The Tenth International Symposium on Operations Research and Its Applications (ISORA 2011)
Dunhuang, China, August 28–31, 2011

A Regularized Optimization Model for Community Identification in Complex Networks

Yong Wang[1] Jiguang Wang[1] Chen Chen[1]
Xiang-Sun Zhang[1]

[1]National Center for Mathematics and Interdisciplinary Sciences
Academy of Mathematics and Systems Science, CAS, Beijing 100190

Abstract Identifying community structure is an important step to reveal the functional characteristics of complex networks. Recently, many models and algorithms have been designed to identify communities in a given network. Here, we propose a general mathematical programming framework to optimize some modularity criteria under certain constraints. We then show that several existing models are special cases of our framework by taking different kinds of modularity criteria and constraints. In addition, a regularization term is introduced as an additional objective to consider the parsimony principle in community structure. Experiments on several toy networks show that our new model is simple yet insightful for the community identification problem.

Keywords Optimization model; Complex network; Community identification; Quantitative modularity

1 Introduction

Networks have become a universal language to understand systems of interacting objects, in diverse disciplines including biological organisms, information communication, and human society. As a result, many new words have been invented, such as network biology and network medicine [1, 2]. To study the structure and dynamics of networks, one important strategy is to identify communities or modules [8, 6]. In topological meaning, a community is qualitatively defined as a sub-network whose nodes are connected tightly inside and sparsely to the outside [11, 10, 12, 13]. In reality, community means groups of related nodes that correspond to functional subunits such as protein complexes, biological pathways, or social clubs. Thus uncovering such community structure not only helps us understanding the topological structure, but also reveals the functionality of large-scale networks.

So far, a large number of methods have been developed for detecting communities in complex networks. Methods include but not limited to clique overlap-based hierarchical clustering [15], clique percolation method [16], subgraph fitness method [17], Potts model [20], information theoretical method [21], random walk methods [22], optimization of modularity measures [19, 18], multiple resolution method [9], and space information based method [4]. However, the connections and differences among these methods

remain unclear. Thus it is in pressing need for some general framework to unify these various methods. This framework will provide further insights for community identification problem and allow new method development.

In this paper, we formulate community identification as a general optimization problem. Specifically, we propose a mathematical programming framework to optimize some modularity criteria under certain constraints. We then show that several existing models are special cases of our framework by taking different kinds of modularity criteria and constraints. In addition, a regularization term is introduced as an additional objective to ensure the sparsity of decision variables. Finally experiments on several toy networks show that our new model is simple yet insightful for community identification problem.

2 A general optimization model

Consider an undirected network $G = (V, E)$, where $A = (a_{ij})_{n\times n}$ is the adjacency matrix of G and n is the number of nodes. To understand and reveal the cluster pattern in the network, we adopt a divide-and-conquer strategy to decompose it into communities, which are informally defined as densely connected sets of nodes. In most existing literature, community identification procedure is modeled in a node or vertex based viewpoint, in which the decision variables are whether one node belongs to a community. In other words, we need to decide which vertex belongs to which community. However, the number of communities in the network is unknown and should be treated as a decision variable to be determined in the optimization procedure. Furthermore, communities in networks often overlap such that nodes simultaneously belong to several groups [3].

Due to this difficulty, conceptually we will switch to an edge based viewpoint and communities are viewed as groups of node pairs rather than nodes. We call it node-pair based model and the decision variable is whether one node pair (i, j) belongs to a specified community. In other words, we will decide if for each pair of vertices, it belongs to a same community. Then for each node pair (i, j), we have one variable x_{ij}. $x_{ij} = 1$ means that i and j belong to the same community, and $x_{ij} = 0$ that i and j are in different communities. In essence, the community identification problem is to optimally choose a certain number of node pairs taking value 1 and eventually decompose the network into several communities.

To optimally choose a set of node pais, we need to rank all the node pairs by weighting their importance in community structure. We note that the widely used modularity measure Q will provide some hints. It's well known that community identification can be achieved by optimizing certain modularity measures and one popular modularity measure is the modularity function Q developed by Newman [19]. It is an index for evaluating how good a community partition is.

Given a partition $P_K = (G_1, G_2, \cdots, G_K) = ((V_1,E_1),(V_2,E_2), \cdots,(V_K, E_K))$, where K is the number of candidate communities, the modularity function Q [19] is defined as

$$Q = \sum_{s=1}^{K} \left[\frac{L(V_s, V_s)}{2L} - \left(\frac{L(V_s, V_s) + L(V_s, \overline{V}_s)}{2L} \right)^2 \right], \tag{1}$$

where $L = L(V, V)/2$ is the total number of links in the network, and $L(V_s, V_s) + L(V_s, \overline{V}_s)$ is the total degree of the nodes in community s. This measure provides a way to determine if or not a partition is good enough to decipher the community structure of a

network. Generally, a larger Q corresponds to a better community structure. Maximizing the modularity Q has been a widely accepted method for detecting community structure of complex networks [12].

Alternatively the modularity function Q [19] is defined as

$$Q = \frac{1}{2m}\sum_{i,j}(a_{ij} - \frac{d_i d_j}{2m})\delta(\ell(i),\ell(j)) \tag{2}$$

Here the adjacency matrix of network G is denoted by $A = (a_{ij})$, thus $a_{ij} = 1$ if node i and j share an edge, $a_{ij} = 0$ otherwise. m is the total number of edges. d_i and d_j are the degrees of node i and j. $\ell(i) \in P_K$ denotes the community belongs of node i. For example, $\ell(i) = G_1$ means node i belongs to community G_1. δ denotes the Kronecker Delta, which is 1 if node i and j belong to the same community, and 0 otherwise.

Recalling the definition of decision variable x_{ij}, the modularity function Q can be reformulated as follows,

$$Q = \frac{1}{2m}\sum_{i,j}(a_{ij} - \frac{d_i d_j}{2m})x_{ij} \tag{3}$$

It's easy to see that node pair (i,j) is weighted by $(a_{ij} - \frac{d_i d_j}{2m})/2m$ and maximizing modularity function Q is equivalent to maximize the weight sum of a subset of node pairs ($x_{ij} = 1$). More generally, we suppose the network is weighted by a matrix W modified from adjacency matrix A, where w_{ij} is the weight on the vertex pair (i,j). Then, the objective function is to maximize the summed weights for the selected node pairs. At the same time the node pair selection is usually guided by a set of constraints. Then the community identification can be written as the following general optimization model.

$$\begin{aligned} \max_{x_{ij}} \quad & \textstyle\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}x_{ij} \\ s.t. \quad & x_{ij} \in F, \quad F \quad \text{is a constraint set} \\ & x_{ij} = 0,1, \quad i,j \in \{1,2,\cdots,n\} \end{aligned} \tag{4}$$

Generally speaking, solving integer programming is NP-hard, and thus not suitable for large scale network analysis. Straightforwardly, we can replace the integer constraint, that each x_{ij} is an integer from $0,1$, with the constraint that each x_{ij} is a real number between 0 and 1. Suppose the problem is unconstrained or F is a linear constraint set, then a linear programming can be obtained and solved in polynomial time by the widely used package CPLEX or LPsolver. After solving the linear programming, we will obtain fractional value x_{ij} for every pair of nodes. Then a post-processing step is necessary to identify communities starting from these fractional assignments. Here, some heuristic strategies have to be adopted. Basically, the x_{ij} can be treated as a metric and interpreted as the "distance" between node i and j. We use these distances to greedily and repeatedly find clusters of "tightly connected" nodes, which are then removed. We note that this procedure is not accurate and the computational performance depends on the structure of x_{ij}. This procedure is also called rounding of the LP and a detailed algorithm is presented in [5].

Our model (4) is a general optimization framework. By changing the definition of W and the constraint set F, we can unify different existing models.

2.1 Choices of constraints set F

Since the choices of F are more straightforward than the choices of W, we firstly introduce the different strategies to incorporate different constraints in our optimization model.

If we don't allow the overlaps among the communities, we need to ensure that the x_{ij} are consistent with each other: if i and j are in the same community, and j and k are in the same community, then so are i and k. This constraint can be written as a linear inequality $(1-x_{ik}) \leq (1-x_{jk})+(1-x_{ij})$. For all triples (i,j,k) we obtain the following integer linear programming,

$$\begin{aligned} \max_{x_{ij}} \quad & \tfrac{1}{2m}\textstyle\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}x_{ij} \\ s.t. \quad & (1-x_{ik}) \leq (1-x_{jk})+(1-x_{ij}), \forall i,j,k \\ & x_{ij} = 0,1, \quad i,j \in \{1,2,\cdots,n\} \end{aligned} \tag{5}$$

where m is the number of edges in network G. The advantage is that we introduce a set of linear constraints, while the disadvantage is that too many constraints are introduced into the model. For example if we have n nodes in the network and there will be n^3 constraints in the above model, which increase the complexity of model even when we relax x_{ij} to a real number between 0 and 1. So this model is more efficient for the case of overlapping community structures.

Sometimes we have prior information for the community identification and can easily incorporate into our model. For example the Correlation Clustering problem. Each vertex pair (i,j) in the network is label either $+$ (i and j are similar or highly correlated, and tend to belong to the same community) or $-$ (i and j are dissimilar or lowly correlated, and tend to belong to different communities). A typical application of this problem is illustrated in [23]. To identify biological meaningful modules in protein-protein interaction network, two interacted proteins correlated to each other in gene expression level should belong to the same module, while two proteins which are anti-correlated should belong to different modules. The goal is to partition the network into communities by considering these prior information, which cannot be simply covered by the adjacency matrix A. Let E_+ and E_- be the sets of edges labeled $+$ and $-$, respectively. If these prior information is reliable enough, the problem can be formulated as

$$\begin{aligned} \max_{x_{ij}} \quad & \tfrac{1}{2m}\textstyle\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}x_{ij} \\ s.t. \quad & x_{ij} = 1, \quad (i,j) \in E_+ \\ & x_{ij} = 0, \quad (i,j) \in E_- \\ & x_{ij} = 0,1, \quad i,j \in \{1,2,\cdots,n\} \end{aligned} \tag{6}$$

However, the prior information is usually noisy and should be treated as soft constraints. We then penalize the constraint to objective function, i.e., to minimize the number of "$-$" pairs inside communities and the number of the "$+$" pairs between communities. Then the problem can be formulated as follows (parameters ν and μ are introduced to

weight the importance of prior information),

$$\begin{aligned} \max_{x_{ij}} \quad & \tfrac{1}{2m}\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}x_{ij} + \mu\sum_{(i,j)\in E_+} x_{ij} - \nu\sum_{(i,j)\in E_-} x_{ij} \\ s.t. \quad & x_{ij} = 0,1, \quad i,j \in \{1,2,\cdots,n\} \end{aligned}$$

The simplest case is that constraint set F is empty, i.e., there is no constraints for decision variable x_{ij}. The advantage is that community overlap is allowed. However our model becomes trival given that W is a positive matrix. We will show later that we can utilize this benefit by introducing additional regularization term in objective function.

2.2 Modularity measure Q optimization

In our model (4), we set $W = \frac{1}{2m}(A-P)$, i.e., $w_{ij} = \frac{1}{2m}(a_{ij} - p_{ij}) = \frac{1}{2m}(a_{ij} - \frac{d_i d_j}{2m})$. Then following modularity optimization model in [5] can be deduced from our general model (4).

$$\begin{aligned} \max_{x_{ij}} \quad & \tfrac{1}{2m}\sum_{i=1}^{n}\sum_{j=1}^{n}(a_{ij} - \tfrac{d_i d_j}{2m})x_{ij} \\ s.t. \quad & (1-x_{ik}) \le (1-x_{jk}) + (1-x_{ij}), \forall i,j,k \\ & x_{ij} = 0,1, \quad i,j \in \{1,2,\cdots,n\} \end{aligned} \tag{7}$$

2.3 Multiple resolution community model

Optimization of modularity function Q has been exposed to suffer from a so-called resolution limit problem, i.e., communities in some special networks may not be resolved by optimization of Q even in an extreme case where the network consists of complete graphs connected by single bridges [7]. In other words, optimization of Q fails to zoom in some small qualified communities.

Resolution limit can be avoided by modifying the definition of W. In our model (4), we set $W_\gamma = \frac{1}{2m}(A - P + \gamma I)$, where I is the identity matrix. γ is the factor for multi-resolution community identification [9]. Then $w_{ij} = \frac{1}{2m}(a_{ij} - p_{ij} + \gamma\delta(i,j)) = \frac{1}{2m}(a_{ij} - \frac{d_i d_j}{2m} + \gamma\delta(i,j))$. δ denotes the Kronecker Delta, which is 1 if its arguments are identical, and 0 otherwise.

$$\begin{aligned} \max_{x_{ij}} \quad & \tfrac{1}{2m}\sum_{i=1}^{n}\sum_{j=1}^{n}(a_{ij} - \tfrac{d_i d_j}{2m} + \gamma\delta(i,j))x_{ij} \\ s.t. \quad & (1-x_{ik}) \le (1-x_{jk}) + (1-x_{ij}), \forall i,j,k \\ & x_{ij} = 0,1, \quad i,j \in \{1,2,\cdots,n\} \end{aligned} \tag{8}$$

This idea is originally introduced in [9]. The new definition of W by introducing parameter γ allows the screening of the modular structure by analyzing the optimal modular structure of W_γ for different values of γ. And the topological structure reveals large groups for small value of γ, and smaller groups for large value of γ. In addition, we note this γ strategy is general and in principle can be used in any quality function.

2.4 Space-independent community model

In a recent paper [4], the authors show that the spatial information plays a crucial role by affecting, directly or indirectly, network community structure. In a similar manner, space information can be readily introduced in the definition W as $W = \frac{1}{2m}(A - P^{spa})$, where the definition of P^{spa} is from Equation 5 in [4]. The physical meaning is that W favors communities made of node i and j such that w_{ij} is large, i.e., pairs of nodes with short distance.

$$\begin{aligned} \max_{x_{ij}} \quad & \tfrac{1}{2m}\textstyle\sum_{i=1}^{n}\sum_{j=1}^{n}(a_{ij} - p_{ij}^{spa})x_{ij} \\ s.t. \quad & (1-x_{ik}) \le (1-x_{jk}) + (1-x_{ij}), \forall i,j,k \\ & x_{ij} = 0,1, \quad i,j \in \{1,2,\cdots,n\} \end{aligned} \tag{9}$$

2.5 More choices of W

Behind the definition of modularity $Q = \frac{1}{2m}\sum(A-P)$, the concept is that links are more abundant within communities than would be expected on the basis of chance. Here the meaning of "by chance" (i.e. the null hypothesis for community identification) is embodied by the matrix $P = (p_{ij})$. In the current definition Q, p_{ij} is the expected weight of a link between node i and j over an ensemble of random networks with certain constraints. The most popular choice, proposed by Newman and Girvan, is $p_{ij} = \frac{d_i d_j}{2m}$, which is theoretical estimation of the true p_{ij}.

The above analysis allows further extension of W. Different null models can be constructed depending on the network under consideration. For example, we can formulate the following optimization model,

$$\begin{aligned} \max_{x_{ij}} \quad & \tfrac{1}{2m}\textstyle\sum_{i=1}^{n}\sum_{j=1}^{n}(a_{ij} - s_{ij})x_{ij} \\ s.t. \quad & (1-x_{ik}) \le (1-x_{jk}) + (1-x_{ij}), \forall i,j,k \\ & x_{ij} = 0,1, \quad i,j \in \{1,2,\cdots,n\} \end{aligned} \tag{10}$$

where s_{ij} is the null hypothesis matrix derived by numerical simulation. For instance, s_{ij} can be calculated by random shuffling the network 1,000 times keeping the degree distribution unchanged. We believe that s_{ij} will provide more accurate approximation for the true p_{ij}.

We also note that adjacent matrix A of network G is used in the above models. Similarity, other substitutive matrix, such as Laplacian matrix and diffusion kernel, will lead to more choices of W.

3 A regularized optimization model

In essence, the community identification problem is to optimally choose a certain number of node pairs and eventually decompose the network into several communities. We have already introduced the objective to maximize the selected weights. Another natural objective function is parsimony, i.e., to minimize the number of selected node pairs (to maximize the number of zero variables in our model (4). The motivation is as follow. Suppose in extreme case we have a network equally decomposed into k isolated

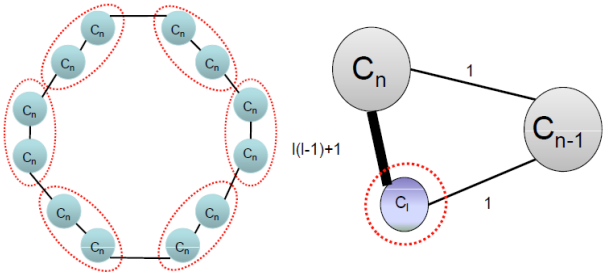

Figure 1: Illustration of two toy examples for the resolution limit and misidentification phenomena. The left network consists of a ring of n cliques ($n \geq 3$), connected through single links. Assuming that there are 2^k cliques. The network has a clear modular structure where each community corresponds to a single clique. But optimizing Q combines two neighboring cliques as one community and fails to obtain the correct partition. The right network consists of three cliques C_n, C_{n-1}, C_l with different sizes. When $n >> l$, the clique C_l is not a qualified community, however, optimization Q will identify C_l as a community in some range of n.

communities. Every community is a clique. Then the percentage of non-zero values in the solution x_{ij} (selected edges) is roughly $1/k$. When k is large, a lot of x_{ij} are equal to zero. Strong regularization is necessary in this case. Here we use regularization term $\sum_{ij} x_{ij}$ to maximize the number of zero variables.

$$\max_{x_{ij}} \quad \sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}x_{ij} - \lambda \sum_{i=1}^{n}\sum_{j=1}^{n} x_{ij} \tag{11}$$

$$s.t. \quad x_{ij} \in F, \quad F \quad \text{is a constraint set} \tag{12}$$

$$x_{ij} = 0,1, \quad i,j \in \{1,2,\cdots,n\}$$

In this way, we can propose a new method for modularity identification. All the above models can be easily improved by introducing regularization term. Next we will consider the simplest case to illustrate the advantage of regularization idea. If there is no constraints in (11), our model can be easily solved. The advantage is that community overlap is allowed. The solution we use is usually called soft thresholding: each x_{ij} is reduced by an amount λ in absolute value and is set to zero if its absolute value is less than zero. Algebraically, soft thresholding is defined by

$$x_{ij} = sign(w_{ij})(|w_{ij}| - \lambda)_+ \tag{13}$$

where $+$ means positive part ($y_+ = y$ if $y > 0$ and zero otherwise).

4 Pilot studies on two toy examples

We use two toy examples to conceptually show the advantage of introducing regularization term in model (11).

We first apply our method in the well-known ring of clique example (left network in Figure 1). Resolution limit problem means that communities in some special networks

may not be resolved by optimization of Q even in an extreme case where the network consists of complete graphs connected by single bridges [7]. In other words, optimization of Q fails to zoom in some small qualified communities. Suppose we have k cliques, each clique has n nodes. Then the number of total edges is $m = kn(n-1)/2 + k$. We calculate the value of matrix W as

$$w_{ij} = \begin{cases} \frac{1}{2m}(1 - \frac{n^2}{2m}) & \text{if } (i,j) \in C_n \text{ and contain no outside connecting node,} \\ \frac{1}{2m}(1 - \frac{n(n+1)}{2m}) & \text{if } (i,j) \in C_n \text{ and contain outside connecting node,} \\ \frac{1}{2m}(1 - \frac{(n+1)^2}{2m}) & \text{if } (i,j) \text{ connects two adjacent cliques,} \\ 0 & \text{otherwise.} \end{cases} \tag{14}$$

We calculate the solution of our model by soft thresholding (13). When we choose $\frac{1}{2m}(1 - \frac{(n+1)^2}{2m}) < \lambda < \frac{1}{2m}(1 - \frac{n(n+1)}{2m})$, every clique C_n can be correctly identified as one single community. This demonstrates that our regularized model does not suffer from resolution limit in this well-known example.

In addition to the resolution limit phenomenon, there is another serious limitation in optimization of Q, the misidentification phenomenon [14], which means that some derived communities do not satisfy the weak community definition[10]. In other words, these communities have sparser connection within them than between them which disobeys the basic intuitive sense for a subnetwork to be a community. A toy example is proposed in [24] and redrew as the right network in Figure 1. Then we consider to calculate the value of matrix W. Since there are many different cases for the connection among the three cliques, the resulting W is complicated and depends on different situations, so we will not show the close form here. Without loss of generality, we consider the average case. Our model by soft thresholding (13) will rank all the node pairs by their w_{ij} values. Averagely the single link connection C_n and C_{n-1} will be firstly set to zero by choosing proper λ. And since $l << n$, the clique C_l will not be identified as a single community. This demonstrates that in principle our regularized optimization model (11) does not suffer from misidentification in this example.

5 Discussions

In summary, we propose a general mathematical programming framework to optimize some modularity criteria under certain constraints, which provides a new visual angle for the community structure problem. In this framework, we can unify several existing models as special cases of our framework by choosing different kinds of modularity criteria and constraints. Then the difference and connections of these methods can be revealed and further insights for improvement is allowed. Importantly, we propose a regularization term as an additional objective function to identify community. In essence, community identification is formulated as a multiple objective optimization problem. We want to maximize the modularity function and minimize the number of the edges at the same time. To achieve this, we introduce parameter λ to balance the two terms. Experiments on two toy networks show that our new model is simple yet insightful for the community identification problem. However there is plenty of room to improve and sharp our pilot model. For example, the procedure to choose proper parameter λ should be carefully

studied. Furthermore, we need to carefully study the post-processing algorithm to identify module after we solve the optimization model. Also experiments in large and real complex networks are also needed. These are all our further research directions.

Acknowledgements

The authors would like to thank ZHANGroup members for insightful discussions. YW, JW, and XSZ are supported by NSFC grant 10801131 and grant kjcx-yw-s7 from CAS. YW is also supported by SRF for ROCS, SEM and the Shanghai Key Laboratory of Intelligent Information Processing (No. IIPL-2010-008).

References

[1] Barabasi A.-L. , Oltvai Z. N. (2004) Network biology: understanding the cell's functional organization Nature Reviews Genetics 5, 101–113.

[2] Barabasi A.-L., Gulbahce N. , Loscalzo J. (2011) Network medicine: a network-based approach to human disease. Nature Reviews Genetics 12, 56–68.

[3] Ahn Y. , Bagrow J. P. , Lehmann S. (2010) Link communities reveal multi-scale complexity in networks. Nature, 466:761.

[4] Expert P., Evans T. S., Blondel V. D., and Lambiotte R. (2011) Uncovering space-independent communities in spatial networks. Proc Natl Acad Sci USA 108(19):7663–7668.

[5] Agarwal G, Kempe D (2008) Modularity-maximizing graph communities via mathematical programming. Eur Phys J B 66(3):409–418

[6] Fortunato S (2010) Community detection in graphs. Phys Rep 486:75-174

[7] Fortunato S, Barthelemy M (2007) Resolution limit in community detection. Proc Natl Acad Sci USA 104(1):36–41

[8] Girvan M, Newman M (2002) Community structure in social and biological networks. Proc Natl Acad Sci USA 99(12):7821–7826

[9] Arenas A, Fernandez A, Gomez S(2008) Analysis of the structure of complex networks at different resolution levels. New J Phys 10:053039

[10] Radicchi F, Castellano C, Cecconi F, Loreto V, Parisi D (2004) Defining and identifying communities in networks. Proc Natl Acad Sci USA 101(9):2658–2663

[11] Ravasz E, Somera A, Mongru D, Oltvai Z, Barabasi A (2002) Hierarchical organization of modularity in metabolic networks. Science 297(5586):1551–1555

[12] Newman M (2006) Modularity and community structure in networks. Proc Natl Acad Sci USA 103(23):8577–8582

[13] Wang RS, Zhang S, Wang Y, Zhang XS, Chen L (2008) Clustering complex networks and biological networks by nonnegative matrix factorization with various similarity measures. Neurocomputing 72:134–141

[14] Zhang XS, Wang RS, Wang Y, Wang J, Qiu Y, Wang L, Chen L (2009) Modularity optimization in community detection of complex networks. Europhys Lett 87(3):38002

[15] Everett M, Borgatti S (1998) Analyzing clique overlap. Connections 21(1):49–61

[16] Palla G, Derényi I, Farkas I, Vicsek T (2005) Uncovering the overlapping community structure of complex networks in nature and society. Nature 435(7043):814–818

[17] Lancichinetti A, Fortunato S, Kertész J (2009) Detecting the overlapping and hierarchical community structure in complex networks. New J Phys 11:033015

[18] Li Z, Zhang S, Wang RS, Zhang XS, Chen L (2008) Quantitative function for community detection. Phys Rev E 77(3):036109

[19] Newman M, Girvan M (2004) Finding and evaluating community structure in networks. Phys Rev E 69(2):026113

[20] Reichardt J, Bornholdt S (2004) Detecting fuzzy community structures in complex networks with a Potts model. Phys Rev Lett 93(21):218701

[21] Rosvall M, Bergstrom C (2007) An information-theoretic framework for resolving community structure in complex networks. Proc Natl Acad Sci USA 104(18):7327–7331

[22] Rosvall M, Bergstrom C (2008) Maps of random walks on complex networks reveal community structure. Proc Natl Acad Sci USA 105(4):1118–1123

[23] Xue H., Xian B., Dong D., Xia K., Zhu S., Zhang Z., Hou L., Zhang Q., Zhang Y., and Han J. D. (2007) A modular network model of aging. Molecular Systems Biology, 3:147.

[24] Zhang X-S., Li Z., Wang R-S., and Wang Y. (2011) A combinatorial model and algorithm for globally searching community structure in complex networks, Journal of Combinatorial Optimization, in press.