Community detection based on significance optimization in complex networks

Ju Xiang$^{1,5}$, Zhi-Zhong Wang$^2$, Hui-Jia Li$^{3,6}$, Yan Zhang$^4$, Fang Li$^{1,5}$, Li-Ping Dong$^1$, Jian-Ming Li$^{1,6}$ and Li-Juan Guo$^{5,6}$

1 Department of Anatomy, Histology and Embryology, Neuroscience Research Center, Changsha Medical University, Changsha, Hunan 410219, People’s Republic of China
2 South City College, Hunan First Normal University, Changsha, Hunan 410205, People’s Republic of China
3 School of Management Science and Engineering, Central University of Finance and Economics, Beijing 100080, People’s Republic of China
4 Department of Computer Science, Changsha Medical University, Changsha, Hunan 410219, People’s Republic of China
5 Department of Basic Medical Sciences, Changsha Medical University, Changsha, Hunan 410219, People’s Republic of China
E-mail: hjli@amss.ac.cn (HJL), ljming0901@sina.com (JML), lisa690544@163.com (LJG), xiang.ju@foxmail.com and xiangju@aliyun.com (JX)

Received 25 January 2017, revised 2 April 2017
Accepted for publication 3 April 2017
Published 25 May 2017

Online at stacks.iop.org/JSTAT/2017/053213
https://doi.org/10.1088/1742-5468/aa6b2c

Abstract. Community structure is an important topological property that extensively exists in various complex networks. In the past decade, much attention has been paid to the design of community-detection methods, while analyzing the behaviors of the methods is also of interest in theoretical research and real applications. Here, we focus on an important measure for community structure, i.e. significance (2013 Sci. Rep. 3 2930). Specifically, we study the effect of various network parameters on this measure, analyze the critical behaviors in partition transition, and then deduce the formula of the critical points and the phase diagrams theoretically. The results show that the critical number of communities in partition transition increases dramatically with the difference between inter-community and intra-community link densities, and

$^6$ Authors to whom any correspondence should be addressed.
thus significance optimization displays higher resolution in community detection than many other methods, but it also may lead to the excessive splitting of communities. By employing the Louvain algorithm to optimize the significance, we confirm the theoretical results on artificial and real-world networks, and further perform a series of comparisons with some classical methods.

**Keywords:** random graphs, networks, clustering techniques

---

1. **Introduction**

Complex networks are effective to understand the structure and function of various complex systems in real world, such as the metabolic networks and protein-protein interaction networks [1]. In the past decade, many common topological properties of complex networks were discovered and investigated, e.g. clustering, degree correlation and community structure [1, 2], which implies the existence of possible organization principles in the systems. The appearance of community structure means that the complex networks generally consist of groups of vertices within dense inner connections and sparse external connections, called communities or modules [1]. Community structure in complex networks is closely related to real functional grouping in real-world systems [3, 4] and it can affect such dynamic processes as information diffusions and synchronizations [5, 6]. For example, Yan et al recently found that local targeted immunization outperforms global targeted immunization, if apparent community structure exists in a network [7]; and Wu et al have shown that the abundance of communities in the

https://doi.org/10.1088/1742-5468/aa6b2c
social network can evidently foster the formation of cooperation under strong selection [8]. Because of the importance of community structure, a large number of methods have been proposed to detect the communities in complex networks based on various approaches, including spectral analysis [9, 10], random walk dynamics [11–13], phase dynamics [14], diffusion dynamics [15], label propagation [16–18], statistical models [19, 20], structural perturbation [21] and Modularity optimization [22–24] (see [1, 25, 26] for reviews).

Although much attention has been paid to the design of community-detection methods, there is only a few works designed to analyze the intrinsic behaviors of the methods. Studying these behaviors is also of interest in both the theoretical research and real applications. On the one hand, it could be helpful for understanding the methods themselves, because the methods also have the scope of application themselves, though they are helpful for detecting and analyzing the structures of complex networks. On the other hand, it could promote the improvement of the methods or the development of more effective methods. For example, in methods based on Modularity optimization and Bayesian inference it was found that phase transitions from detectable to undetectable structures exist in community detection, which provides a bound on the achievable performance of the methods [27–29]. Botta et al presented a detailed analysis of Modularity density, showing its superiority and drawbacks [30]. The limits of Modularity, such as the resolution limit [31–33], implied the possible existence of multi-scale structures in networks, and promoted the proposal of various (improved) methods, especially the multi-resolution Modularity or Hamiltonians [34–38]. Various approaches have been used to improve the performance of Modularity-based methods [39–41]. Lai et al proposed the improved Modularity-based method by random walk network preprocessing [40], and then enhanced the Modularity-based belief propagation method by using the correlation between communities to improve the estimate of number of communities [41]. Chakraborty et al proposed a new post-processing technique by which many existing community-detection methods for hard partitions can be extended to soft partitions, based on the resemblance between identified non-overlapping and actual overlapping community structure [42].

Optimizing quality functions for community structures is very popular for community detection, such as Modularity [22–24, 43–45], Hamiltonians [19], and ‘fitness’ functions [46, 47]. In [48], Traag et al proposed an important measure for community detection, called significance. It can be used to estimate the quality of community structures, by looking at how likely dense communities appear in random networks, and is defined as,

$$S = \sum_s \left( \frac{n_s}{2} \right) D(p_s || p)$$

$$= \sum_s \frac{n_s(n_s - 1)}{2} \left( p_s \ln \frac{p_s}{p} + (1 - p_s) \ln \frac{1 - p_s}{1 - p} \right),$$

(1)

where the sum runs over all communities; $n_s$ represents the number of vertices in community $s$; $p_s$ represents the ratio of the number of existing edges to the maximum in the community; $p$ represents the ratio of the number of existing edges to the maximum in the whole network. This measure was initially proposed to determine the significant
scale of community structures, which also could be optimized as the objective function to find the optimal community partitions [48].

In this paper, we will analyze the effect of various network parameters on the significance measure, study the critical behaviors on partition transition, and give the formula of the critical points and the phase diagrams theoretically. By employing the Louvain algorithm to optimize the significance, we verify the theoretical results on artificial and real-world networks, and perform a series of comparisons with classical methods, including Infomap, Walktrap, OSLOM, LP and Modularity. Finally, we come to a conclusion.

2. Method

In this section, firstly we introduce a set of model networks (see figure 1) and the detailed expression of significance in the networks, and then analyze the relation between significance and various network parameters. Finally the critical behaviors of significance in partition transition is investigated in detail, and the formula of the critical points and the phase diagram is deduced theoretically.

2.1. Definition of model networks

For the convenience of theoretical analysis, we constructed a set of community-loop model networks with r communities connected one by one (see figure 1). For the pre-defined original community partition in the networks, which contains r communities with $n_c$ vertices, the value of significance reads,

$$S_{\text{origin}} = r \left( \frac{n_c}{2} \right) D(p_1||p)$$

$$\approx \frac{r \cdot n_c^2}{2} \left( p_1 \ln \frac{p_1}{p} + (1-p_1) \ln \frac{1-p_1}{1-p} \right)$$

$$\approx \frac{r \cdot n_c^2}{2} \left( p_1 \ln \frac{r \cdot p_i}{p_i + 2p_o} + (1-p_1) \ln(1-p_i) \right), \quad (2)$$

where $p_1 = p_i$, $p = (p_i + 2p_o)/r$, and $1-p \approx 1$ generally.
In order to analyze the critical behaviors of \textit{significance} in partition transition, we consider a kind of partitions that consists of \( r/2 \) groups, where each group contains two adjacent communities and thus has 2\( n_c \) vertices. Therefore, the value of \textit{significance} for the partition with community merging reads,

\[
S_{\text{merge}} = \frac{r}{2} \left( \frac{2n_c}{2} \right) D(p_2||p) \\
\approx r \cdot n_c^2 \left( p_2 \ln \frac{p_2}{p} + (1 - p_2) \ln \frac{1 - p_2}{1 - p} \right) \\
\approx r \cdot n_c^2 \left( \frac{p_i + p_o}{2} \ln \frac{r \cdot \frac{p_i + p_o}{2}}{p_i + 2p_o} + \left(1 - \frac{p_i + p_o}{2}\right) \ln \left(1 - \frac{p_i + p_o}{2}\right) \right),
\]

where \( p_2 = (p_i + p_o)/2 \) and \( p = (p_i + 2p_o)/r \).

\[ \text{(3)} \]

\subsection*{2.2. Relationship between significance and network parameters}

For the sake of visual illustration, figures 2 and 3 plot the curves of \textit{significance} with various network parameters, though equations (2) and (3) contain the relations between \textit{significance} and network parameters.

Firstly, suppose \( 1 - p \approx 1 \) for large \( r \)-values, thus \( S \propto r \ln r \) for \( S_{\text{origin}} \) and \( S_{\text{merge}} \). Figures 2 and 3 also clearly show that the values of \textit{significance}, normalized by \( r \cdot n_c^2 \), are linearly increasing with \( r \).
Secondly, for $S_{\text{origin}}$, the slopes of the curves are affected only by the inner-community link probability $p_i$, while $p_o$ only affects the intercepts of the curves (note the intercepts are also affected by $p_i$). So we can see the family of curves for different $p_i$, which are a series of parallel straight lines for different $p_o$ (see figure 2(a)).

Thirdly, for $S_{\text{merge}}$, the slopes of the curves are affected by $p_i$ and $p_o$, while the intercepts are also. So we see the family of curves for different $p_i$, which contains the straight lines with different slopes and intercepts for different $p_o$ (see figure 2(b)).

Finally, by comparing the curves for different $p_i$ (see figure 2), $S_{\text{origin}}$ and $S_{\text{merge}}$ increase with the increase of $p_i$. By comparing the curves for different $p_o$, $S_{\text{origin}}$ decreases with the increase of $p_o$, while $S_{\text{merge}}$ increases with the increase of $p_o$. Figure 3 further displays the conclusions more clearly. It also implies that the larger the $p_i$-values, the more difficult the community merging, since the needed $p_o$-values will be larger.

2.3. Critical behaviors in partition transition

In the section, we study the transition from the predefined partition to the partition with community merging. When $S_{\text{merge}} - S_{\text{origin}} > 0$, the identified partition should change to be the above partition with community merging while not the pre-defined original partition. Figure 4 shows that $S_{\text{merge}}/S_{\text{origin}}$ will be greater than 1, when the number $r$ of communities is large enough, and the critical points are different for different $p_o$-values (e.g. from 0.0 to 1.0). As we see that, the smaller the $p_o$-values, the larger the needed $r$-values, meaning that the community merging is more difficult. For smaller $p_i$ (e.g. $p_i = 0.6$), the needed $r$-values decrease correspondingly.

On the basis of the above qualitative analysis, in the following, we give the analytic expression of the critical points in the partition transition. By equations (2) and (3), the critical condition in the transition reads,

$$
\left(\binom{2n_c}{2} D(p_2||p) - 2 \binom{n_c}{2} D(p_1||p)\right) \geq 0.
$$

Figure 3. For the original partition (Orig) and the partition with community merging (Merg), curves of significance as a function of $p_o$ for different $p_i$: (a) $r = 100$ and (b) $r = 1000$. 
Theorem. By solving equation (4) for \( r \), the critical number of communities for significance in the partition transition, reads,

\[
 r_{\text{critical}} = p' \cdot \exp \left( \frac{2H(p_2) - H(p_1)}{2p_2 - p_1} \right) \\
= p' \cdot \exp \left( \frac{H(p_1) + 2\Delta H}{p_1 + 2\Delta p} \right) \\
= p' \cdot \exp \left( \frac{1 + 2\Delta H/H(p_1)}{1 + 2\Delta p/p_1} \cdot \frac{H(p_1)}{p_1} \right), \tag{5}
\]

where the information entropy \( H(x) = -x \ln(x) - (1 - x) \ln(1 - x) \), \( \Delta H = H(p_2) - H(p_1) \), \( \Delta p = p_2 - p_1 \) and \( p' = p_1 + 2p_o \). The critical number of communities in the transition is closely related to the changes of the information entropy caused by the inner-link probability in the communities. Moreover, for Modularity, \( r_{\text{critical}} = p_1/p_o + 2 \).

Proof. Suppose that \( n_c - 1 \approx n_c, 2n_c - 1 \approx 2n_c, 1 - p \approx 1 \) for large \( r \)-values, and define \( p = p'/r \). By equation (4),

\[
0 = \frac{2n_c(2n_c - 1)}{2} D(p_2||p) - \frac{2n_c(n_c - 1)}{2} D(p_1||p)
\]

\[
0 = 2 \left( p_2 \ln \left( \frac{p_2}{p'} \right) + (1 - p_2) \ln (1 - p_2) \right) - \left( p_1 \ln \left( \frac{p_1}{p'} \right) + (1 - p_1) \ln (1 - p_1) \right)
\]

\[
(2p_2 - p_1) \ln(r) = p_1 \ln(p_1) + (1 - p_1) \ln(1 - p_1) - 2(2p_2 \ln p_2 + (1 - p_2) \ln(1 - p_2)) + 2(p_2 - p_1) \ln \left( \frac{1}{2p_2 - p_1} \right)
\]

\[
r = \exp \left( \frac{1}{2p_2 - p_1} \left( p_1 \ln \left( \frac{p_2}{p'} \right) + (1 - p_1) \ln (1 - p_1) - 2 \left( p_2 \ln \left( \frac{p_2}{p'} \right) + (1 - p_2) \ln (1 - p_2) \right) \right) + \ln(p') \right) = p' \cdot \exp \left( \frac{2H(P_2) - H(P_1)}{2p_2 - p_1} \right)
\]

\[
(6)
\]
For illustration, figure 5 displays the relation between $r_{\text{critical}}$ and network parameters. We can see that $r_{\text{critical}}$ decreases with the increase of $p_o$. This is reasonable, because the increase of the number of links between communities will make the communities merging easier. For large $p_o$-values, the $r_{\text{critical}}$-values are very small, which are close to the critical values of Modularity. However, for small $p_o$-values, the $r_{\text{critical}}$-values dramatically increase with the decrease of $p_o$, which is far greater than that of Modularity. As a result, significance generally tends to split the communities in the networks, especially with small inter-community link density, and find more communities than other methods, such as Modularity. This is verified by the experimental results in the next section.

Moreover, we see that for fixing $p_o$-values, the larger the $p_i$-values, the larger the $r_{\text{critical}}$-values (see figure 5(b)). This means that the denser the links inside communities, the more difficult the communities merging. On the whole, the difference between inter-community and intra-community link density is easily to result in the disconnecting of communities. The slight link-density heterogeneity in community is also possible to lead to the split of the community. In some cases, some high link-density regions may be separated from the communities in the networks.

3. Experimental results

In this section, we provide a series of comparisons of significance with some classical methods such as Modularity [24], Infomap [11], Walktrap [49], OSLOM [50] and LP [51] on artificial and real-world networks.

3.1. Artificial networks

Firstly, we identify the communities in the above community-loop networks by using the Louvain algorithm to optimize significance. Table 1 shows that (1) when the
number of pre-defined communities \( r \) is large enough, communities merging will appear, e.g. for \( p_i = 1.0 \) and \( p_o = 0.4 \); (2) when \( p_o \) is large enough, communities merging will appear, e.g. for \( p_i = 1.0 \) and \( r = 128 \); (3) the decrease of \( p_i \) makes communities more easily merge. Table 2 shows similar results for Modularity in the same networks, but Modularity is easier to merge the communities in the networks than significance. These results are consistent with the above theoretical analysis.

Figure 6 compares the accumulative number of identified communities by different methods in the community-loop networks with different parameters. It confirms that significance can identify more communities than other methods. Or say, significance has higher resolution in community detection.

However, the high resolution of significance may lead to another problem—the excessive splitting of communities. In some cases, it may not be able to identify the community structures, which can be identified by some classical methods. We test a set of examples for this problem. Table 3 shows the ratio of the number of communities identified by different methods to the number of predefined communities, in the LFR networks [52]. With the decrease of the mean degree \( k_m \) in the networks, the split of communities is getting worse, because of the increase of heterogeneity inside communities.
3.2. Real-world networks

Finally, we apply the above methods to a set of real-world networks. In the real-world networks, it is difficult to directly compare the performance of different methods. In table 4, therefore, we list the number of communities identified by different methods. The results show that significance intensively splits the networks into communities. This confirms that significance also tends to generate more communities in the general real-world networks than other methods.

4. Discussion

In general, significance has higher resolution in community detection than many other methods, such as Modularity, while it also may result in the excessive splitting of communities. The excessive splitting of communities often originates from the inhomogeneity of link density. The inhomogeneity of link density may be due to the noise of data (false positive and false negative data), but it also is possible to imply the existence of inner substructures, which can be identified by high-resolution methods. For the first case, it will be a kind of effective approach to reweight the links of network or reconstruct the topology of network. In recent years, the link-rewriting strategy has...
been proved to be able to effectively enhance the performance of community-detection methods [39, 40]. While, compared to the link-reweighting strategy, the network reconstruction will be more suitable for significance, because it does not require the weighted version of the methods. Systematically studying the effect of the network reconstruction on community detection will be an interesting topic.

In fact, it generally is difficult to discriminate between the two cases of the excessive splitting: the mistake caused by the noise of data and the existence of inner substructures, e.g. multi-scale community structures. It is believed that multi-scale structures widely exist in various complex networks. Different methods may identify different-scale structures. For example, in the H13-4 network with community structures at two scales (see figure 7), Modularity identifies the first-level community structure, while significance identifies the second-level one, which also means that the first-level

### Table 4. The number of communities in various real-world networks, identified by different methods.

| Networks       | Modularity | Infomap | Walktrap | OSLOM | LP  | Significance |
|----------------|------------|---------|----------|-------|-----|--------------|
| Dolphin [53]   | 5          | 6       | 5        | 2     | 10  | 22           |
| Polbooksa      | 5          | 5       | 4        | 2     | 6   | 28           |
| Football [54]  | 9          | 12      | 10       | 11    | 11  | 15           |
| Jazz [55]      | 4          | 6       | 8        | 5     | 3   | 36           |
| C. Elegans neural [56] | 5   | 8       | 22       | 3     | 2   | 67           |
| Email [57]     | 11         | 63      | 47       | 24    | 9   | 233          |

V Krebs, www.orgnet.com/ (from www-personal.umich.edu/~mejn/netdata/).

![Figure 7. H13-4, a network with 256 vertices and two scale community structures](https://doi.org/10.1088/1742-5468/aa6b2c)
structure is ignored by significance. Interestingly, by introducing a tunable resolution parameter into Modularity, it can find the communities at two scales, while significance is just a single-scale method. It is an interesting topic to extend the existing single-scale methods (e.g., significance) to the multi-scale case.

Multiplex network is a type of multilayer network where all layers may have different or similar topology and share some inter-layer links. The structure of the multiplex networks allows for vertices between two adjacent layers to be connected by inter-layer links. These links represent some types of associations between layers, and can be specified either by using information extracted directly from the multiplex or external data. There is a wide class of real networks that can be represented by a multiplex network [58].

In recent years, a handful of algorithms have been described in the literature to address the problem of finding community structures in multiplex networks. A straightforward approach is that we apply a community detection algorithm to each separate layer, and then combine all the resulting partitions either by using cluster ensemble approaches or by merging communities across all layers. However, most quality function optimization methods, such as Modularity, can not be directly applied to the whole multiplex network, because modularity optimization may produce some ‘inter-layer’ communities which may cross two or more layers (see figure 8). This is not reasonable and can not be explained easily in the real situations.

One can avoid this ‘inter-layer’ problem by providing new algorithms to produce more ‘finer’ communities, since finer community will belong to only exactly one layer with a higher probability. In this paper, we analyze the result of community detection based on the significance measure, and it is easily to lead to the excessive splitting of communities, which thus may naturally avoid the ‘inter-layer’ community problem.

Furthermore, for the self-similar structures, empirical studies show that some networks display the property that high- or low-degree nodes tend to connect to other nodes with similar degrees, which is referred to as ‘assortative mixing’ [59]. This assortative mixing phenomenon is usually associated with the hierarchical structure, which leads to a strong community structure—the higher degree nodes locate on the core position of the community and the lower degree nodes locate on the border. It is interesting to reveal the effect of the property on the behaviors of community detection methods.

Figure 8. The red community is ‘inter-layer’ one, and the blue and green communities are finer ones.
5. Conclusion

Community structure extensively exists in various complex networks. Detecting communities (or modules) in complex networks is very important for the research of complex networks. In the past decades, much attention was paid to the development of methods for community detection in complex networks. However, the detailed analysis of the methods’ behaviors is also of interest, which could help in understanding the methods themselves, and promote the development of more effective methods.

In this paper, we focus on an important measure for estimating the quality of community structures, called significance. It was proposed to initially determine significant scale of community structures, but it can also be used as a target function to search the optimal community partitions. We studied the effect of various network parameters on this measure in detail, analyzed the critical behaviors of it in partition transition, and analytically gave the formula of the critical points and the phase diagram. The results were confirmed on artificial and real-world networks, and a series of comparisons with some classical methods were also given.

The difference between inter- and intra-community link density is crucial to the disconnecting or splitting of communities in networks. The results show that the critical number \( r_{\text{critical}} \) of communities in partition transition is to increase dramatically with the decrease of the inter-community link density for each intra-community link density. When the inter-community link density is very large, the \( r_{\text{critical}} \)-value is very small, which is close to but still larger than that of Modularity, but when the inter-community link density becomes small, the \( r_{\text{critical}} \)-value quickly increases, and is far greater than that of Modularity.

On the whole, significance tends to split the communities in the networks, and find more communities than other methods, such as Modularity. So it generally has higher resolution in community detection than many other methods, but it also may lead to the problem of excessive splitting of communities. In some cases, the low link-density heterogeneity in a community is also possible to lead to the split of the community. It is still an open issue how to find the appropriate balance between the high resolution and excessive splitting in community detection.

Finally, we expect that the above detailed analysis could be helpful for the understanding of the behaviors of significance in community detection and provide useful insight into designing effective methods for detecting communities in complex networks.

Acknowledgment

This work was supported by the construct program of the key discipline in Hunan province, the Scientific Research Fund of Education Department of Hunan Province (Grant No. 17A024, 17B034 and 17C0180), the Scientific Research Project of Hunan Provincial Health and Family Planning Commission of China (Grant No. C2017013), the Scientific Research Fund of Education Department of Hunan Province (Grant Nos.14C0112 and 14B024), the Department of Education of Hunan Province (Grant No. 15A023), the Hunan Provincial Natural Science Foundation of China (Grant No. 2015JJ6010), the Hunan Provincial Natural Science Foundation of China (Grant No. 13JJ4045), the National Natural Science Foundation of China (Grant No. 71401194 and 1140178) and Young Elite Teacher Project of Central University of Finance and Economics (Grant No. QYP1603).

https://doi.org/10.1088/1742-5468/aa6b2c
References

[1] Fortunato S 2010 Community detection in graphs Phys. Rep. 486 75
[2] Xiang J, Hu K, Zhang Y, Hu T and Li J M 2015 Analysis and perturbation of degree correlation in complex networks Europhys. Lett. 111 6
[3] Chen P and Redner S 2010 Community structure of the physical review citation network J. Informetrics 4 278
[4] Zhang S-H, Ning X-M, Ding C and Zhang X-S 2010 Determining modular organization of protein interaction networks by maximizing modularity density BMC Syst. Biol. 4 1
[5] Nematzadeh A, Ferrara E, Flammini A and Ahn Y-Y 2014 Optimal network modularity for information diffusion Phys. Rev. Lett. 113 088701
[6] Ren G and Wang X 2014 Epidemic spreading in time-varying community networks Chaos 24 023116
[7] Shu Y, Shao t in T, Wenyi F, Sen P and Zhiming Z 2015 Global and local targeted immunization in networks with community structure J. Stat. Mech. Theor. Exp. 2015 P08010
[8] Wu Z-X, Rong Z and Yang H-X 2015 Impact of heterogeneous activity and community structure on the evolutionary success of cooperators in social networks Phys. Rev. E 91 012802
[9] Cheng J-J, Li L-J, Leng M-W, Lu W-G, Yao Y-K and Chen X-Y 2016 A divisive spectral method for network community detection J. Stat. Mech. 033403
[10] Donetti L and Munoz M A 2004 Detecting network communities: a new systematic and efficient algorithm J. Stat. Mech. 2004 P10012
[11] Rosvall M and Bergstrom C T 2008 Maps of random walks on complex networks reveal community structure Proc. Natl Acad. Sci. USA 105 1118
[12] Jin D, Yang B, Baquero C, Liu D, He D and Liu J A 2011 Markov random walk under constraint for discovering overlapping communities in complex networks J. Stat. Mech. 05301
[13] Delvenne J-C, Yaliraki S N and Barahona M 2010 Stability of graph communities across time scales Proc. Natl Acad. Sci. USA 107 12755
[14] Chen J, Wang H, Wang L and Liu W 2016 A dynamic evolutionary clustering perspective: Community detection in signed networks by reconstructing neighbor sets Physica A 447 482
[15] Cheng X-Q and Shen H-W 2010 Uncovering the community structure associated with the diffusion dynamics on networks J. Stat. Mech. 2010 P04024
[16] Barber M J and Clark J W 2009 Detecting network communities by propagating labels under constraints Phys. Rev. E 80 026129
[17] Raghavan U N, Albert R and Kumara S 2007 Near linear time algorithm to detect community structures in large-scale networks Phys. Rev. E 76 036106
[18] Šubelj L and Bajec M 2012 Ubiquitousness of link-density and link-pattern communities in real-world networks Eur. Phys. J. B 85 1
[19] Reichardt J and Bornholdt S 2006 Statistical mechanics of community detection Phys. Rev. E 74 016110
[20] Karer B and Newman M E J 2011 Stochastic blockmodels and community structure in networks Phys. Rev. E 83 016107
[21] Wang P, Gao L and Ma X 2017 Dynamic community detection based on network structural perturbation and topological similarity J. Stat. Mech. 013401
[22] Agarwal G and Kempe D 2008 Modularity-maximizing graph communities via mathematical programming Eur. Phys. J. B 66 409
[23] Lehmann S and Hansen L K 2007 Deterministic modularity optimization Eur. Phys. J. B 60 83
[24] Blondel V D, Guillaume J-L, Lambiotte R and Lefebvre E 2008 Fast unfolding of communities in large networks J. Stat. Mech. 2008 P10008
[25] Bagrow J P 2008 Evaluating local community methods in networks J. Stat. Mech. 05001
[26] Danon L, Diaz-Guilera A, Duch J and Arenas A 2005 Comparing community structure identification J. Stat. Mech. 2005 P09008
[27] Nadakuditi R R and Newman M E J 2012 Graph spectra and the detectability of community structure in networks Phys. Rev. Lett. 108 188701
[28] Decelle A, Krzakala F, Moore C and Zdeborová L 2011 Inference and phase transitions in the detection of modules in sparse networks Phys. Rev. Lett. 107 065701
[29] Reichardt J and Leone M 2008 (Un)detectable cluster structure in sparse networks Phys. Rev. Lett. 101 078701
[30] Botta F and del Genio C I 2016 Finding network communities using modularity density J. Stat. Mech. 123402

https://doi.org/10.1088/1742-5468/aa6b2c
Community detection based on significance optimization in complex networks

[31] Zhang X S, Wang R S, Wang Y, Wang J, Qiu Y, Wang L and Chen L 2009 Modularity optimization in community detection of complex networks Europhys. Lett. 87 38002
[32] Good B H, de Montjoye Y-A and Clauset A 2010 Performance of modularity maximization in practical contexts Phys. Rev. E 81 046106
[33] Fortunato S and Barthélémy M 2007 Resolution limit in community detection Proc. Natl. Acad. Sci. USA 104 36
[34] Xiang J, Hu T, Zhang Y, Hu K, Li J-M, Xu X-K, Liu C-C and Chen S 2016 Local modularity for community detection in complex networks Physica A 443 451
[35] Xiang J, Tang Y-N, Gao Y-Y, Zhang Y, Deng K, Xu X-K and Hu K 2015 Multi-resolution community detection based on generalized self-loop rescaling strategy Physica A 432 127
[36] Ronhovde P and Nussinov Z 2010 Local resolution-limit-free potts model for community detection Phys. Rev. E 81 046114
[37] Traag V A, Van Dooren P and Nesterov Y 2011 Narrow scope for resolution-limit-free community detection Phys. Rev. E 84 016114
[38] Arenas A, Fernández A and Gómez S 2008 Analysis of the structure of complex networks at different resolution levels New J. Phys. 10 053039
[39] Xiang J, Hu K, Zhang Y, Bao M-H, Tang L, Tang Y-N, Gao Y-Y, Li J-M, Chen B and Hu J-B 2016 Enhancing community detection by using local structural information J. Stat. Mech. 033405
[40] Lai D, Lu H and Nardini C 2010 Enhanced modularity-based community detection by random walk network preprocessing Phys. Rev. E 81 066118
[41] Lai D, Shu X and Nardini C 2016 Correlation enhanced modularity-based belief propagation method for community detection in networks J. Stat. Mech. 053301
[42] Chakraborty T 2015 Leveraging disjoint communities for detecting overlapping community structure J. Stat. Mech. 05017
[43] Newman M E J and Girvan M 2004 Finding and evaluating community structure in networks Phys. Rev. E 69 026113
[44] Hu Y, Wu J and Di Z 2009 Enhance the efficiency of heuristic algorithms for maximizing the modularity Q Europhys. Lett. 85 18009
[45] Sun Y, Danila B, Josic K and Bassler K E 2009 Improved community structure detection using a modified fine-tuning strategy Europhys. Lett. 86 28004
[46] Lancichinetti A, Fortunato S and Kertész J 2009 Detecting the overlapping and hierarchical community structure in complex networks New J. Phys. 11 033015
[47] Havemann F, Heinz M, Struck A and Gläser J 2011 Identification of overlapping communities and their hierarchy by locally calculating community-changing resolution levels J. Stat. Mech. 01023
[48] Traag V A, Kring G and Van Dooren P 2013 Significant scales in community structure Sci. Rep. 3 2930
[49] Pons P and Latapy M 2005 Computer and Information Sciences—ISCIS 2005 vol 3733, ed P Yolum— (Berlin: Springer) p 284
[50] Lancichinetti A, Radicchi F, Ramasco J J and Fortunato S 2011 Finding statistically significant communities in networks PLoS One 6 e18961
[51] Steve G 2010 Finding overlapping communities in networks by label propagation New J. Phys. 12 103018
[52] Lancichinetti A, Fortunato S and Radicchi F 2008 Benchmark graphs for testing community detection algorithms Phys. Rev. E 78 046110
[53] Lusseau D, Schneider K, Boisseau O J, Haase P, Slooten E and Dawson S M 2003 The bottlenose dolphin community of doubtful sound features a large proportion of long-lasting associations Behav. Ecol. Sociobiol. 54 396
[54] Girvan M and Newman M E J 2002 Community structure in social and biological networks Proc. Natl Acad. Sci. USA 99 7821
[55] Gleiser P and Danon L 2003 Community structure in Jazz Adv. Complex Syst. 6 565
[56] Watts D and Strogatz S 1998 Collective dynamics of ‘small-world’ networks Nature 393 440
[57] Guimerà R, Danon L, Díaz-Guilera A, Giralt F and Arenas A 2003 Self-similar community structure in a network of human interactions Phys. Rev. E 68 065103
[58] Hu Y, Havlin S and Makse H A 2014 Conditions for viral influence spreading through multiplex correlated social networks Phys. Rev. X 4 021031
[59] Gallos L K, Song C and Makse H A 2008 Scaling of degree correlations and its influence on diffusion in scale-free networks Phys. Rev. Lett. 100 248701
[60] Arenas A, Díaz-Guilera A and Pérez-Vicente C J 2006 Synchronization reveals topological scales in complex networks Phys. Rev. Lett. 96 114102