# **Make Community Detection More Human**

Motahareh EslamiMehdiabadi University of Illinois at Urbana-Champaign eslamim2@illinois.edu Amirhossein Aleyasen University of Illinois at Urbana-Champaign aleyase2@illinois.edu

Karrie Karahalios University of Illinois at Urbana-Champaign kkarahal@illinois.edu

# **ABSTRACT**

Although there are dozens of community detection algorithms, we lack real community structures necessary to create reliable benchmarks for evaluation of such clustering algorithms. We developed a Facebook application to explore three common community detection algorithms. With this tool, we learn from people and real networks to evaluate existing approaches.

# **Keywords**

Community

#### 1. INTRODUCTION

Among the social networks, communities are one of the most salient parts which attract conspicuous notice. Communities, groups, clusters or modules all refer to the structures with two or more humans and their interactions based on common behaviours, interests or goals [1]. Having the most similarity within the communities besides being very dissimilar to members of other communities [2] is one of the most important features of a community structure. In order to show the relationships within a group, "Sociogram" has been used as a graphic representation of social links that a person has. Sociograms were developed by Moreno as a scheme for structural analysis of a community [3]. Figure 1 shows a typical representation of a social group by the sociogram.

Detecting communities within a social network is a notable problem in analysing people collective behaviour in the real world. As Freeman [5] mentions, many enormous efforts have been done to uncover the social groups. In Figure 2, a sociogram has been constructed to detect the groups of students in a classroom [6]. In order to build this sociogram, each student was asked to mention confidentially to two students who work with her/him on an activity. This visualization can give a teacher a clear view of relationships in the class by considering different groups.

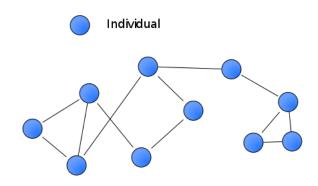


Figure 1: An example of representing a social network group by sociogram [4]

Finding the groups and community structures can be important in different applications. Clustering the web clients who are geographically near to each other and identifying the clusters of customers with similar interests are some of these applications [1]. Using sociograms to detect the groups can be useful in the relatively small networks. However, the large scale of real networks makes community detection process hard. Therefore, a large number of community detection or clustering algorithms has been proposed to find the groups within social networks.

In spite of the great attention of developing community detection algorithms, the issue of evaluating these algorithms in a fair way does not received enough consideration yet[1]. Hence, it is still impossible to determine an algorithm or a subset of algorithms as the best one in community detection field. This problem arises from the need of community detection algorithms to the "Ground-truth" where Big Data makes it hard or impossible in many areas. Therefore, some synthetic benchmark graphs [8, 9] by typical structure of communities have been considered to find the accuracy of a community detection algorithm. However, these benchmarks do not satisfy all the features of real communities which make the evaluation process of community detection methods unfair [7]. Besides these benchmarks, some evaluation metrics for community detection algorithms have been defined. Nevertheless, none of these evaluation metrics can measure the exact performance of the algorithms in uncovering communities.

Considering these issues in the evaluation of current com-

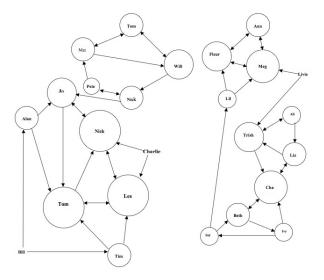


Figure 2: A classroom sociogram[6]

munity detection algorithms leads us to find a better way to assess these methods. If we have the ability of finding the ground-truth with the known group structures, it will be easy to evaluate an algorithm performance by measuring its accuracy. Additionally, using visualization as a tool can help to determine the structure of the communities better. Therefore, we decide to use a new evaluation approach which humanizes the community detection evaluation process. To make this idea true, we use Facebook as a real network structure for finding groups. Utilizing three different community algorithms, we develop a humanized Community Detection Application (CDA) that tries to evaluate these algorithm by applying them over different Facebook networks. This framework will be used by different people to visualize their Facebook network group structure and let them evaluate the algorithms themselves. We believe that this approach will provide a promising step towards evaluating the current community detection algorithms in a very different way. This new method will lead us to find the most appropriate community detection algorithms over the bed of social networks.

The rest of the paper is organized as follows. Section 2 elaborates the community detection problem by describing common community detection algorithms and different evaluation approaches. The proposed framework of humanized community detection is presented in Section 3. In this section, different networks' group structures have been investigated by some community detection algorithms. Then, the community detection application (CDA) and its user interface have been explained. Section 4 describes the evaluation metrics and the very first results of the study using CDA. The concluding remarks and future work are provided in Section 5.

# 2. COMMUNITY DETECTION

During the last decades, a great deal of studies has been done about communities in networks. These studies cover three main angles: community concept definition, community detection algorithms and evaluation approaches. Although there is no unique definition of community, most of studies agree on a general view of community in which the objects within a community are similar to each other while they are dissimilar to the objects in other communities [2]. Based on this definition, a huge number of community detection algorithms in discovering communities of networks have been developed.

In spite of this large number of methods, the evaluation approach as one of the main pivots of community detection field has received little attention. The most influencing factor on this lack of attention is the unavailability of real community structures [1]. Thus, many studies use synthetic benchmarks as the ground-truth community structure to evaluate their proposed methods. However, these benchmarks do not cover all of the features of real community structures which result in biased evaluations. Besides these synthetic benchmarks, there are a few number of real networks with known community structure. Nevertheless, these networks have their own issues such as small scales and unreliability of their community structure.

In the following, we will first explain about the community detection algorithms and specifically the ones we consider in our study. Then, we will describe the existing synthetic and real benchmarks as well as current metrics and measures for evaluation approach of community detection methods.

# 2.1 Algorithms

Based on the common definition of community, various objective functions have been defined which approximation or heuristic algorithms try to optimize them [20]. Considering these objective functions and algorithms, many community detection methods have been developed. These methods can be categorized by different classifications based on several criteria. One criterion in categorizing clustering methods is the information have been used to identify communities. While some of the methods use network structure, another group utilize nodes' features and attributes to detect communities. As the third group, there are some methods which use both network structure and nodes' features as input information. While having more features about nodes in addition of graph structure can be helpful to identify communities with more accuracy, it has its own costs such as gathering and processing this data. Here, we consider the first group which are structure-based groups since using the same information of graph structure makes these methods comparable to each other.

Another important criterion for categorizing clustering methods is the type of membership in communities. In classical community detection methods, each object can be the member of just one community. We called these methods "disjoint clustering" methods. However in many real-world networks, it has been observed that an object can be a member of more than one community. Many recent community detection methods consider this fact in their procedure which are named "overlapping clustering". Additionally, communities can have a multilevel structure in real-world networks which means a community can be a subset of another community [21]. This group which is called "hierarchical clustering" has been used widely in social network analysis [1].

Since there are a huge number of clustering methods, it is not applicable to use all of them in our study to evaluate them in a humanized manner. Therefore, we try to choose a representative method from each category to cover different types of membership. In the following, we introduce the candidates from different groups and chosen methods for our study.

- Girvan-Newman (GN) [9]: This algorithm is one of the
  most important methods in the history of community
  detection since it started a new age in this field [1]. It
  uses the concept of edge betweenness as a measure of
  edge importance to find community boundaries in networks. In spite of good accuracy, its time complexity
  is relatively high (O(n³)) in comparison with other algorithms. Therefore, some fast versions of this method
  have been introduced during recent years to improve
  its speed.
- Markov Clustering (MCL): Using the concept of Markov chains besides simulation of stochastic flows in graphs builds the fast and scalable unsupervised Markov clustering algorithm. This algorithm has a relatively high performance and been used widely in bioinformatics such as protein networks [10].
- Clauset-Newman-Moore (CNM): This algorithm is a hierarchical algorithm which tries to increase modularity metric as a density measure in each step. Similar to MCL, this algorithm can also perform fast on large scale networks [11].
- OSLOM: The Order Statistics Local Optimization Method (OSLOM) is the first community detection algorithm which accounts for edge directions and weights, overlapping communities, hierarchies and community dynamics. It is one of the most new methods with a high performance similar to the best existing methods besides a good scalability for large networks [12].
- Louvain: This algorithm as a hierarchical algorithm uses modularity as the objective function and try to maximize it heuristically. It has a very low computation time as well as a good quality. Utilizing it for social networks makes it a good candidate to use in our study [15].

Considering different membership methods, we decide to cover the three different approaches of disjoint, overlapping and hierarchical clustering. Therefore, we choose MCL from disjoint clustering group, Louvain method as a hierarchical algorithm and OSLOM as an overlapping method.

#### 2.2 Evaluation Approaches

The next step after developing a community detection algorithm is its evaluation. In ideal case, the identified communities by an algorithm on a network should be compared to the real communities of that network. However, in most cases obtaining the ground-truth of communities is a hard or impossible task. Therefore, the evaluation approaches of community detection methods are divided to two categories based on the availability of the ground-truth. The category which uses the ground-truth to evaluate a method is named

Extrinsic methods. The other category which is based on the unavailability of the ground-truth is called *intrinsic methods*. In the following, these groups of evaluation approaches have been introduced briefly.

#### 2.2.1 Extrinsic Methods

While accessing to the ground-truth of a grouping structure is desirable to evaluate a community detection algorithm performance, the lack of this ground-truth results in generating many synthetic benchmarks. Girvan-Newman benchmark [9] and LFR benchmark [8] are two of the most popular benchmarks which tries to cover real grouping structures' features. However, these benchmarks has many differences with real grouping structures that makes the evaluation process not realistic.

To make the evaluation process more practical, some real networks with known community structure have been introduced. The social network of Zachary's karate club [19] and the social network of bottlenose dolphins living in Doubtful Sound (New Zealand)[18] are two of the most common real benchmarks. However, the small scale of these networks and their different nature confine the evaluation procedure. Recently, [21] provides 10 real grouping structure of Facebook by asking people to create their groups by hand. Nevertheless, this study does not consider some challenges of grouping a large number of friends by hand such as human errors by getting tired.

Considering the introduced synthetic and real benchmarks, the extrinsic measures have been provided to compare a ground-truth of groupings with the resulted grouping structure of an algorithm. Since there could be no exactly association between real communities and resulted communities, three different categories have been introduced to compare two groupings: pair counting, cluster matching and information theory [1]. In the Experiments section, a metric which has a good quality in finding the accuracy will be introduced.

#### 2.2.2 Intrinsic Methods

On the unavailability of the ground-truth, we have to use some measures that evaluate a clustering approach by determining its ability in separating the clusters besides making inside of the clusters dense. *Modularity* is one of the most common metrics in community detection evaluation. This metric measures the quality of the network division into different communities by computing the difference of the number of edges between nodes in a special community and the expected number of such edges in a random graph with identical degree sequence [13]. This measure has been used in this project to evaluate the introduced community detection algorithms over different networks structures.

#### 3. THE PROPOSED FRAMEWORK

As we mentioned before, one of the most important problems of community detection algorithms is the unavailability of the ground-truth for testing their performance. Therefore, having a network with labeled nodes by their communities can be a valuable ground-truth. However, these groundtruths are not available easily due to the unavailability or unreliability of nodes' labels. In this paper, we decide to consider online social networks as an appropriate bed of constructing labeled networks by different communities. Here, we build a framework on Facebook as a popular online social network to ask people identify the groups which their friends belong too. To achieve this goal, we apply the aforementioned three community detection methods of MCL, Lovain and OSLOM on a user's Facebook network of friends and then ask him/her to revise the grouping structures. We believe this data can be a reliable and valuable labeled data in the community detection area.

Before introducing the humanized community detection framework, we will analyze the structure of Facebook network with two other networks with different structures to figure out how community detection algorithms work on different structures. Then, the user interface of the built application will be described.

# 3.1 Different Grouping Structures

Since there are various social networks with different structures, we first compare the group structure of Facebook network with some different networks. This investigation will help us to find out how current community detection algorithms can find groups in different structures. Currently, we consider one of the author's Facebook network to compare with other structures. As the second structure, the network of American football games between Division IA colleges during regular season fall 2000 [9] has been considered. The last is one the co-authorship network of theory scientists [14] which is named NetScience. Table 1 shows the properties of these networks.

Table 1: The Different Network Structures

Network	Node#	Edge#	Description	
Facebook	290	5041	Facebook Network	
Football	115	615	US football network	
NetScience	1589	2742	Collaboration network	

Here, we choose the methods of GN and CNM from the disjoint clustering category to apply on these networks and analyze the quality of resulted grouping structures. Since for Football and NetScience networks, there is no ground-truth of grouping structure, we use 'modularity' as an intrinsic measure to compare the generated grouping structures by the mentioned algorithms.

Figure 3 shows the amount of modularity on the resulted grouping structures of these networks. As higher modularity shows better grouping structure, it can be seen that the community detection algorithms perform much better on NetScience and Football network in comparison with Facebook network. This fact shows the more complicated group structure of Facebook network respect to the scientist and football networks. Indeed, the relationships between people in Facbook is more complex than the simple and formal relationships of scientists or football players. Therefore, finding an appropriate community detection algorithm for social networks such as Facebook will be an important issue which should be addressed. Here, we build a humanized community detection framework as a step toward this goal which will be explained in the following section.

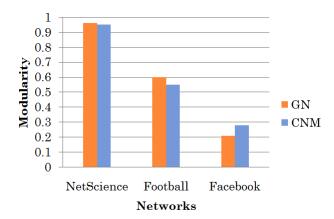


Figure 3: Comparison of modularity as an evaluation metric over different networks by different community detection algorithms

#### 3.2 User Interface

We develop Community Detection Application (CDA)<sup>1</sup> which applies three community detection algorithms of 'MCL' as a disjoint clustering, 'Louvain' as a hierarchical clustering and 'OSLOM' as an overlapping clustering over Facebook network. Each algorithm generates a different structure of groups or communities from Facebook network. The user can modify these group structures by labeling the groups and editing their members.

In the first step, user logins to the application by using his Facebook account. Since each user should evaluate three different community detection algorithms, we consider a random permutation of these algorithms to apply over the user Facebook respectively. It will make the evaluations fair due to removing some human impacts such as exhaustion and being affected by the very first results. After the user logins to the application, he will see the result of the first algorithm which will be a clustering of his Facebook friends. Figure 4 shows a snapshot of the application on one of the authors' Facebook network after applying the disjoint clustering method of MCL.

The left panel named 'Groups Panel' shows the groups of friends that the algorithm detects. Each group has been named by a number which the user can modify it. The number of the groups depends on the algorithm which is executed and also the number of friends the user has. At the bottom of the Groups Panel, a group named 'Un-Grouped' can be seen. This group contains the friends that the algorithm cannot find any appropriate group to put them in. This group is similar to the "rag bag" which is used in may practical scenarios for the objects which cannot be assigned to any specific group [2].

In the right panel named 'Members Panel', the members of each group have been shown by clicking on that group. The user can label a group if he finds at least  $\frac{2}{3}$  of its members from a meaningful group. Additionally, the user is able to

http://social.cs.illinois.edu/users/motahhare/CDA/index.php/frontend/intro

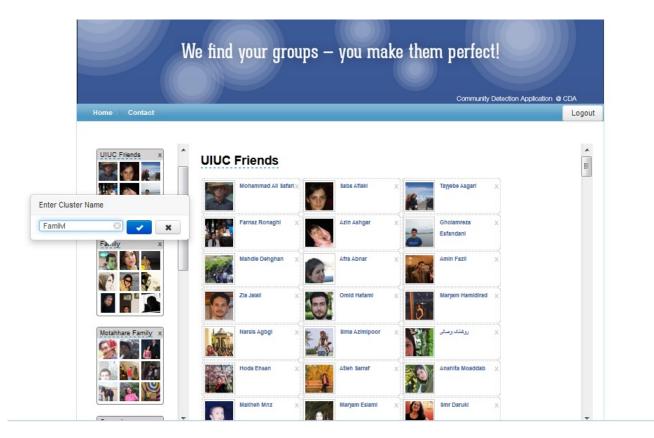


Figure 4: A Snapshot of Community Detection Application

delete or move a group's member. All the deleted members will be automatically moved to 'Un-Grouped' group. If the executed algorithm is an overlapping algorithm, dragging a friend from a group to another group will not remove it from the original group since a friend can belong to more than one group.

Since the generated groups by an algorithm can be not comprehensive, we let user to create his own groups if he thinks the current groups do not cover all of his Facebook friends. Additionally, we ask user to check the Un-Grouped members and move them to an appropriate group if it is possible. Finally, if there is any friend that the user does not know which group it belongs too, he can put it in Un-Grouped group. The aforementioned procedures will be repeated for all of the three community detection algorithms.

#### 4. EXPERIMENTS

As the goal of this study is providing a humanized evaluation of community detection methods, we start to conduct user studies by asking people to use CDA and evaluate the results. We have plan to ask 30 people to use our application and currently, we have done the study by 3 users (the authors).

# 4.1 Evaluation Metrics

**BCubed.** Since a user provides the ground-truth, we will use the extrinsic methods to evaluate the results of each algorithm. Considering different categories of extrinsic methods, we choose BCubed metric as the metric that covers the

essential criteria of community detection evaluation such as cluster homogeneity, cluster completeness and considering Rag Bag [22]. This metric can be also extended for overlapping clustering methods which is needed for our study as it covers all kinds of membership.

Let define L(f) as the assigned cluster to friend f by a clustering algorithm and C(f) as the cluster of friend f the user has been assigned to it as the ground-truth. In the disjoint clustering, the correctness of two friends of  $f_i$  and  $f_j$  has been defined as:

$$Correctness(f_i, f_j) = \begin{cases} 1 & \text{if } L(f_i) = L(f_j) \Leftrightarrow C(f_i) = C(f_j) \\ 0 & \text{Otherwise} \end{cases}$$
(1)

To aggregate precision and recall, we use F-measure metric:

$$F = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$
 (2)

**Un-Grouped Category.** Besides BCubed metric, the Un-Grouped category can also be a measure of an algorithm performance. If an algorithm puts a member in un-grouped category and the user can put it in another group, it means an error for the clustering algorithm. Additionally, it will be an interesting question to investigate how many members will remain in un-grouped category after the user modifies an algorithm's result.

User's Rate. Besides these statistical metrics, we define

a user's rate for each algorithm which shows the level of satisfaction of the user from that algorithm before and after modification. We believe this metric can be of great value in humanizing the evaluation of community detection methods.

#### 4.2 Performance Evaluation

The results of the study over 3 Facebook networks in average have been provided for two methods of MCL and OSLOM in table 2.

Table 2: The evaluation results over my Facebook network

Algorithm	Precision	Recall	F-measure
MCL	91%	82%	86%
OSLOM	95%	77%	84%

The more comprehensive results and analysis will be provided further after completing user studies. The next step after analyzing and comparing algorithms' performance by different evaluation metrics will be comparing resulted ground-truths by users for each algorithm. We believe there is no unique ground-truth for communities on social networks and have plan to test this hypothesis by comparing the ground-truths generated by users after modifying the three community detection algorithms' results.

We also want to compare different membership types over Facebook to find out the dominant type of membership in social networks. Additionally, we want to analyze overlapping clustering to find the percentage of multiple membership versus disjoint membership. Finally, we are interested to study the structure of hierarchical communities over social networks by using the user studies results.

We believe conducting user studies can go beyond the goal of finding ground-truths of community structures since it can help people to get a big picture from their online social relationships in the form of communities. This picture can help them to gain insight about their relationships. Here there is a quote from one of the participants while he confronts with some communities which he has no idea about its members: "I do not believe I have 3 large groups here that I do not know any of their members! I should change my way of accepting friend requests in future. I might also look over my friends list and delete some of the unknown people."

# 5. CONCLUSION AND FUTURE WORK

This paper provides a new different view over community detection algorithms by a humanized approach. The difficulty of accessing to the ground-truth is one of the main issues in community detection problem which makes its evaluation hard. Here, we propose a humanized framework which tries to find a new way of evaluating the community detection algorithms. The Community Detection Application (CDA) uses the three common algorithms of MCL, Lovain and OSLOM to find the groups of a Facebook user's network. Evaluating the results of these algorithms will be done by the user as the best one for assessing his/her groups structure.

There are several directions for future work that should be addressed. First, conducting the user studies will give us valuable insights toward community detection algorithms and their challenges. Additionally, some new evaluation metrics and questions should be designed to make the assessment process more reliable. Utilizing the results of the interviews can lead us to a new community detection algorithm which considers some new metrics in finding communities in online social networks.

### 6. REFERENCES

- Fortunato, Santo. Community detection in graphs. Physics Reports 486.3 (2010): 75-174.
- [2] Han, Jiawei, and Micheline Kamber. Data mining: concepts and techniques. Morgan Kaufmann, 2006.
- [3] Moreno, Jakob L. Sociometry in relation to other social sciences. Sociometry 1.1/2 (1937): 206-219.
- [4] http://en.wikipedia.org/wiki/Sociogram, Retrieved at 1st November, 2012.
- [5] Freeman, Linton C. Visualizing social networks. Journal of social structure 1.1 (2000): 4.
- [6] http://www.6seconds.org/2012/05/08/sociograms-mapping-the-emotional-dynamics-of-a-classroom/, Retrieved at 10th November, 2012.
- [7] A. Lancichinetti, S. Fortunato, and F. Radicchi, Benchmark graphs for testing community detection algorithms, Physical Review E, APS, 2008.
- [8] Lancichinetti, Andrea, and Santo Fortunato. "Community detection algorithms: A comparative analysis." Physical review E 80.5 (2009): 056117.
- [9] Girvan, Michelle, and Mark EJ Newman. Community structure in social and biological networks. Proceedings of the National Academy of Sciences 99.12 (2002): 7821-7826.
- [10] van Dongen, Stijn Marinus. Graph clustering by flow simulation., PhD Thesis, University of Utrecht, The Netherlands (2000).
- [11] Clauset, Aaron, Mark EJ Newman, and Cristopher Moore. Finding community structure in very large networks. Physical review E 70.6 (2004): 066111.
- [12] Lancichinetti, Andrea, et al. "Finding statistically significant communities in networks." PloS one 6.4 (2011): e18961.
- [13] Yang, Jaewon, and Jure Leskovec. Defining and evaluating network communities based on ground-truth. Proceedings of the ACM SIGKDD Workshop on Mining Data Semantics. ACM, 2012.
- [14] M. E. J. Newman, Finding community structure in networks using the eigenvectors of matrices, Preprint physics/0605087, 2006.
- [15] Blondel, Vincent D., et al. "Fast unfolding of communities in large networks." Journal of Statistical Mechanics: Theory and Experiment 2008.10 (2008): P10008.
- [16] SNAP, Community Detection Library. http://snap.stanford.edu/. Retrieved on September 2012.
- [17] MCL Tool. http://micans.org/mcl/. Retrieved on September 2012.
- [18] Lusseau, David. "The emergent properties of a dolphin social network." Proceedings of the Royal Society of

- London. Series B: Biological Sciences 270. Suppl<br/> 2 (2003): S186-S188.
- [19] Zachary, Wayne W. "An information flow model for conflict and fission in small groups." Journal of anthropological research (1977): 452-473.
- [20] Leskovec, Jure, Kevin J. Lang, and Michael Mahoney. "Empirical comparison of algorithms for network community detection." Proceedings of the 19th international conference on World wide web. ACM, 2010
- [21] McAuley, Julian, and Jure Leskovec. "Learning to discover social circles in ego networks." Advances in Neural Information Processing Systems 25. 2012.
- [22] Amigo, Enrique, et al. "A comparison of extrinsic clustering evaluation metrics based on formal constraints." Information retrieval 12.4 (2009): 461-486.