frac{\sum_i A_{ij} q_{ir}}{\sum_i k_i q_{ir}}, \quad (6)$$

Where  $k_i$  is the total degree of node  $i$ . The still unknown probabilities  $q_{ir}$  are then determined a posterior by noting that:

$$q_{ir} = \Pr(g_i = r | A, \pi, \theta) = \frac{\Pr(A, g_i = r | \pi, \theta)}{\Pr(A | \pi, \theta)}, \quad (7)$$

From which one obtains:

$$q_{ir} = \frac{\pi_r \prod_j \theta_{rj}^{A_{ij}}}{\sum_s \pi_s \prod_j \theta_{rj}^{A_{ij}}} \quad (8)$$

Equation (5), (6), and (8) form a set of self consistent equations for  $\pi_r$ ,  $\theta_{rj}$  and  $q_{ir}$  that any extreme of the expected log-likelihood must satisfy.

Thus, given a graph  $G$ , the EM algorithm consists of picking a number of classes  $K$  into which the nodes are to be classified and searching for solutions of Equation (5), (6), and (8). These equations are derived by Newman et al.

[2]. They also show that when applied to diverse type of networks the resulting,  $\theta_{rj}$  and  $q_{ir}$  yield useful information about the internal structure of the network. Note that only a minimal amount of a priori information is supplied: the number of classes  $K$  and the networks.

### INCREMENTAL EM METHOD

Despite the EM method's wide-spread popularity, practical usefulness of the EM method is often limited by computational inefficiency. The EM method makes a pass through all of the available data in every iteration. Thus, if the size of the networks is large, every iteration can be computationally intensive. We introduce an incremental EM algorithm for fast computation based on random sub-sampling which is denoted by the acronym IEM from now on. Using only a subset rather than the entire networks allows for significant computational improvements since many fewer data points need to be evaluated in every iteration. We also argue that one can select the subsets intelligently by appealing to EM's highly-appreciated likelihood judgment condition and increment factor.

Given a graph  $G$  of  $N$  nodes, we first select  $M$  ( $M \subset N$ ) nodes as the initial sample subset, and then we will train the initial subset by using NL-EM method. After the training, we will add  $d$  ( $d \subset N - M$ ) nodes from the remaining samples to the initial subsets, and then we will train the new formed subsets. The similar iterative operation is repeated until the subset is identical to the entire samples. The quantities in our theory thus fall into three classes: (1) How to define parameter  $M$ ? In other words, how many nodes should be first chosen as the initial subsets? (2) How to define parameter  $d$ ? That is to say, how many nodes should be complemented after last training? (3) When will  $d$  nodes be added to the subset? Namely what conditions should be satisfied when the subset changes? We will give some reasonable solutions as followed

#### *The definition of parameter $M$*

Parameter  $M$  means the number of nodes in the initial subset. The initial subset selection is an important part of IEM which has a great influence on the results. Our goal is to select some nodes as the initial subset which is most representative of the entire data, and therefore the selected subset can well describe the global features.

There is a popular view in network analysis that the important nodes are most representative of the entire networks. Consequently we will select the important nodes of the whole networks. Centrality analysis provides answers with measures that define the importance of nodes. There are many classical and commonly methods used ones [9]: degree centrality, closeness centrality, and betweenness centrality. These centrality measures capture the importance of nodes in different perspectives. With large-scale networks, the computation of centrality measures can be expensive except for degree centrality. We define  $n$  to be the number of the nodes and  $m$  to be the number of edges between nodes. Then we can get time complexity and space complexity about the centrality measures. Closeness centrality, for instance, involves the computation of all the pairwise shortest paths, with time complexity of  $O(n^2)$  and space complexity of  $O(n^3)$  with the Floyd-Warshall algorithm [10] or  $O(n^2 \log n + nm)$  time complexity with Johnson's algorithm [11]. The betweenness centrality requires  $O(nm)$  computational time following [12]. For large-scale networks, efficient computation of centrality measures is critical and requires further research.

We propose a new method of measuring the centrality which is a compromise between complexity and efficiency. Now we study degree centrality which is the simplest measures. For degree centrality, the importance of a node is determined by the number of nodes adjacent to it. The larger the degree of one node, the more important the node is. The degree centrality of node  $v$  is defined as:

$$C_D(v) = d_i / (n - 1) \quad (9)$$

where  $d_i$  is defined as the number of nodes adjacent to  $v$ , and  $n$  is defined as the number of nodes in the network.

However, the measure is not comprehensive enough, i.e., some important nodes (i.e., bridge contacts connect with merely two edges) don't have high degree centrality. Based on the idea, we argue that the importance of one node is determined by its connection model as well as its role in the networks. Accordingly we consider two factors, namely the connection model of the node and its role in the network. The connection model of one node can be described by its degree centrality, and the role of one node can be described by its cohesion centrality.

Definition1: The connectivity of node  $v$  is defined as the number of the edges between  $v$  and the nodes directly connected with  $v$ .

The connectivity of a node measures how close it is to the nodes which are directly connected with it, and reflects the local connection property of the node. Obviously, the span of connectivity is between 0 and  $C_D(v)(C_D(v)-1)/2$ .

Definition 2: The cohesion centrality of node  $v$  is defined as follows:

$$C_c(v) = \frac{C_D(v)(C_D(v)-1)}{2c_i} \quad (10)$$

Where  $C_D(v)$  is the degree centrality of node  $v$ , and  $c_i$  is the connectivity of node  $v$ .

According to the relations between the nodes and the edges in the network, the value of  $C_c(v)$  satisfies the conditions:

$$C_c(v) \geq 1 \quad (11)$$

We find that the larger the connectivity of one node, the less important the node is. This is because the deletion of the node with larger connectivity will make less affection on the network. Thus according to equation (10), the more importance one node is, the larger the cohesion centrality of the node is. Therefore, the cohesion centrality is the positive evaluation index of the node.

To integrate the two factors (i.e., connection model of one node and its role in networks), an importance function is introduced to measure the importance of the node, where the importance consists of two parts-a degree centrality and a cohesion centrality:

$$I(v) = \alpha \cdot C_D(v) + (1 - \alpha) \cdot C_c(v) \quad (12)$$

where  $\alpha$  satisfies  $0 \leq \alpha \leq 1$ .

In this importance function, the degree centrality  $C_D(v)$  measures the connection model of node  $v$ , and the cohesion centrality  $C_c(v)$  measures the role of node  $v$ . The parameter  $\alpha$  is set by the user to control the level of emphasis on each part of the total importance function.

Thus according to equation (12) we can select important nodes with high value as the initial subset.

### **The definition of parameter $d$**

Parameter  $d$  means the number of nodes which will be complemented to the subset in every iteration. The definition of parameter  $d$  is another crucial question in IEM. The parameter should make the subset fit the real model as much as possible. Here we propose the concept of incremental factor to describe parameter  $d$  based on

information entropy.

According to information theory, the entropy measures the uncertainty of the system. The larger the entropy is, the more uncertain the system is. If the density function values of every node in the subset are approximately equal, the uncertainty of the distribution for the entire data is largest (i.e., the subset has maximum entropy). Conversely, if the density function values of every node are very asymmetric, the subset has minimal uncertainty. Therefore, we introduce the concept of density entropy to measure incremental factor.

Definition 3: Given nodes set  $D = \{x_1, x_2, \dots, x_N\}$  which has  $N$  nodes, the density function value of every node is  $f(x_i)$ ,  $i = 1, 2, \dots, N$ , and  $\delta$  is the sample variance, then the density entropy is defined as follows:

$$DenEn(\delta) = - \sum_{i=1}^N \frac{f(x_i)}{Sum} \ln \frac{f(x_i)}{Sum} \quad (13)$$

where  $Sum$  is the normalized factor defined as follows:

$$Sum = \sum_{i=1}^N f(x_i) \quad (14)$$

The density entropy has two properties:

Property 1:  $0 \leq DenEn(\delta) \leq \ln(N)$ ;

Property 2:  $DenEn(\delta) = \ln(N)$  when and only when  $f(x_1) = f(x_2) = \dots = f(x_N)$ , therefore  $\lim_{\delta \rightarrow 0} DenEn(\delta) = \ln(N) = \max_{\delta} (DenEn(\delta))$ .

From Property 2, when  $DenEn(\delta) = \ln(N)$ , the nodes in the subset are consistent with the real distribution which is ideal case. With the increase of  $\delta$ , the value of  $DenEn(\delta)$  decreases which will reach a minimum subsequently, and then the value of  $DenEn(\delta)$  will become larger which will reach a maximum  $\ln(N)$  when  $\delta \rightarrow 0$  and  $f(x_1) = f(x_2) = \dots = f(x_N)$ .

The change of the sample variance  $\delta$  is similar to the density entropy, and we take into account the middle value of  $DenEn(\delta)$ . Consequently, we propose incremental factor  $\beta$  as:

$$\beta = DenEn(\delta) / 2 = \ln(N) / 2 \quad (15)$$

According to equation (15), the parameter  $d$  can be described as follows:

$$d = N / \beta \quad (16)$$

Once parameter  $d$  is determined, the iterative process of IEM can be carried out as follows: when the samples in the subset fit the real networks,  $d$  nodes are added to the subset, and then new fitting process goes on. The incremental process will end until the subset is equal to the entire data. In the process, the subset gradually approaches the entire data.

It is worth mentioning that the complement nodes in every iteration are selected from the entire nodes. Hence the number of nodes in subset turns out to be:

$$M = M + d = M + N / \beta \quad (17)$$

#### **The conditions when the subsets changes**

From Section 3, we can see that the EM method is an iterative procedure for Maximizing  $\bar{L}(\pi, \theta)$  which we

describe in Equation (4). Assume that after the  $n^{\text{th}}$  iteration the current estimate for  $\pi$  and  $\theta$  is given by  $\pi_n$  and  $\theta_n$ . Since the objective is maximize  $\bar{L}(\pi, \theta)$ , we wish to computer an updated estimate  $\pi$  and  $\theta$  such that:

$$\bar{L}(\pi, \theta) > \bar{L}(\pi_n, \theta_n) \quad (18)$$

Equation (18) means the iteration condition of the EM method, namely if the update likelihood is not more than the current likelihood then the iteration will end.

Inspired by Equation (18), we propose the iteration condition of IEM. We define  $\bar{L}_t(\pi, \theta)$  as the maximum likelihood after  $t^{\text{th}}$  iteration and  $\bar{L}_{(t+1)}(\pi, \theta)$  as the maximum likelihood after  $(t+1)^{\text{th}}$  iteration. The iteration condition can be defined as:

$$\bar{L}_{(t+1)}(\pi, \theta) > \bar{L}_t(\pi, \theta) \quad (19)$$

Equivalently we want to maximize the difference:

$$D = \bar{L}_{(t+1)}(\pi, \theta) - \bar{L}_t(\pi, \theta) \quad (20)$$

Assume  $\gamma$  is a positive number which is small enough, then if  $D > \gamma$ , we argue that the current estimate is undesirable and the iteration should go on. If  $D \leq \gamma$ , then the subset in  $t^{\text{th}}$  iteration is fit to the model of the real data, and the new samples should be complemented to carry out the next iteration. When the subset is equal to the entire data, the terminate condition is consistent to the EM method.

### *The determining of $\alpha$*

How to determine the  $\alpha$  in Equation (12) is a challenging issue. When the ground truth is available, standard validation procedures can be used to select an optimal  $\alpha$ . However, in many cases there is no ground truth and the community detection performance depends on the user's subjective performance. In the respect, through the parameter  $\alpha$ , our IEM provide the user a mechanism to push the community detection results toward his or her preferred outcomes. The problems of whether a "correct"  $\alpha$  exists and how to automatically find the best  $\alpha$  when there is no ground truth are beyond the scope of this paper. To simply the experiments, we will set  $\alpha$  as 0.5 in the following example applications.

### **EXAMPLE APPLICATIONS**

In this section, we use several synthetic dataset to study the performance of our IEM from different aspects. In section 5.1 we will first verify the correctness of IEM, and then we will compare our IEM with baseline algorithm-EM in section 5.2.

#### *First Example Application*

We start with the first synthetic dataset, which is a static network, to illustrate some good properties of our IEM. This dataset is first studied by White and Smyth[13] and is shown in Figure 1(a). The network contains 15 nodes which roughly form 3 communities-C1, C2, and C3-where edges tend to occur between nodes in the same community.

We first apply our algorithm to the network with various community numbers  $m$  and the resulting  $Q$  values are plotted in Figure1 (b).  $Q$  values can be interpreted as modularity values which is a measure of the deviation between the observed edge-cluster probabilities and what one would predict under an independence model. Newman etc. [14] show that larger  $Q$  values are correlated with better graph clustering. In Figure 1(b) we also show the modularity values  $Q'$  that are reported by White and Smyth. As can be seen from Figure1 (b), both  $Q$  and  $Q'$  show distinct peaks when  $m = 3$ , which corresponds to the correct community number. Our IEM algorithm gets higher modularity values which indicate that IEM can classify the network better.

Next, after our IEM algorithm correctly partitions the nodes into three communities, we illustrate the soft community membership by studying two communities-C2 and C3. In Figure 1(b) we use triangle shape to represent the nodes in C2, and circle shape to represent the nodes in C3. But we use different gray levels to indicate their community membership—we use white color to illustrate the level that a node belongs to C2 and dark color to show the level that a node belongs to C3. As can be seen, while the nodes which are white or black have very clear community memberships, the nodes which are on the boundary between C2 and C3, have rather fuzzy membership. The shallower the nodes are, the more likely the nodes belong to C2, Centrally, the deeper the nodes are, the more likely the nodes belong to C3. In other words, our IEM algorithm is capable of assigning meaningful soft membership to a node to indicate to which level the node belongs to a certain community.

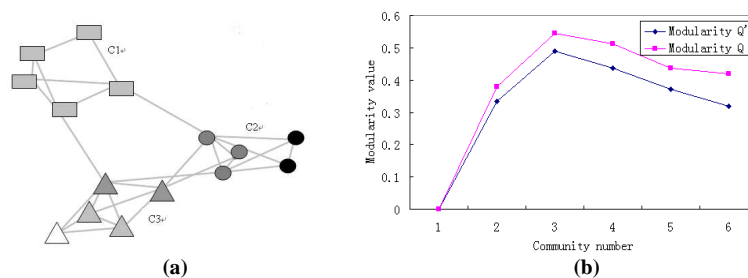


Figure1. First example application: (a) applications of IEM method, (b) Modularity value  $Q$  and Modularity value  $Q'$  under different community numbers

### Second Example Application

We secondly apply our IEM algorithm to a small network—"Karate club" network [14]. The network contains 34 nodes which roughly form 2 communities-C1 and C2. The network is of particular interest because the club split in two during the course of Zachary's observations as a result of an internal dispute and Zachary recorded the membership of the two factions after the split.

Figure 2 shows the result of our IEM algorithm which the number of clusters is set to 2. We use different gray levels to indicate their community membership as first example application. In Figure 2 we use circle shape and rectangle shape to represent the nodes in C1 and C2 respectively. But we use different gray levels to indicate their community membership—we use white color to illustrate the level that a node belongs to C1 and dark color to show the level that a node belongs to C2. As can be seen, node 9, 3, 14, 20 are on the boundary between C1 and C2, which have rather fuzzy membership.

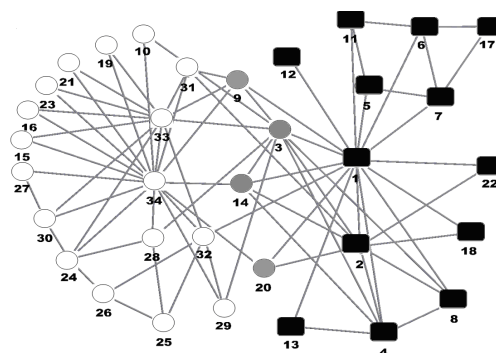


Figure2. Second example application: applications of IEM method

Next, after our IEM algorithm correctly partitions the nodes into two communities, we compare our IEM with baseline algorithm NL-EM. The compared result is shown as Table 1. As we can see from Table 1, under the same computing environment the time of IEM need only 0.471 second, which is much less than it of NL-EM; the iterations of IEM is only 35, which is greatly less than it of NL-EM. From the table we have the following observations. On the dataset, among the two algorithms (NL-EM and IEM), IEM outperforms NL-EM. In other words, our IEM can reach the neighborhood faster than NL-EM, and is high efficient because of fast convergence rate.

Table 1. The comparison between NL-EM and IEM

Dataset	Parameter	NL-EM	IEM
Karate club	Time(s)	0.782	0.471
	Iterations	87	35
	Likelihood Estimate	-6.321	-6.435



## CONCLUSION

Community detection is a challenging research problem with broad applications. In this paper we have described an incremental EM method-IEM for community detection. IEM uses the machinery of probabilistic mixture models and the incremental EM algorithm to generalize a feasible model fit the observed network without prior knowledge except the networks and the number of groups. The method is more efficient than previous NL-EM, making use of a new incremental approach which is more close to the optimal solutions. We use only a subset rather than the entire networks allows for significant computational improvements since many fewer data points need to be evaluated in every iteration. We also argue that one can select the subsets intelligently by appealing to EM's highly-appreciated likelihood-judgment condition and increment factor. We have demonstrated the method with applications to some simple examples, including computer-generated and real-world networks. The method's strength is its efficiency which leads to high convergence rate and good clustering effect.

As part of future work, we plan to extend our framework in two directions. First, our current model only applied on static networks where no temporal analysis is used for evolution study. We are using our model in dynamic networks to detect communities. Second, so far we only considered the link information. In many applications, the content information is also very important. We are investigating how to incorporate content information into our model.

## REFERENCES

- [1] Girvan M and Newman MEJ. *Proceedings of the National Academy of Sciences*, **2002**, 99(12), 7821-7826.
- [2] Newman MEJ and Leicht EA. *Proceedings of the National Academy of Sciences*, **2007**, 104(23), 9564-9569.
- [3] Ramasco JJ and Mungan M. *Physics and Society E*, **2008**, 77(3), 036122.
- [4] Mungan M and Ramasco JJ. *Journal of Statistical Mechanics: Theory and Experiment*, **2010**, 4, 04028.
- [5] Vazquez A. *Populations and Evolution*, **2008**, 77(6), 066106.
- [6] Kirkpatrick S, Gelatt CD, Vecchi MP. *Optimization by simulated annealing*, *Science*, **1983**, 220(4598), 671-680.
- [7] Duch J and Arenas A. *Physical Review E*, **2005**, 72(2), 027104.
- [8] Condon A and Karp RM. *Random structures and algorithms*, **2001**, 18(2), 116-140.
- [9] Zhang B.; Zhang S.; Lu G. *Journal of Chemical and Pharmaceutical Research*, **2013**, 5(9), 256-262.
- [10] Zhang B.; *International Journal of Applied Mathematics and Statistics*, **2013**, 44(14), 422-430.
- [11] Zhang B.; Yue H.. *International Journal of Applied Mathematics and Statistics*, **2013**, 40(10), 469-476.
- [12] Zhang B.; Feng Y.. *International Journal of Applied Mathematics and Statistics*, **2013**, 40(10), 136-143.
- [13] Bing Zhang. *Journal of Chemical and Pharmaceutical Research*, **2014**, 5(2), 649-659.
- [14] Bing Zhang; Zhang S.; Lu G. *Journal of Chemical and Pharmaceutical Research*, **2013**, 5(9), 256-262.