Significance analysis for community partition across multiple optimization models

Hui-Jia Li[1,2] [(a)], Luonan Chen[2]

[1] *School of Management Science and Engineering, Central University of Finance and Economics, Beijing 100080, China.*
[2] *Key Laboratory of Systems Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200233, China.*

PACS 89.75.Hc – First pacs description
PACS 89.75.Fb – Second pacs description

Abstract –The study of community structure is an important problem in a wide range of applications, which can help us understand the real network system deeply. However, due to the exist of random factors and error edges in real networks, how to measure the significance of community structure efficiently is a crucial question. In this paper, we present a novel statistical framework computing significance of community structure across multiple optimization methods. Different from the universal approaches, we calculate the similarity between a given node and its leader and employ the distribution of link tightness to derive the significance score, instead of a direct comparison to a randomized model. Based on the distribution of community tightness, a new "p-value" form significance measure is proposed for community structure analysis. Specially, the well-known approaches and their corresponding quality functions are unified to a novel general formulation, which facilitate providing a detail comparison across them. To determine the position of leaders and their corresponding followers, an efficient algorithm is proposed based on the spectral theory. Finally, we apply the significance analysis to some famous benchmark networks and the good performance verified the effectiveness and efficiency of our framework.

1 Introduction . – In many real networks, a common feature observable is the presence of community structures [1]- [8], i.e. subset of vertices which are densely connected to each other while less connected to the vertices outside. In many scenarios, community detection methods can help to unveil the functional properties of the complex networks, thus there is a necessity to devise better community detection methods which meet both speed and accuracy requirements simultaneously [9]- [13]. In order to estimate how much a decomposition of a network which is found by a community detection algorithm is meaningful, we need a quality measure. Consequently, for a particular measure, the community detection algorithms can be ranked. To this end, various measures have been proposed in the literature, so far. The most prevalent measure which has been used extensively in the literature is due to Newman & Girvan [4]. This measure, called modularity, quantifies how much the density of the edges inside identified communities differs from the expected edge density in an equivalent network with similar number of vertices and edges but randomized edge placement, which is taken as the null model for statistical tests. Considering the modularity measure, the community detection problem is transformed to the modularity maximization problem. Moreover, some optimization algorithms based on Potts models which used to detect community structure have attracted attention. Communities correspond to Potts model spin states, and the associated system energy indicates the quality of a candidate partition. For more optimization functions in detail, please find in Supplementary Material [28].

Although a lot of optimization method and their functions are proposed, some important questions remain unclearly answered, that are about the significance of the communities in real networks. Are the communities partitioned by different optimization methods are truly significant or they are just the coincidence of edge positions in the network [14] [15]? How to determine the significance of a given community effectively? To answers these crucial questions, in this paper, we present a novel statis-

[(a)] Corresponding authors: Hjli@amss.ac.cn

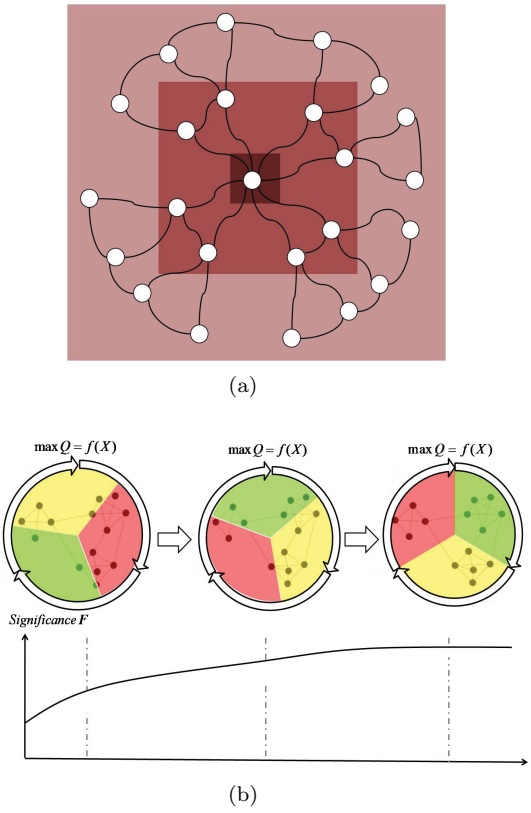
tical framework comparing the significance of community structure across various optimization methods. Different from the universal approaches, we calculate the similarity of a given node to its leader and employ the distribution of link tightness to derive the significance score, instead of a direct comparison to a randomized model. A small example is shown in Fig.1(a), which illustrates that tighter the following nodes link to its leader, more significant the community is. Based on the distribution of community tightness, a new "p-value" form significance measure is proposed for community structure analysis. Specially, the well-known approaches and their corresponding quality functions are unified to a novel general formulation, to provide a detail comparison across them. Then, we can choose the most suitable form of the function by set the parameters properly. To determine the position of leaders and their corresponding followers, an efficient detection algorithm is proposed based on the spectral theory. Finally, we apply the significance analysis to some famous benchmark networks and the good performance verified the effectiveness and efficiency of our framework.

2 The framework.

2.1 community structure and the leader.

Leader-driven algorithms [20] [21] constitute a special case of seed-centric approaches. These methods show that, in many real world, especial the social networks, nodes of a network are usually classified into two categories: leaders and followers. For each community, the most central node is selected as a leader, and a given leader node represents a specific community. Follower nodes are assigned to the most nearby leader node and together form a community. The leaders should have two properties: they are well connected to the members of their group, and they are able to communicate with other leaders when necessary. If the distributed algorithm is carried out in each group separately and the leaders communicate at a higher level, the nodes can enjoy faster convergence rate.

For example, considering the famous Karate network [19], nodes 1 and 33 are two significant leaders and corresponding communities are built around them. If two leaders are removed, these communities will be split up, as they link to most followers and keep the community together. Since community are consequence of information spreading, a given community can be defined as the area in which a leader has most influence. So, one can uncover the community partition by finding all natural leaders and their corresponding followers on which they influence. We believe if followers are more tightly linked to the leader, or leader spreads more influence on their followers, this community are more significant or robust. When we use a given optimization method to evolve the community configure, the significance of communities also evolves correspondingly, which shown in Fig.1(b). The function and



(a)

(b)

Fig. 1: (a) For a given community, the leader node usually locates on the highest level, representing the most influential node. Circles depict different levels in the network hierarchy, with the darkest color denoting the highest level. Tighter the following nodes link with its leader, more significant the community is. (b) In every circle, sectors with different colors represent different communities. It can be noticed that the community partition in the rightmost circle is strongest due to the fewest intercommunity edges. When we use a given optimization method to evolve the community configure X (describe by different sectors) based on maximizing the objective function $\max Q = f(X)$, the significance of X also evolves correspondingly. The F score is utilized to measure the significance of community configure X. Here, the global maximum of F is maybe an asymptotically stable fixed point of dynamical system associates to community configure X in the rightmost circle.

computation methods of significance are valuable which can be utilized to measure the quality of community configure. Specifically, the global maximum of significance maybe an asymptotically stable fixed point associates to community configure dynamics, which deserves us to study it deeply.

2.2 The community detection algorithm based on leader position.

In this study, the relative positions of leader and corresponding followers are crucial to analyze the significance situation. In order to obtain the leader of corresponding community, we extract the candidate community member-

ship by minimizing the following objective function

$$J_m = \sum_{i=1}^{n}\sum_{j=1}^{k} x_{ij}\|d_i - c_j\|^2, \quad (1)$$

where variables x_{ij} is the membership that node i in community j, with $\sum_j x_{ij} = 1$. This method is similar as the famous k-means method and can be obtain both center and assignment iteratively. d_i is the ith n-dimensional data point, c_j is the n-dimensional center(leader) of the community j, and $\| * \|$ is any norm expressing the similarity between a given node and the center. One can use an iterative optimization of the objective function shown above, to obtain the network partition by the update of membership x_{ij} and the community leaders c_j . This procedure converges to a local minimum or a saddle point of J_m.

Suppose K is the upper bound of number of clusters and $A = (a_{ij})_{n\times n}$ is the adjacent matrix of a network, then the algorithm is stated straightforwardly as follows(for the detailed algorithm framework, please find in Supplementary Material [28]):

Step 1: for a given K

(i) Calculate the diagonal matrix $D = (d_{ii})$, where $d_{ii} = \sum_k a_{ik}$.

(ii) Computing the top K eigenvectors based on generalized eigensystem $Ax = tDx$, and then establish the eigenvector matrix $E_K = [e_1, e_2, ..., e_K]$ by .

Step 2: for each number of communities $2 \le k \le K$:

(i) Establish the matrix $E_K = [e_2, e_3, ..., e_K]$ from the matrix E_K.

(ii) Normalize the rows of E_K to unit length using Euclidean distance norm.

(iii) Cluster the row vectors of E_K using any community detection method by minimizing Eq.(1) to obtain a membership matrix X_k and corresponding leaders.

Step 3: Maximizing the modular function: Pick the optimal number of communities k and the corresponding partition X_k that maximizes $Q(X_k)$.

In step 1, given the adjacent matrix $A = (a_{ij})_{n\times n}$ and a diagonal matrix $D = (d_{ii})$, $d_{ii} = \sum_k a_{ik}$, two matrices $D^{-1/2}AD^{-1/2}$ and $D^{-1}A$ are used. This is motivated by Ref. [22], which uses the top K eigenvectors of the generalized eigensystem $Ax = tDx$ instead of the K eigenvectors of the adjacent matrix. It shows that after normalizing the rows using Euclidean norm, their eigenvectors are mathematically identical and emphasize that this is a numerically more stable method. Although their result is designed to cluster real-valued points [22] [23], it is also appropriate for network clustering.

In step 2, we choose the initial the starting centers to be as orthogonal as possible which already used in k-means clustering method [23] [24]. This way of choosing centers(leaders) does not cost additional time complexity, and also improve the quality of the partition, thus at the same time reduces the need for restarting the random initialization process. Specially, recording the label of leaders is crucial to compute the significance score which will be illustrated in the next two sections.

In step 3, the Q function measures the quality of a given community structure organization of a network and can be used to automatically select the optimal number of communities k according to the maximum Q value [24], we will discuss the multiple optimization methods and their corresponding Q function in detail in the following section.

2.3 The general and expanded formation of function Q.

For many community detection algorithms, the target function Q is critical. Here, we find that Q can be tried to be optimized has the following general form:

$$Q = -\frac{1}{2}\sum_{\mu}\sum_{i=1}^{n}(\sum_{j=1}^{n} f_\mu^+ a_{ij}x_{i\mu}x_{j\mu} - \sum_{j=1}^{n} f_\mu^-(1-a_{ij})x_{i\mu}x_{j\mu}) + \sum_{\mu} R_\mu, \quad (2)$$

which can be rephrased as,

$$\begin{aligned} Q &= -\frac{1}{2}\sum_{\mu}[\frac{2\sum_{j=1}^{n} x_{j\mu}}{l_\mu}R_\mu + \sum_{i=1}^{n}(\sum_{j=1}^{n} f_\mu^+ a_{ij}x_{i\mu}x_{j\mu} \\ &\quad - \sum_{j=1}^{n} f_\mu^-(1-a_{ij})x_{i\mu}x_{j\mu})] \\ &= -\frac{1}{2}\sum_{\mu}\sum_{i=1}^{n} x_{i\mu}(\sum_{j=1}^{n} f_\mu^+ a_{ij}x_{j\mu} - \sum_{j=1}^{n} f_\mu^-(1-a_{ij})x_{j\mu} \\ &\quad + \frac{2}{l_\mu}R_\mu), \end{aligned} \quad (3)$$

where $l_\mu = \sum_{j=1}^{n} x_{j\mu}$ is the size of the community μ. In fact, one can interpret these kinds of measures as different rewarding-punishing strategies. Each choice of parameters has its own intuition, strengths and drawbacks.

Based on Eq.(3), let us define the following function,

$$Q_{i\mu} = \sum_{j=1}^{n} f_\mu^+ a_{ij}x_{j\mu} - \sum_{j=1}^{n} f_\mu^-(1-a_{ij})x_{j\mu} + R_{i\mu}, \quad (4)$$

and choose $R_{i\mu}$ such that $\partial R_{i\mu}/\partial x_{i\mu} = 0$ and $R_{mu} = \sum_{i=1}^{n} R_{i\mu}$, e.g. $R_{i\mu} = \frac{2}{l_\mu}R_{mu}$.

Interestingly, when all $x_{i\mu}$ are in hard membership state, the H function with $Q_{i\mu}$ defined as Eq.(4) can be reduced to well-known optimization measures by following considerations:

(1) Hofman & Wiggins [6]

$$f_\mu^+ = \log\frac{p^{in}}{p^{out}}, f_\mu^- = \log\frac{1-p^{out}}{1-p^{in}}, R_\mu = l_\mu \log\pi_\mu. \quad (5)$$

(2) Ronhovde & Nussinov [7]

$$f_\mu^+ = 1, f_\mu^- = \min_\mu p_{in,\mu}, R_\mu = 0. \quad (6)$$

(3) RB Potts model (Erdős-Rényi null model) [5]

$$f_\mu^+ = 1 - \gamma_{RB} p, f_\mu^- = \gamma_{RB} p, R_\mu = 0. \quad (7)$$

(4) RB Potts model (Configuration null model) [5]

$$f_\mu^+ = 1 - \frac{\gamma_{RB}}{2m}, f_\mu^- = \frac{\gamma_{RB}}{2m}, R_\mu = \sum_{i>j} \frac{\gamma_{RB}}{2m}(k_i k_j - 1) x_{i\mu} x_{j\mu}. \quad (8)$$

where k_i is the degree of node i and m is the number of all edges in the network.

(5) Modularity [4]

$$f_\mu^+ = 1, f_\mu^- = \frac{k_i k_j}{2m}, R_\mu = \sum_{i>j} \frac{1}{2m}(k_i k_j - 1) x_{i\mu} x_{j\mu}. \quad (9)$$

where k_i is the degree of node i and m is the number of all edges in the network.

(6) Label propagation [9]

$$f_\mu^+ = 1, f_\mu^- = 0, R_\mu = 0. \quad (10)$$

where k_i is the degree of node i and m is the number of all edges in the network.

3 Significance of community structure. – It is essential to establish a detail framework analyzing the significance of community structure, since real networks own specific characteristics [16] [17] [18]. In this section, we discuss these important characteristics and give a detailed introduction of the framework.

3.1 Node similarity.

We define the similarity of nodes i and j, $sim(i, j)$, as the ratio between the intersection and the union of their neighborhoods $\Gamma(i)$ and $\Gamma(j)$,

$$sim(i, j) = \frac{|\Gamma(i) \cap \Gamma(j)|}{|\Gamma(i) \cup \Gamma(j)|}, \quad (11)$$

By employing Eq.(11), we can calculate the expected similarity between a given node and the community leader z,

$$E[sim(x, z)] = \int_{\mathbb{R}^M} sim(x, z) Q(x|z) dx, \quad (12)$$

where $Q(x|z)$ is a distribution of nodes in a community with leader z.

Next, Using the maximum entropy principle(See the Section 4 in Supplementary Material), the statistical unbiased distribution fulfilling constraint can be obtained using the maximum entropy principle:

$$Q(x|z, \eta) = \frac{1}{Z_\eta} P_0(x) e^{\eta sim(x,z)} dx, \quad (13)$$

where $P_0(x)$ is the background distribution used to contrast with an alternative hypothesis: node x being part of a community, a group of nodes distinguished by enhanced mutual similarity. Z_η is the normalisation constant depends on the value of the *scoring parameter* η:

$$\frac{\partial}{\partial \eta} \log Z_\eta = E[sim(x, z)]. \quad (14)$$

η is the parameter which used to control the "width" of a community and the larger the value of η, the smaller the expected width or scale of a given community. Specially, the distribution $Q(x|z, \eta)$ is the same as the background model $P_0(x)$ when $\eta = 0$.

3.2 Log-likelihood score and community tightness.

We define the log-likelihood score as the deviations of the community distribution from the null model

$$s(x|z, \eta) \equiv \log \frac{Q(x|z, \eta)}{P_0(x)} = \eta sim(x, z) - \log Z_\eta. \quad (15)$$

By Eq.(15), nodes which are more likely to be in a community with center z and scoring parameter η own larger positive value, than in the null background model. Given a community with nodes set $\{1, ..., N\}$, for a given leader z and a scoring parameter η, the log-likelihood scores $s(i|z, \eta)$ are positive. The community tightness is the sum of the scores of the community elements,

$$S(1, ..., N|z, \eta) = \sum_i \max[s(i|z, \eta), 0]. \quad (16)$$

However, we can't determine the scoring parameter η easily. Here, the tightness function of Eq.(16) can be simplified as:

$$S(1, ..., N|z, \eta) = \sum_{i=1}^{N} \max[s(i|z) - \mu, 0], \quad (17)$$

where $s(i|z) = sim(i, z)$. By this transformation, one can control the width of community using parameter μ simply. The community tightness is determined both by the number of elements and by their similarities with the leader, that is, tighter communities with fewer elements own comparable more tightness to looser but larger communities.

3.3 Calculation of Significance score.

We can the quantified the quality of the true and random communities by characterize the distribution of the tightness score $p(S)$ from the background distribution. A new "p-value" form measure [25] can be used to define the statistical significance of score S_0, as the probability that a random chosen nodes set contains a community with score

greater than or equal to S_0. This "p-value" form significance can be explained by a null hypothesis: "These nodes are drawn from the background distribution". To test this hypothesis, we compute the statistical significance of score S_0: low value suggests that the null hypothesis is unlikely and allows for rejecting it. This method provides a new connection between statistical p-value theory and network analysis and then get an interesting significance measure.

If the network is large enough, according to the mean field theory, $s_i = s(i|z)$ owns an approximate Gaussian-distribution with variance M, $P(s(i|z)) = \sqrt{1/(2M\pi)}\exp\{-s^2/(2M)\}$. The distribution of the tightness S can be calculated straightforwardly using the derivation shown in Supplementary Material [28]. Specifically, we need to compute the following quality function:

$$\begin{aligned} &Z_c(\beta,\mu) = \textstyle\int_{\mathbb{R}^N} e^{\beta S(1,\ldots,N|z,\eta)} P(s_1)\ldots P(s_N) ds_1 \ldots ds_N \\ &= [\textstyle\int_{-\infty}^{+\infty} e^{\beta \max[s_i-\mu,0]} P(s) ds]^N \\ &= [\textstyle\int_{-\infty}^{\mu} P(s) ds + \int_{\mu}^{+\infty} e^{\beta(s_i-\mu)} P(s) ds]^N \\ &= [(1-H(\mu)) + e^{\frac{(\beta)^2}{2}-\beta\mu} H(\mu-\beta)]^N, \end{aligned} \tag{18}$$

where $H(x) = \int_x^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2}$ is the complementary cumulative Gaussian distribution. In Eq.(18), two intervals are divided: below the score threshold μ, the score is zero, which contributes the cumulative distribution $\int_{-\infty}^{\mu} ds/(2\pi)^{1/2}\exp[-s^2/2]$ to the generating function. Above μ, the score is positive, which generates a contribution of $\int_{\mu}^{+\infty} ds/(2\pi)^{1/2}\exp[-s^2/2+\beta(s-\mu)]$. The free energy function reads

$$-\beta f(\beta,\mu) = \log[(1-H(\mu)) + e^{\frac{(\beta)^2}{2}-\beta\mu} H(\mu-\beta)], \tag{19}$$

and the entropy is

$$\omega(s,\mu) = -\max_{\beta}[\beta s + \beta f(\beta,\mu)]. \tag{20}$$

According to the distribution of community tightness(See the Section 5 in Supplementary Material),

$$\log p(S,\mu) \simeq N\omega(S/N,\mu) - \frac{1}{2}\log N. \tag{21}$$

Given a specifical community, we can calculated the significance score F using the probability that the community tightness S, $p(S)$, larger than or equal to S,

$$F(S,\mu) = \int_S^{+\infty} p(S',\mu) dS'. \tag{22}$$

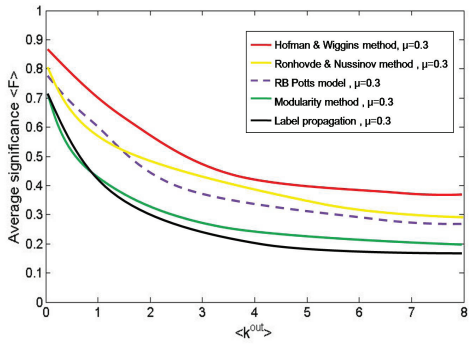
Furthermore, from the perspective of the whole network, we use the average significance score $\langle F\rangle_Q$ to indicate the robustness of a partition, defined as the average value among F values of all communities partitioned by maximizing a particular quality function Q shown in section.

4 Experiments. – We will test the validity of our framework on some famous benchmark network and real networks. Experiments are designed and implemented for two main purposes: (1) to evaluate the performance of a given optimization algorithm; (2) to test the effectiveness and efficiency of our method. Here, we use famous Girven-Newman benchmark as example, for more results on benchmarks such as LFR network, stochastic block model and real networks, please find in Supplementary Material [28].

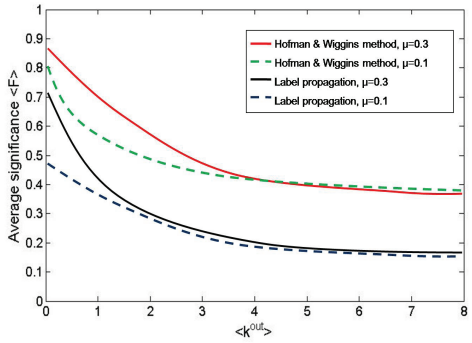
First, we apply to the classical Girven-Newman benchmark [26], where the network with $n = 128$ nodes are divided into four 32 nodes communities. Edges are established with different probabilities according to belong to the same community or not. Every node owns average $\langle k^{in}\rangle$ links with nodes in its own group and $\langle k^{out}\rangle$ links with the rest of the network. According to the establish mechanism, the community structure will fuzzier and thus when $\langle k^{out}\rangle$ increases, it is more difficult to identify them correctly. Hence, the significance of communities will tend to be weaker and the value of F index will also decrease. The comparison results of F value corresponding to all five optimization algorithms are shown in Fig.2(a) when $\mu = 0.3$. It can be observed that the index F has a great performance on GN benchmark: when $\langle k^{out}\rangle$ approaching 0, the community structure is quite strong and all corresponding $\langle F\rangle$ value is close to 1; while when the network is fuzzy enough, the corresponding $\langle F\rangle$ value of all algorithm is low, extremely for Modularity optimization method and Label propagation method, only near 0.2-0.3.

Moreover, by comparing five algorithms, we find in Fig.2(a) that the $\langle F\rangle$ values corresponding to Hofman & Wiggins method is largest, and the Label propagation method is the lowest. This may because Label propagation method emphasize the simplicity of calculation too much while ignoring the accuracy of results. Furthermore, the $\langle F\rangle$ values between Modularity optimization method and Label propagation method are similar when $\langle k^{out}\rangle$ becomes lower. This result is similar as Ref [10] and [27], which verifies the inner correlation between these two methods. These observations are no evidence of overall superiority of one method over another, but an example of how to compare the significance and use the different partitioning algorithms on a given network.

Furthermore, when $\langle k^{out}\rangle$ increases, the topology becomes fuzzier and the sizes of communities will become more and more smaller correspondingly. At the same time, as the width parameter μ increases, the significance will favor tighter communities with fewer elements. We test the Hofman & Wiggins method and Label propagation method in Fig.2(b), the value of $\langle F\rangle$ corresponding to $\mu = 0.3$ will be larger than $\mu = 0.1$ for all two examples. As a conclusion, we argue that when the corresponding $\langle F\rangle$ is smaller than 0.3 on average($\langle k^{out}\rangle \approx 4$), it is not safe to say there exists significant community structure for a given network.



(a)



(b)

Fig. 2: The experimental results of significance $\langle F\rangle$ on GN benchmark network and each point in curves is obtained by testing 100 times. (a) For all five optimization methods, $\langle F\rangle$ decreases with increasing of $\langle k^{out}\rangle$. For a given network, when $\langle F\rangle$ is larger than 0.3 on average($\langle k^{out}\rangle \approx 4$), one can say there exit significant community structure. (b)The value of $\langle F\rangle$ corresponding to $\mu = 0.3$ will be larger than $\mu = 0.1$ for the Hofman & Wiggins method and Label propagation method. This implies as the width parameter μ increases, the significance favors tighter communities with fewer elements.

5 Discussion. – It is unreasonable to analyze the community structure only using the topology information with considering the significance. In this paper, we present a novel framework calculating the significance of community structure revealed by multiple optimization functions. As part of the future work, it is necessary to take a deeper look into how different similarity measures impact the results of this method. Additionally, this framework can be easily extended to a weighted, directed and overlapping form, which only needs to modify the formation of the quality function Q. In conclusion, this method has a great performance and deserves more attention from us.

* * *

We are grateful to the anonymous reviewers for their valuable suggestions. The authors are separately supported by NSFC grants 71401194, 11131009, 91324203 and "121" Youth Development Fund of CUFE grants QBJ1410.

REFERENCES

[1] S. Fortunato. *Phys Rep*, 486(3-5), (2010), 75-174.
[2] Y. Y. Ahn, J. P. Bagrow and S. Lehmann. *Nature*, 466(7307), (2010), 761-764.
[3] F. Folino and C. Pizzuti. *IEEE T Knowl Data En*, 26(8), (2014), 1838-1852.
[4] M. E. J. Newman and M. Girvan. *Phys. Rev. E*, 69(2), (2004), 026113.
[5] J. Reichardt and S. Bornholdt. *Phys. Rev. E*, 74(1), (2006), 016110.
[6] J. M. Hofman and C. H. Wiggins. *Phys. Rev. Lett.*, 100(25), (2008), 258701.
[7] P. Ronhovde and Z. Nussinov. *Phys. Rev. E*, 81(4), (2010), 046114.
[8] M. E. J. Newman. *Phys. Rev. E*, 69(6), (2004), 066133.
[9] U. N. Raghavan, R. Albert and S. Kumara. *Phys. Rev. E*, 76(3), (2007), 036106.
[10] G. Tibély and J. Kertész. *Physica A*, 387(19-20), (2008), 4982-4984.
[11] R. Guimera and L. A. Nunes Amaral. *Nature*, 433(7028), (2005), 895-900.
[12] J. Duch and A. Arenas. *Phys. Rev. E*, 72(2), (2005), 027104.
[13] M. E. J. Newman and E. A. Leicht. *Proc. Natl. Acad. Sci. U.S.A.*, 104(23), (2007), 9564-9569.
[14] B. Karrer, E. Levina and M. E. J. Newman. *Phys. Rev. E*, 77, (2008), 046119.
[15] G. Bianconi, P. Pin and M. Marsili. *Proc. Natl. Acad. Sci. U.S.A.*, 106(28), (2009), 11433-11438.
[16] A. Lancichinetti, F. Radicchi, J. J. Ramasco and S. Fortunato. *PloS one*, 6(4), (2011), e18961.
[17] Y. Hu., Y. Nie, H. Yang, J. Cheng, Y. Fan and Z. Di. *Europhys Lett*, 82(6), (2010), 066106.
[18] H. J. Li and J. J. Daniels. *Phys. Rev. E*, 91(1), (2015), 012801.
[19] W. Zachary. *J. Anthropol. Res.*, (1977), 452-473.
[20] H. J. Li, Y. Wang, L. Y. Wu, J. Zhang and X. S. Zhang. *Phys. Rev. E*, 86(1), (2012), 016109.
[21] H. J. Li and X. S. Zhang. *Europhys Lett*, 103(5), (2013), 58002.
[22] Verma. D and Meila. M. *University of Washington Tech Rep UWCSE*, 1(1-18), (2003), 030501.
[23] Ng. A Y, Jordan. M. I and Weiss. Y. *Advances in neural information processing systems*, 2, (2002), 849-856.
[24] S. Zhang, R. S. Wang and X. S. Zhang. *Physica A*, 374(1), (2007), 483-490.
[25] J. D. Wilson, S. Wang, P. J. Mucha, S. Bhamidi and A. B. Nobel. *Ann Appl Stat*, 8(3), (2013), 1853-1891.
[26] M. Girvan and M. E. J. Newman. *Proc. Natl. Acad. Sci. U.S.A.*, 99(12), (2002), 7821-7826.
[27] M. J. Barber and J. W. Clark. *Phys. Rev. E*, 80(2), (2009), 026129.
[28] Please download the supplementary material file from the website http://doc.aporc.org/wiki/Robustness