Community detection methods can discover better structural clusters than ground-truth communities

Vinh-Loc Dao, Cécile Bothorel, Philippe Lenca

▶ **To cite this version:**

Vinh-Loc Dao, Cécile Bothorel, Philippe Lenca. Community detection methods can discover better structural clusters than ground-truth communities. 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, Jul 2017, Sydney, Australia. 10.1145/3110025.3110053 . hal-01577343

HAL Id: hal-01577343
https://hal.science/hal-01577343

Submitted on 25 Aug 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Community detection methods can discover better structural clusters than ground-truth communities

Vinh-Loc Dao, Cécile Bothorel and Philippe Lenca

***Abstract*— Community detection emerged as an important exploratory task in complex networks analysis across many scientific domains. Many methods have been proposed to solve this problem, each one with its own mechanism and sometimes with a different notion of community. In this article, we bring most common methods in the literature together in a comparative approach and reveal their performances in both real-world networks and synthetic networks. Surprisingly, many of those methods discovered better communities than the declared ground-truth communities in terms of some topological goodness features, even on benchmarking networks with built-in communities. We illustrate different structural characteristics that these methods could identify in order to support users to choose an appropriate method according to their specific requirements on different structural qualities.**

I. INTRODUCTION

Community detection is a fundamental task in the exploration of networks and has received an exceptional attention in complex network analysis in recent years [1], [2], [3]. It is widely accepted that nodes in networks have a tendency to connect preferably with the similar ones in order to establish operational groups, which are sometimes called clusters, modules or communities. Understanding these community structures of networks plays an essential role in the study of their functionality in many domains such as social network analysis [4], biochemistry [5], communication [6], etc.

The word community in real life represents semantic concepts, which define groups of individuals such as groups of friends in social networks, biological modules in protein-protein interaction networks or classes of malicious web domains in world wide web. In this paper, we call these communities *metadata communities*. On the other hand, in network theory, a *community* implies a group of nodes in a network that can be detected using some topological conditions, which are usually and widely based on edge density criteria. The confusion between these two independent definitions in two different contexts could lead to serious misconstructions in the study of complex networks. Here, when we refer to the word *community*, with an exception of metadata ones, we imply the latter sense of community, i.e. clusters discovered by algorithms.

Many community detection algorithms with different approaches have been proposed in the last decades [7], [8], [9]. Each approach has a different partitioning strategy and sometimes different points of view about topological requirements of a good community. Consequently, communities detected on a same network by different methods are usually different in their structural patterns and sometimes release unpredictable characteristics. That is the reason why the analysis of community discovery method behaviors still faces many challenges.

In networks where metadata communities are available, the performance of community detection methods is usually evaluated based on the similarity between communities that they discover with metadata communities considered as *ground-truth*. The most frequently used similarity metrics that are worth mentioning include Normalized Mutual Information (NMI), Recall score, Precision score, F1-measure, *etc.*. However, declared metadata communities are not always good references in terms of topology. For that reason, considering them as ground-truth can lead to irrelevant conclusions about the properties of topological community detection methods and eliminates good structures that are not similar to metadata communities. A meticulous distinction between these two notions has been pointed out in a recent research by Aaron Clauset *et al.* [10].

In this paper, we present a comparative analysis of most common community detection methods in the literature by studying a wide range of structural characteristics of communities that they produce. In this way, a community is appraised by many criteria that correspond to different goodness notions. Although the analysis that we conduct in this paper uses metadata as a point of reference, the goodness of a community is only depended on its structure and not the similarity with its associated metadata groups.

Then, we show that metadata communities in real-world networks are not always good in terms of some structural qualities, whereas using them as groud-truth is commonly acknowledged. More importantly, we prove that there exists communities in benchmarking networks that are even structurally better than planted communities, which are widely believed as ground-truth and represent the best solution that could be found in synthetic networks.

The rest of this paper is organized as follows: Section II introduces some researches relating to the evaluation of community quality and community discovery methods. In Section III, we describe the dataset that is used in this work and community detection methods under consideration. Section IV focuses on the analysis process using goodness metrics and important results obtained from this analysis. Finally, we conclude this paper and point out some essential contributions and perspectives in Section V.

IMT Atlantique, UMR 6285 Lab-STICC, Université Bretagne Loire, F-29238 Brest, France. {vinh.dao, cecile.bothorel, philippe.lenca}@imt-atlantique.fr

II. RELATED WORK

Leskovec *et al.* [11] analyze the performance of many scoring functions based on their capacities to identify some community goodness characteristics. The authors also include many real life stimulation strategies on network communities such as node swapping, member replacing, community expanding, community shrinking in order to test their consistency through perturbations.

Fortunato *et al.* published an analysis [12] where they compare metadata communities with structural communities throughout a dozen of community detection methods. They use NMI metric to assess the performance of these methods through correlation score matrices. Besides, hierarchical structures in communities are also considered in this analysis.

Rossetti *et al.* [13] introduce a novel approach to evaluate the efficiency of algorithms by comparing detected communities with a given ground-truth information. They use precision score, recall score and their harmonic mean, known as F1-measure, to assess the quality of community discovery methods using scatter plots. This approach is close to the descriptive approach [14] where authors use combinations of quality metrics to point out and expose structures of real-world communities in a comprehensive way.

Cherifi *et al.* provide evidence of the distinctness between traditional detection performance in terms of similarity metrics (Rand Index, NMI, etc.) and topological goodness quality. They indicate that high performance scores do not necessarily imply good structures [15].

III. NETWORKS DATASET AND COMMUNITY DETECTION METHODS

In this paper, we rely on the hypothesis that there exist communities whose structures are better than ground-truth or real communities. To avoid any further confusion, semantic communities in real-world networks or planted communities in synthetic networks are from now called metadata communities.

A. Network dataset

Since our first priority is to compare as many methods as possible, in this paper we only investigate undirected and unweighted networks. Table I describes the dataset that we use in this paper including very well-known real-world networks in the first part and synthetic networks with built-in communities from LFR benchmark [17] in the second part. These synthetic networks are created in a way that their structural parameters approach those of real-world networks. Such that node degrees follow the power-law distribution [18] with exponent coefficients around -2.5. Average node degrees are set from 5 to 20 to acquire relatively sparse networks. Besides, in order to obtain a variety of community quality, the lfr1-lfr5 networks are configured with mixing parameters $\mu \in \{0.1, 0.2, 0.3, 0.4, 0.5\}$ which represents the probability that an edge of a node is connected to nodes outside of its community. Moreover, built-in communities can be disjointed or overlapped with different average participation rates A from 1.0 to 2.6 groups per node.

Graph	N	E	d	$\widehat{k}$	$\bar{\alpha}$	CCF
zachary[1]	34	78	5	4.59	-2.16	0.256
football[1]	115	613	4	10.67	-9.09	0.407
polblog[1]	1222	16714	8	27.36	-3.67	0.226
youtube[2*]	39841	224235	15	11.26	-2.78	0.063
livejournal[2*]	84438	1521988	27	36.05	-2.35	0.773
dblp[2]	317080	1049866	23	6.63	-3.26	0.306
amazon[2]	334863	925872	47	5.53	-3.59	0.205
lfr1	5000	26836	7	10.73	-2.98	0.192
lfr2	10000	24617	18	4.92	-3.08	0.312
lfr3	25000	133429	6	10.67	-3.05	0.024
lfr4	100000	480978	14	9.62	-2.51	0.178
lfr5	100000	1056963	6	21.14	-2.50	0.055

[1] *http://www-personal.umich.edu/ mejn/netdata/*
[2] *https.//snap.stanford.edu/data/* and **subgraph of 5000 communities*
N number of nodes, E number of edges, d network diameter, $\widehat{k}$ average degree of nodes, $\bar{\alpha}$ estimated power law exponent of node degree sequence [18], CCF clustering coefficient

TABLE I
A DESCRIPTION OF DATASET USED IN THIS ANALYSIS

The metadata communities of the previous presented networks are described in table II. Metadata groups in real-world networks are considered as communities such as: those in amazon represent product categories on Amazon website, while metadata communities in dblp are publication venues. These communities can be overlapped since a product on Amazon can belong to several categories and an author in DBLP can publish at several publication venues.

Graph	C	S	A	$\bar{\beta}$	Metadata
zachary	2	17.00	1.00	NA	Group memberships
football	12	9.58	1.00	NA	Team groups
polblog	2	611	1.00	NA	Political alignments
youtube	5000	14.59	1.83	-2.21	Subscription groups
livejournal	5000	27.80	1.65	-2.83	Group memberships
amazon	75149	30.23	6.78	-2.08	Product categories
dblp	13477	53.41	2.27	-3.06	Publication venues
lfr1	515	13.59	1.40	-3.08	Planted groups
lfr2	1473	6.79	1.00	-3.29	Planted groups
lfr3	3002	21.65	2.60	-2.09	Planted groups
lfr4	9434	13.14	1.24	-2.50	Planted groups
lfr5	4729	21.15	1.00	-2.57	Planted groups

C number of communities, S average community size, A community membership per node, $\bar{\beta}$ estimated power law exponent of community size distribution [18].

TABLE II
A DESCRIPTION OF TOPOLOGICAL FEATURES OF METADATA GROUPS

B. Community detection methods

We include a wide spectrum of popular methods in the literature which can handle the networks in a reasonable time. Except for further mentions, all the analysis based on the following methods are conducted with default parameters provided from their authors.

1) Structure-based approaches:

- Cfinder [19] searches for groups of nodes which are constituted by sets of k-cliques that can be reached from each other through a sequence of adjacent k-cliques. We set $k = 3$ in this analysis (referred as cfinder3).

2) Dynamics-based approaches:

- Conclude [20] weights edges of the input network by using a random walker and then calculates distances between each pair of connected nodes. Finally, it searches out communities based on these distances.
- Infomap [21] turns the community detection task into the problem of compressing the description length of a random walker's trajectory so that one can recover as closely as possible the original structure of the network.
- Walktrap [22] measures similarities of vertices based on random walks and then computes community structures in an agglomerative process. This approach is based on the intuition that random walkers tend to get trapped into densely parts corresponding to communities.

3) Diffusion-based approaches:

- Label [23]: The label propagation method is based on the idea that every node is likely to belong to the community where reside most of its neighbors. Nodes are configured with unique labels at the initial step, then they adopt the labels that most of their neighbors have.
- Copra [24] is an extension of the label propagation method which can include overlapping information as an input parameter. Each vertex can belong up to v communities. In this analysis we set $v = 4$ (referred as copra4).
- Demon [25] fuses overlapping communities that are explored locally by using the label propagation algorithm together based on a merging function. Fusing parameter is chosen at $\varepsilon = 0.25$.
- Ganxis is a general speaker-listener label propagation algorithm based on information propagation process, which spreads labels around nodes according to some interaction rules [26].

4) Optimization-based approaches:

- Fastgreedy is a hierarchical agglomeration algorithm for detecting community structures based on the maximization of modularity increment in each iteration [32].
- Louvain: This method is a multi-step hierarchical modularity optimization [27]. In each step, it aggregates closely connected nodes that yield an optimization of modularity to create a intermediate partition.
- Oslom is based on a local optimization of a fitness function expressing the statistical significance of clusters with respect to random fluctuations [28].

IV. An analysis of structural features of community detection methods

The main objective of this paper is to quantify the behavior of community detection methods in terms of some specific community topological features. In other words, we evaluate the functionality of a method by analyzing structural characteristics of all the communities that it discovers. In order to do that, we apply all community detection methods presented in Section III-B to each network in the dataset and then measure various structural features of identified communities. The obtained values will be then compared with their corresponding values in metadata communities in order to evaluate detection performance. These structural features, which will be introduced in Section IV-A, also represent goodness metrics as they describe various notions of a good community. In this paper, these terms are used interchangeably.

A. Definitions of structural goodness metrics

A graph $G = (V, E)$ is composed of a set of $n = |V|$ nodes and $m = |E|$ edges where $E = (u, v) : u, v \in V$. Given a cluster S of n_S nodes, which is a subgraph of G, a function $g(S)$ quantifies a structural goodness feature of S according to a particular expectation of community quality. Let m_S be the number of edges inside S, $m_S = |(u, v) \in E : u \in S, v \in S|$; c_S be the number of edges that connect S to other nodes outside of S, $c_S = |(u, v) \in E : u \in S, v \notin S|$. Since quality scores of communities are normally correlated [11], some following representative goodness metrics are considered:

- *Separability* [11] is based on the concept that a good community should be well separated by a rupture in edges distribution. This function measures the ratio between internal connections and external connections of nodes inside a community: $g(S) = \frac{m_S}{c_S}$.
- *Embeddedness* [12] reflects how much the direct neighbors of a node belong to its community. It is measured as the ratio of internal degree to the total degree of a community: $g(S) = \frac{2m_S}{2m_S + c_S}$.
- *Density* [11] captures the idea that nodes in a community must be well connected. It quantifies the fraction of edges inside S over the total possible edges could be established in S: $g(S) = \frac{m_S}{n_S(n_S-1)/2}$.
- *Compactness* [29] suggests that good communities should be at the same time dense and easily reachable from nodes to nodes. This quality is calculated by: $g(S) = \frac{m_S}{d(S)}$, where $d(S)$ is the diameter of S.
- *Clustering coefficient* is a very well-known metric [30] which is used to evaluate community quality. It is based on the concept that pairs of nodes with common neighbors are more likely to be connected.
- *Cluster modularity* measures the difference between edges inside S and the expected number of such edges in a random network with the same degree distribution. $g(S) = \frac{m_S}{|E|} - \frac{(m_S + c_S)^2}{|E|^2}$ [1], [2].

B. Properties of metadata communities in terms of structural goodness metrics

As previously mentioned, we employ metadata community goodness scores as a reference for testing the performance of community detection methods. Thus, we are now interested in analyzing the distribution of these scores throughout our set of metadata communities. We compute the goodness scores defined in Section IV-A for all metadata communities in the dataset and present the distribution of these scores in Figure 1 and 2 for real-world communities and synthetic communities respectively. Each subfigure represents a distribution of one goodness metric's scores.

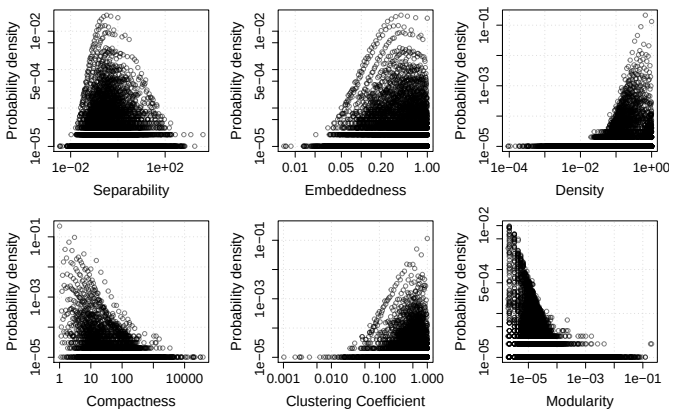

Fig. 1. Distribution of goodness scores of real-world metadata communities

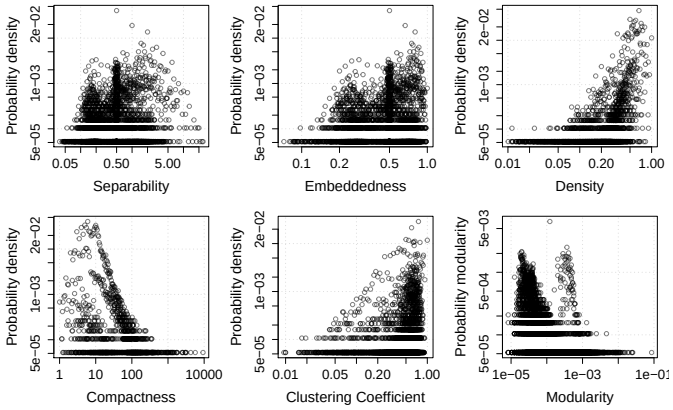

Fig. 2. Distribution of goodness scores of synthetic metadata communities

As we can see, the distributions differ notably from one metric to another in both real-world communities and synthetic communities as they measure different structural aspect of communities. However within a goodness metric, these distributions are quite similar from one metric of quality to another. A possible explanation for this similarity could be that metadata communities in these networks are correlated with structural features, so that goodness scores are distributed in a more organized way than in a topological meaningless partition. Note that by the construction mechanism of LFR benchmark based on a constant mixing parameter μ, there are some ruptures and peaks appeared on the side of synthetic communities [31].

These distributions could also give a prediction about the possibility to discover good structural communities. It is based on a premise that one expect to get partitions with high goodness scores. If goodness score distributions are right-skewed, there is a chance that community detection methods could identify communities whose scores are moved towards the right hand side of the distribution landscapes. Although the performance of community detection algorithms also depends on topological characteristics of input graphs, this preliminary survey provides significant information about metadata structural properties and a potential perspective.

Method	Sep.	Emb.	Den.	Com.	CCF.	Q.
cfinder3	1.68	0.94	**3.97**	1.16	**1.90**	0.44
conclude	1.42	1.13	2.52	0.72	1.33	0.63
infomap	2.24	1.26	3.34	0.96	0.90	0.75
walktrap	1.87	1.19	3.35	0.78	0.93	0.65
label	2.72	1.40	1.84	1.15	1.11	1.06
copra4	1.35	0.87	2.94	1.76	1.14	0.53
demon	4.28	1.26	0.63	**19.28**	1.05	1.44
ganxis	5.34	1.39	2.54	1.19	1.03	0.84
fastgreedy	6.18	1.46	2.79	1.79	0.99	2.71
louvain	**11.01**	**1.50**	2.68	6.02	0.94	**12.67**
oslom	1.69	1.10	1.29	1.21	1.05	0.83
Average	3.62	1.23	2.54	3.27	1.13	2.19

Ratio between average structural goodness scores of detected communities over those of metadata communities calculated based on equation (1). Sep. *separability*, Emb. *embeddedness*, Den. *density*, Com. *compactness*, CCF. *clustering coefficient*, Q. Cluster modularity, Avg. *average quality improvement score*. The best method of each quality feature are bolded.

TABLE III

AVERAGE GOODNESS SCORE RATIOS ON REAL-WORLD NETWORKS

C. An analysis of community detection methods' behaviors based on structural goodness metrics

The principle interest in this section is answering the question: "How good are communities identified by methods presented in Section III-B compared to the metadata communities in terms of structural goodness features defined in Section IV-A?" We quantify a ratio between the average goodness scores of discovered communities and those of metadata communities to evaluate these methods.

Suppose that method M discovers C_M communities in a network, which contains C_O metadata communities. The goodness ratio which represents the improvement of feature F promoted by method M in this dataset is measured by:

$$R(M,F) = \frac{\left[\sum_{i=1}^{C_M} g_F(S_i)\right]/C_M}{\left[\sum_{j=1}^{C_O} g_F(S_j)\right]/C_O} \quad (1)$$

This ratio can vary from zero to infinity. $R(M,F) = 1$ indicates that the method M provides communities that are as good as metadata communities in terms of feature F, while $R(M,F) > 1$ and $R(M,F) < 1$ implies an enhancement and a degradation respectively.

We measure all goodness scores of detected communities and calculate goodness ratios based on Equation (1). The average ratios are showed in Table III and IV for real-world networks and synthetic networks respectively. Each row corresponds to a method and each column corresponds to a goodness metric.

Surprisingly, we observe a significant quality improvement in most methods and goodness metrics, even in synthetic networks where it is widely believed that planted communities are the best, there are still improvements in general. This phenomenon is explainable since LFR benchmark only create communities based on a mixing parameter condition which is not always preferred by all goodness functions. Hence, it is generally possible to get higher goodness scores just by some simple actions such as merging, dividing communities, etc.

Method	Sep.	Emb.	Den.	Com.	CCF.	Q.
cfinder3	0.89	0.89	**1.57**	0.88	**1.63**	0.70
conclude	1.00	0.97	1.26	0.76	1.22	0.77
infomap	1.22	1.08	0.95	0.99	0.98	1.00
walktrap	1.22	1.07	1.27	0.79	1.04	0.83
label	1.20	**1.64**	0.88	0.87	0.87	0.79
copra4	0.89	1.45	0.82	0.74	0.75	0.66
demon	0.76	0.77	0.87	2.20	1.16	2.78
ganxis	1.14	1.04	1.04	0.84	0.96	1.03
fastgreedy	**1.94**	1.22	0.35	**14.97**	0.80	**36.70**
louvain	1.57	1.21	0.09	7.10	0.73	26.62
oslom	1.10	1.04	0.86	1.13	0.97	1.25
Average	1.17	1.13	0.91	2.84	1.01	6.65

The abbreviations are reused from Table III.
The best method of each quality feature are bolded.

TABLE IV
AVERAGE GOODNESS SCORE RATIOS ON SYNTHETIC NETWORKS

However, we can see in Table IV that no method can improve all goodness scores of synthetic communities at the same time. Each method will normally improve some goodness scores while reduce some others.

Importantly, it can be seen that the average improvement of goodness scores in real-world networks is generally higher than that of in synthetic networks. This is completely reasonable since metadata communities in synthetic networks are planted based on many topological conditions while metadata communities in real-world networks are often chosen by semantic meanings, functional criteria or sometimes in a subjective way. As a consequence, there are obviously less correlation between metadata communities with their topology in real-world networks. For this reason, community detection algorithms could ameliorate more remarkably structural goodness scores of real-world metadata groups.

Density is the only goodness metric that shows a global degradation in synthetic networks. Though, it is totally explainable since the link density of a subgraph S is measured by the ratio between the number of links inside S and the total number maximum of links that could be formed, which is $n_S(n_S-1)/2$. While link density increases linearly with the size of S in sparse graphs, the number of possible links increases quadratically. So it is clear that *density* favors small communities in general. Since synthetic community sizes follow power-law distributions with high exponent coefficients, there are plenty of tiny communities in synthetic networks. This explains why *density* is not often improved. However, one can note that cfinder3 always improves significantly the *density* in the two cases as it found a numerous number of small cliques. This clique detection mechanism also makes cfinder3 to be the best to improve clustering coefficient.

Since both metrics *separability* and *embeddedness* are both built on the notion that good communities have relatively higher number of internal edges than number of external edges, one can remark that there is a correlation between these two metrics throughout all methods in both cases.

As presented, *compactness* favors short diameter communities where nodes are easily accessible from one to another. This conception explains why methods that discover locally nodes in networks such as demon, louvain, fastgreedy usually improve significantly this goodness feature.

Modularity is the most atypical among the studied metrics as it is improved more significantly in synthetic networks. Since modularity is designed as an accumulative function, it can be misbehaved by using an average ratio. However, it is not unpredictable to see that modularity optimization-based methods such as fastgreedy and louvain enhance remarkably this metric.

D. A ranking of community detection by structural goodness metrics

In this section, we focus on the ranking of community detection methods in function of the above goodness metrics. There are no changes in the experiment setup except that at this time, we consider the dataset as a whole. That means that the ranking of a method should reveal its functionality in both real-world networks and synthetic networks.

Table V shows the ranking of studied methods according to the six different qualities. A quick glance through the table shows that *separability* and *embeddedness* quality can be best discovered by louvain, fastgreedy and ganxis methods while high *density* communities can be detected by a dynamics-based method. *Compactness* and *modularity* are notably ameliorated by using an optimization-based method whereas clustering coefficient prefers cfinder3 method.

We can not notice any method that outperforms all the others in a general way. All studied community detection methods are doing well according to one or sometimes many criteria, but there is not a best method for all the studied cases. The modularity-based methods obtain the highest average ranking score of 4.3 over 11 methods and perform well in terms of most criteria but poorly for *density* and *clustering coefficient* quality. Globally, the rankings of these methods fluctuate around a medium value, which presents them as very highly competitive. However, if one has an interest in a specific criteria of structural community, this reference could provide an important information. It answers the question: "Which method one should choose to identify communities with specific characteristic requirements?" and assist analysts to choose the most appropriate method that corresponds to their expected notion of community.

One should note that this ranking is relative. A change of ranking criteria, goodness metric set or input dataset could change the orders of these methods. However, it does not mean that the ranking have no significance. Since we analyzed a very large number of communities with a variety of topological features, the result showed a statistical consistence in the functionality of community detection methods.

It is also worth mentioning that many methods are variable, which means communities that they discover may be different from one try to another. Hence, there could be some slight fluctuations in goodness scores that one can obtain during various experiments.

V. CONCLUSIONS AND PERSPECTIVES

In the paper, we analyze the functionality of a wide range of community detection methods using structural character-

Method	Sep.	Emb.	Den.	Com.	CCF.	Q.	Avg.
cfinder3	9	11	**1**	7	**1**	11	6.7
conclude	10	9	5	11	2	9	7.7
infomap	6	5	3	9	9	7	6.5
walktrap	7	6	2	10	8	8	6.8
label	5	**1**	9	6	5	5	5.2
copra4	11	7	4	4	7	10	7.2
demon	4	10	11	**1**	3	3	5.3
ganxis	3	4	6	8	6	6	5.5
fastgreedy	2	3	7	2	10	2	**4.3**
louvain	**1**	2	8	3	11	**1**	**4.3**
oslom	8	8	10	5	4	4	6.5

The abbreviations are reused from Table III, Avg. - *average ranking*. The highest ranked method by each goodness metric are bolded.

TABLE V

RAKING OF COMMUNITY DETECTION METHODS OVERALL THE DATASET

istics that correspond to many notions of good community structures. We yield two following contributions:

Firstly, we found that communities identified by these methods are structurally better than given metadata even in synthetic networks. In consequence, comparing structural communities with metadata as ground-truth by similarity metrics is not fair for evaluating the performance of algorithms that use merely topological information. This is not the first paper that indicate this issue, but with a new approach we reinforce the discovery with more evidences.

Secondly, to the best of our knowledge, there is actually no straightforward answer for the question: "Which community detection methods one should choose to optimize a topological goodness metric?". Even if there are many methods, which try to optimize some objective functions such as Modularity [2] or Conductance, many of them are usually biased especially in large-scale networks and produce unexpected partitions [33]. Actually, the choice of a method is often based on its simplicity (in terms of deployment and time consumption) and its availability. This paper brings the community structural goodness notion into consideration.

REFERENCES

[1] S. Fortunato. Community detection in graphs. Physics Reports 486, pp. 75-174 (2010)

[2] M. Girvan and M.E.J. Newman. Community structure in social and biological networks. Proc. Natl. Acad. Sci. USA 99 (PNAS), pp. 7821-7826 (2002)

[3] M.E.J Newman and M. Girvan. Finding and evaluating community structure in networks. Phys. Rev. E 69, vol. 69, pp. 026113, (2004)

[4] F. Musciotto, S. Delpriori, P. Castagno and E. Pournaras. Mining social interactions in privacy-preserving temporal networks. IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, pp. 1103-1110 (2016)

[5] R. Guimer and L.A.N. Amaral. Cartography of complex networks: modules and universal roles. Journal of statistical mechanics (2005)

[6] J. Diesner, T.L. Frantz and K.M. Carley. Communication Networks from the Enron Email Corpus It's Always About the People. Enron is no Different. Computational and Mathematical Organization Theory. vol. 11, issue 3, pp. 201-228 (2005)

[7] M. Coscia, F. Giannotti and D. Pedreschi. A classification for community discovery methods in complex networks. Statistical Analysis and Data Mining journal, Special Issue: Networks. vol. 4, issue 5, pp. 512-546 (2011)

[8] J. Reichardt and S. Bornholdt. Statistical Mechanics of Community Detection, Phys. Rev. E, 74, 016110 (2006)

[9] V.A. Traag and J. Bruggeman. Community detection in networks with positive and negative links, Phys. Rev. E 80, 036115, (2009)

[10] L. Peel, D.B. Larremore and A. Clauset. The ground truth about metadata and community detection in networks. The 5th International Workshop on Complex Networks and Their Applications. (2016)

[11] J. Yang and J. Leskovec. Defining and evaluating network communities based on ground-truth. Knowledge and Information Systems, vol. 42, issue 1, pp. 181-213 (2015)

[12] D. Hric, R.K. Darst and S. Fortunato. Community detection in networks: structural communities versus ground truth. Phys. Rev. E 90, 062805 (2014)

[13] G. Rossetti, L. Pappalardo and S. Rinzivillo. A novel approach to evaluate community detection algorithms on ground truth. Complex Networks VII: Proceedings of the 7th Workshop on Complex Networks, pp. 133-144 (2016)

[14] V.-L Dao, C. Bothorel and P. Lenca. Community structures evaluation in complex networks: a descriptive approach. NetSci-X 2017 - 3rd International winter school and Conference on Network Science. Springer Proceedings in complexity, pp. 11-19 (2017)

[15] G. Orman, V. Labatut and H. Cherifi. Comparative Evaluation of Community Detection Algorithms: A Topological Approach. Journal of Statistical Mechanics: Theory and Experiment, P08001 (2012)

[16] J. Yang and J. Leskovec. Defining and evaluating network communities based on ground-truth. Proceedings of 2012 IEEE International Conference on Data Mining (2012)

[17] A. Lancichinetti, S. Fortunato and F. Radicchi. Benchmark graphs for testing community detection algorithms. Physical Review E 78, 046110 (2008)

[18] A. Clauset, C.R. Shalizi, and M.E.J. Newman. Power-law distributions in empirical data. SIAM Review 51 (4), pp. 661-703 (2009)

[19] G. Palla, I. Dernyi, I. Farkas and T. Vicsek. Uncovering the overlapping community structure of complex networks in nature and society. Nature 835, pp. 814-818 (2005)

[20] P.D. Meo, E. Ferrara, G. Fiumara and A. Provetti. Mixing local and global information for community detection in large networks. Journal of Computer and System Science 80, pp. 72-87 (2014)

[21] M. Rosvall and C. T. Bergstrom. Maps of random walks on complex networks reveal community structure. Proc. Natl. Acad. Sci. USA 105, pp. 1118-1123 (2008)

[22] P. Pons and M. Latapy. Computing communities in large networks using random walks. Journal of Graph Algorithms and Applications, vol. 10, no. 2, pp. 191-218 (2006)

[23] U.N. Raghavanand, R. Albert and S. Kumara. Near linear time algorithm to detect community structures in large-scale networks. Phys. Rev. E 76, 036106 (2007)

[24] S. Gregory. Finding overlapping communities in networks by label propagation. New J. Phys. 12 103018 (2010)

[25] M. Coscia, G. Rossetti, F. Giannotti and D. Pedreschi. DEMON: a Local-First Discovery Method for Overlapping Communities. Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 615-623 (2012)

[26] J. Xie and B.K. Szymanski. Towards Linear Time Overlapping Community Detection in Social Networks. Advances in Knowledge Discovery and Data Mining, pp. 25-36, Springer (2012)

[27] V.D. Blondel, J.-L. Guillaume, R. Lambiotte and E. Lefebvre. Fast unfolding of communities in large networks. J. Stat. Mech. (2008)

[28] A. Lancichinetti, F. Radicchi, JJ. Ramasco, S. Fortunato. Finding Statistically Significant Communities in Networks. PLoS ONE 6(4): e18961 (2011)

[29] J. Creusefond, T. Largillier and S. Peyronnet. Finding compact communities in large graphs. IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, pp. 1457-1464 (2015)

[30] D.J. Watts and S.H. Strogatz. Collective dynamics of "small-world" networks. Nature, 393(1), pp. 440-442 (1998)

[31] G. Orman, V. Labatut and H. Cherifi. Qualitative Comparison of Community Detection Algorithms. Communications in Computer and Information Science, 167 pp. 265-279 (2011)

[32] A. Clauset, M.E.J. Newman and C. Moore. Finding community structure in very large networks. Phys. Rev. E 70, 066111 (2004)

[33] H. Almeida, D. Guedes, W. Meira and M.J. Zaki. Is There a Best Quality Metric for Graph Clusters?. Machine Learning and Knowledge Discovery in Databases. Lecture Notes in Computer Science, vol. 6911, pp. 44-59, Springer (2011)