Real World Community Analysis and Evaluation

CS 579 Online Social Network Analysis

Dr. Cindy Hood
11/13/25

Final Project Proposals – Due 11/8

The final project is an extension of your analysis from HW 4. You are required to include the following: Chicago Community Areas, Census data, Network modeling and analysis, and Geographic mapping/visualization. Beyond this, there are many different ways to extend your analysis from HW 4 (e.g. include all CAs, include additional census data, include data from other sources, include additional analysis methods, etc.). Your proposal should include the following aspects:

- What is the goal of your project?
- How will you extend what you did for HW 4?
- What motivates this extension?
- What data will you use?
- What will be the most challenging part of the project?

The final project may be done individually or in a group of up to 4 students. Groups should submit one proposal that includes technical content for each team member.

The proposal should be one page or less, group proposals may be longer as needed to convey the technical proposal for each group member. Groups should submit one proposal that includes the names of all group members.

Review Complete – most received ok, a few revise and resubmits

Remaining Exam and Deliverables

- Final project progress report and video
  - Due 11/21
- HW #5
  - Problems that will help you prepare for Exam 2
  - Due 11/24 (No late days)
- Exam 2
  - 12/2
- Final project poster presentation/video (online students)
  - 12/4
- Final project report
  - Week of 12/8
    - Specific date tbd

Reference

Journal of Scientometric Research, 2025; 14(1):62-74.
https://www.jscires.org

Research Article

Evaluating Community Detection Algorithms: A Focus on Effectiveness and Efficiency

Srilatha Tokala[1], Murali Krishna Enduri[1,*], T. Jaya Lakshmi[1,2], Koduru Hajarathaiah[1]

[1]Department of Computer Science and Engineering, Algorithms and Complexity Theory Lab, SRM University, Amaravati, Andhra Pradesh, INDIA.
[2]Department of Computing and Software Engineering, Sheffield Hallam University, Sheffield, England, UNITED KINGDOM.

ABSTRACT
Many practical problems and applications are characterized in the form of a network. If the network becomes huge and complex, it becomes very difficult to identify the partitions and the relationships among each of the network's nodes. As a result, the graph is divided into communities and several community detection methods are proposed to associate those communities. The formation of virtual clusters or communities often occurs in networks due to the likelihood of individuals with similar choices and desires associating with one another. Detecting these communities holds significant benefits across various applications, such as identifying shared research areas in collaboration networks, detecting protein interaction in biological networks and finding like-minded individuals for marketing and suggestions. Numerous community detection algorithms are applied in different domains. This paper gives a brief explanation of existing algorithms and approaches for community detection like Louvain, Kernighan-Lin, Girvan Neuman, Label Propagation and Leiden algorithms as well as discusses various applications of community detection. We have evaluated our comparison with six different datasets namely biocelegans, ca-netscience, usair97, webpolblogs, email-univ and powergrid for comparing the efficiency of the methods. The modularity and conductance scores are used to assess the caliber of the partitioned community. A special emphasis on the comparison of these community detection methods is concerned and how the quality resembles and the time taken for its evaluation. We have evaluated all these algorithms and concluded that Louvain and Leiden community detection algorithms are used for effective community division in terms of its structure and time.

Keywords: Network, Community, Community detection and Modularity score.

Correspondence:
Murali Krishna Enduri
Department of Computer Science and Engineering, Algorithms and Complexity Theory Lab, SRM University, Amaravati, Andhra Pradesh, INDIA.
Email: muralikrishna.e@srmap.edu.in

Received: 26-07-2023;
Revised: 17-01-2024;
Accepted: 01-05-2024.

INTRODUCTION

https://jscires.org/10.5530/jscires.20250839

Merriam-Webster definition of community

community noun

com·mu·ni·ty | kə-ˈmyü-nə-tē

plural **communities**

often attributive

Synonyms of *community* >

1 : a unified body of individuals: such as

a : the people with common interests living in a particular area

> ... there was to be a school party on Saturday night for the entire *community*.
> – Richard Peck

broadly : the area itself

> ... Linfield, a farming *community* about seven miles away.
> – James McBride

b : a group of people with a common characteristic or interest living together within a larger society

> a *community* of retired persons

> a monastic *community*

c : a body of persons of common and especially professional interests scattered through a larger society

> the academic *community*

> the scientific *community*

> That kind of success has a ripple effect in the creative *community*.
> – Vivian Song

d : a body of persons or nations having a common history or common social, economic, and political interests

> the international *community*

> ethnic *communities*

> The series won acclaim ... for its affectionate portrayal of the LGBTQ+ *community* ...
> – Lisa Respers France

e : a group linked by a common policy

f : an interacting population of various kinds of individuals (such as species) in a common location

> a diverse *community* of microorganisms

> ... an impressive plant *community* of marshes, cattail stands, bulrush beds and mesquite bosques ...
> – Jake Frederico

g : a political unit (such as a nation or state) : **COMMONWEALTH**

Community Detection Algorithms

Table 1: Exploring the Landscape: An Overview of Community Detection Algorithms.

Community Type	Algorithm	Year	Contribution
Non-Overlapping	Girvan Neuman	2002	Computes communities based on edge-betweenness.
	Louvain	2008	Optimizes the modularity by improving the community structure within neighborhoods of nodes.
	Kernighan-Lin	1970 (existed from 2009).	Iterative partitioning approach that minimizes the total edge cut between communities.
	Leiden	2019	Modularity optimization is accomplished by eliminating local maximum traps and assuring more robust community assignments, which provide computational efficiency and scalability for large networks.
Overlapping	Label Propagation	2007	Label updating process, which propagates node labels based on the most common labels among their neighbors.
	Infomap	2008	Use information theory and random walks to identify communities in a network.
	Walktrap	2011	Identifies communities by detecting dense regions of the network where random walks are likely to remain in a given neighborhood.

https://jscires.org/10.5530/jscires.20250839

Community Detection Algorithms

Table 2: Analyzing Community detection techniques: Strengths vs. Weaknesses.

Community Detection Algorithm	Strengths	Weaknesses
Girvan Neuman Community detection.	The network is decomposed into communities by iteratively removing high-betweenness edges. This algorithm is applicable for weighted and unweighted as well as directed and undirected graphs.	The performance is very sensitive for its size and for the density of the network. The calculation of betweenness centrality for the edges is very expensive.
Louvain Community detection.	Within the short-limited time, Louvain algorithm can handle large networks. Based on the structure of the network the communities are identified without any prior knowledge.	The communities are more dependent on the initial partition. Due to the non-deterministic nature of the network, different results will be produced for different iterations.
Kernighan-Lin Community detection.	If the partition is not for two groups this algorithm is better, as well as get well-defined communities in the network. It improves the quality of the partition by iteratively exchanging the nodes.	The algorithm struggles in defining the community structures for large networks. As it is of only bipartition, this algorithm is not preferred in most of the cases.
Leiden Community detection.	It can handle overlapping community structures which help in analyzing complex communities. The algorithm produces more deterministic results for the given initial conditions.	It mainly focuses on modularity optimization, which is not preferred for all network structures. The memory and computational requirements has become a main drawback which results in multiple resolutions.
Label Propagation Community detection.	This algorithm is simple, scalable and efficient which is suitable for large networks. It can adopt to various community structures.	The algorithm feels vulnerable to assign label at initial stages. Incase if node belongs to more community structures, the algorithm struggles in detecting the communities.
Infomap Community detection.	The algorithm has the capability of identifying local and global community structures with different scales. If a node belongs to more communities, this algorithm can detect the overlapping communities.	It is computationally demand for large networks. Identifying the similar communities are more difficult for the larger structures if their sizes are comparable.
Walktrap Community detection.	The emphasis of local neighborhood is considered for random walks of short lengths. The run time is manageable as the algorithm is scalable and efficient for analyzing large networks.	It is sensitive to noise which leads to detecting irrelevant nodes in the community structures. The choice of selecting parameters requires some practical analysis.

Assumptions of Community Detection Algorithms

Focus is on similarities

- Does this tell the whole story?
- Is this the correct question/focus/lens for *our assignment*?

What makes a good community in real life?

How would *you* define a good community?

- Get in groups
  - Discuss
  - Research
- Create 1 slide of the key characteristics of a good community
- Put a title slide on it with everyone’s names and put in the shared folder