Community Detection in Social Networks

Author : Punam Bedi Chavvi Sharma

Presented by : Dinesh Reddy

Oct 13th, 2020

Problem Statement

The expansion of the web and emergence of a large number of Social networking sites (SNS) have empowered the users to easily interconnect on a shared platform. The tendency of people with similar tastes, choices and preferences to get associated in a social network leads to the formation of virtual clusters or communities. Detection of these communities can be beneficial for numerous applications such as finding a common research area in collaboration networks, finding set of likeminded users for marketing and recommendations. A large number of community detection algorithms have been proposed and applied to several domains in the literature. This paper presents a survey of the existing algorithms and approaches for detection of communities in social networks.

Social Network:

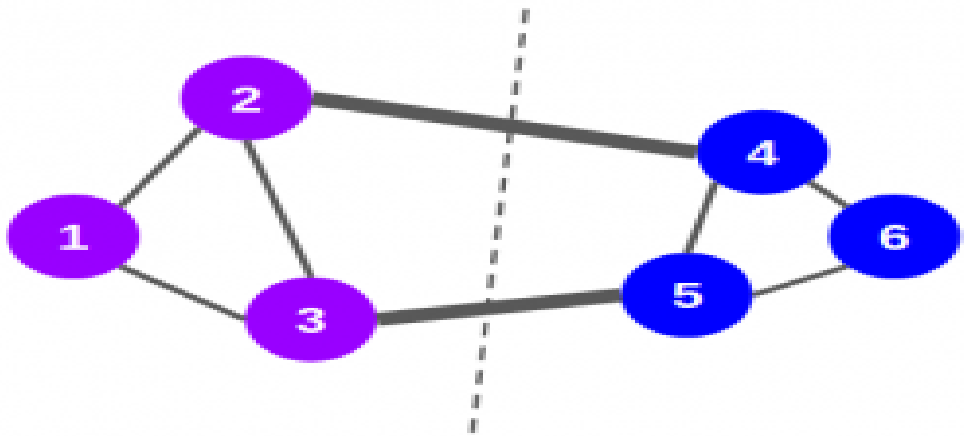
A Social network can be represented by a graph consisting of a set of nodes and edges connecting these nodes. The nodes represent the individuals/entities and the edges correspond to the interactions among them.

Community:

A community, with respect to graphs, can be defined as a subset of nodes that are densely connected to each other and loosely connected to the nodes in the other communities in the same graph.

Community Detection:

In a large scale network, such as an online social network, we could have millions of nodes and edges. Detecting communities in such networks becomes a herculean task. Therefore, we need community detection algorithms that can partition the network into multiple communities.



Methodology

There are primarily two types of methods for detecting communities in graphs:

Agglomerative Methods: In agglomerative methods, we start with an empty graph that consists of nodes of the original graph but no edges. Next, the edges are added one-by-one to the graph, starting from "stronger" to "weaker" edges. This strength of the edge, or the weight of the edge, can be calculated in different ways.

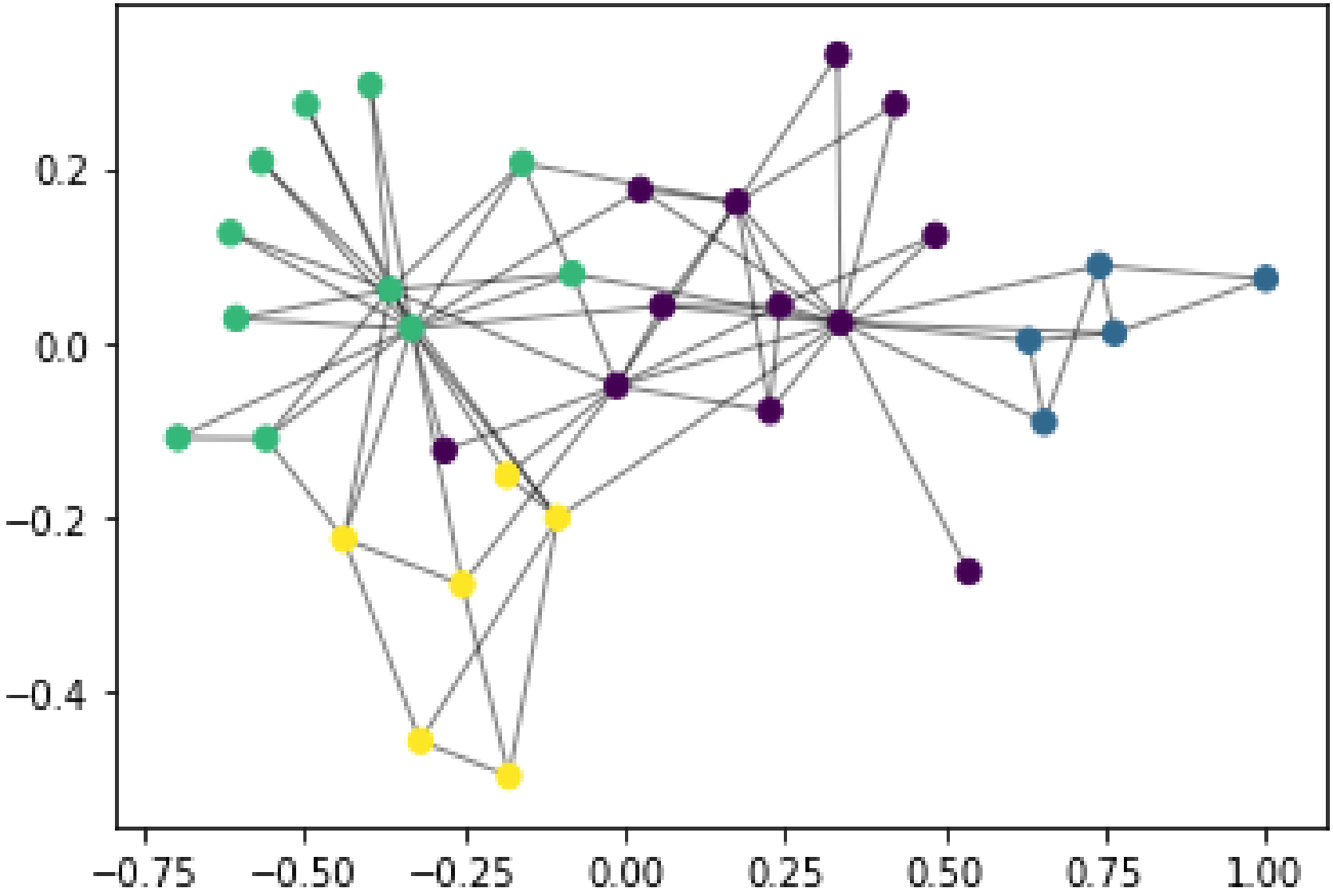
Divisive Methods: In divisive methods, we go the other way round. We start with the complete graph and take off the edges iteratively. The edge with the highest weight is removed first. At every step, the edge-weight calculation is repeated, since the weight of the remaining edges changes after an edge is removed. After a certain number of steps, we get clusters of densely connected nodes.

Methodology (cont.,)

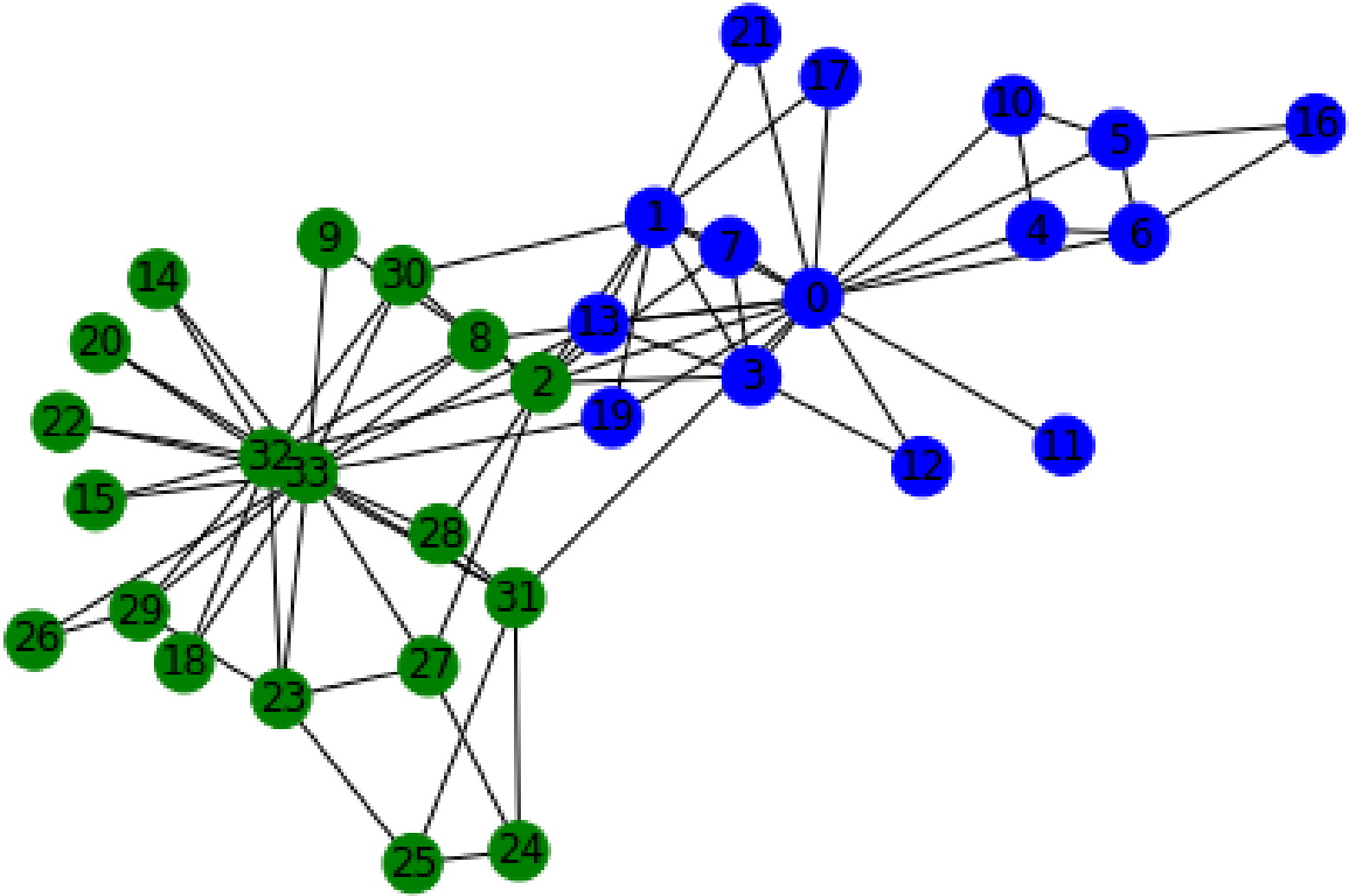
Girvan-Newman Method: It is an example of Divisive Method. The communities in a graph are discovered by iteratively removing the edges of the graph, based on the edge betweenness centrality value. The edge with the highest edge betweenness is removed first.

Louvain Method: It is an example of Agglomerative Method which iteratively optimizes local communities until global modularity can no longer be improved given perturbations to the current community state. It is crazy fast, and significantly more accurate than other fast methods.

Results



Results (cont.,)



Conclusion

The area of community detection holds a vast potential for discovery of communities in today's exponentially growing social networks. The basic concepts of social networks, community structure and methods for grouping similar items are presented in this paper. Application of the algorithms to detect communities in actual networks of Facebook, Twitter, and LinkedIn etc. can provide substantial amount of information for myriad purposes. The discovery and analysis of communities is used in biology, sociology and many other branches of science. Such information may prove to be useful for commercial, educational or developmental purposes.

THANK YOU