Journal of Information Engineering and Applications www.iiste.org
ISSN 2224-5782 (print) ISSN 2225-0506 (online)
Vol.3, No.6, 2013- Selected from International Conference on Recent Trends in Applied Sciences with Engineering Applications

Community Detection in Social Networking

Yachana Bhawsar, Dr. G.S.Thakur
Research Scholar, Assistant Professor, MANIT Bhopal
yachanabhawsar@gmail.com, ghanshyamthakur@gmail.com

Abstract
A social structure made of nodes (individuals or organizations) that are related to each other by various interdependencies like friendship, kinship, etc. The notion of social networks, where relationships between entities are represented as *links* in a graph, has attracted increasing attention in the past decades. Thus social network analysis, from a data mining perspective, is also called *link analysis* or *link mining*. Social creatures interact in diverse ways: forming groups, sending emails, sharing ideas, and mating. Some of the interactions are accidental while others are a consequence of the underlying explicit or implicit social structures. In order to understand social interactions, it is therefore crucial to identify these social structures or "communities," which are loosely defined as collections of individuals who interact unusually frequently. How can we find communities in dynamic networks of social interactions, such as who calls whom, who emails whom, or who sells to whom? In this paper we will discuss various community detection algorithm and measures of community detection in any complex network.
Keywords: Social network, Data mining, Community Detection

1. Introduction

Communities are subset of actors among whom there are relatively strong, direct, intense, frequent or positive ties. People form communities in social media. Community detection is to formalize the strong social groups based on the social network properties. Roughly community detection methods can be divided into 4 categories. These are Node-Centric, Group-Centric, Network-Centric and Hierarchy-Centric.

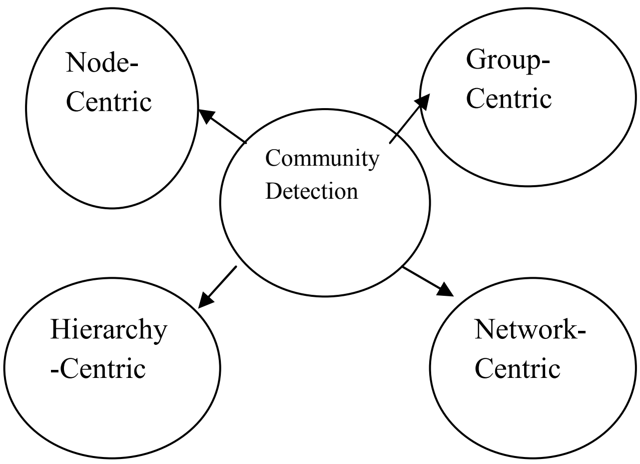

Figure1: Community Detection Criteria

Social media communities can further be described as explicit or implicit. Explicit Communities are created as a result of human decision and acquire members based on human consent. Example of explicit social media communities are Facebook and Flicker groups. Implicit communities, on the other hand, are assumed to exist in the system and "wait" to be discovered. Implicit communities are particular important for two reasons: (a) they do not require human effort and attention for their creation and (b) they enable the study of emergency phenomena within social media system. This survey focuses on the definition and discovery of implicit communities.

2. Performance Comparison

In assessing the performance of community detection methods, there are two fundamental aspects that one needs to consider: (a) the computational complexity and (b) the requirements of the method in term of main memory. The incremental computation properties of community detection method constitute an additional performance consideration, which is increasingly important in the context of Social Media systems. In this section, we provide a theoretical discussion of the first two performance aspects, computational complexity and memory requirements, for a variety of methods, and report the results of an experimental study that compares the performance of eight popular community detection methods on a wide variety of synthetic network.

3.Community structure

Basics

A partition is a division of a graph into disjoint communities, such that each node belongs to a unique community. A division of a graph into overlapping (or fuzzy) communities is called a cover. We use P = {C1, . . . , Cn_c} to denote the partition, which is composed of n_c communities. In P, the community to which the node v belongs to is denoted by σ_v. By definition we have $V = \cup_1^{nc}$ Ci and ∀i = j,Ci ∩ Cj = ∅. We denote by S = {S1, . . . , Sn_c} a cover composed of n_c communities. In S, we may find a pair of community S_i and Sj such that Si ∩ Sj = ∅. Given a community C ⊆ V of a graph G = (V,E), we define the internal degree k^{int}_v (respectively the external degree k^{ext}_v) of a node v ∈ C, as the number of edges connecting v to other nodes belonging to C (respectively to the rest of the graph). If $k^{ext}_v = 0$, the node v has only neighbors within C: assigning v to the current community C is likely to be a good choice. If $k^{int}_v = 0$ instead, the node is disjoint from C and it should better be assigned to a different community[6]. Classically, we note $k_v = k^{int}_v + k^{ext}_v$ the degree of node v. The internal degree k^{int} of C is the sum of the internal degrees of its nodes. Likewise, the external degree k^{ext} of C is the sum of the external degrees of its nodes. The total degree k_C is the sum of the degrees of the nodes of C. By definition:

$$k_C = k^{int}_c + k^{ext}_c$$

Modularity

One may want to measure the quality of a partition through a quality function, which assigns a score to each partition of a graph. In this way, partitions can be ranked based on their score given by the quality function. Partitions with high scores are ”good”, so the one with the highest score is by definition the best[6].

The widest accepted quality function is the modularity introduced by Newman and Girvan. Let e_{ij} be the fraction of edges in the network that connect nodes in community i to those in community j, and $a_i = \sum_j e_{ij}$. The modularity measure is
defined as:

$$Q = \sum_i \left(e_{ii} - a_i^2\right).$$

4. Related work

Community detection is very challenging field of research. There are various community detection algorithms available that includes removal of high-betweenness edges[1], optimization of modularity[2], detection of dense subgraphs, statistical inference, random walk[3], and many others.

5. Conclusion

This paper provides an overview of community detection on social networking.
The two major problems concerning community detection are overlapping community detection and dynamic community detection.

5. References

[1] M. Girvan, M. E. J. Newman. 2002. Community structure in social and biological networks, *Proc. Natl. Acad. Sci.* USA 99, 7821–7826.
[2] M. E. J.Newman . 2004. Fast algorithm for detecting community structure in networks. *Phys. Rev. E* 69 066133.
[3] P. Pons, M. Latapy. 2005. Computing communities in large networks using random walks. *Physics and Society* (physics.soc-ph).
[4] Qiong Chen, Ming Fang. 2012. An Efficient Algorithm for Community Detection in Complex Networks. THE 6TH sna-kdd Workshop'12, Beijing, China.
[5] Erwan Le Martelot, Chris Hankin. 2012. Fast Multi-Scale Detection of Relevant Communities in Large-Scale Networks.
[6] Q.Wang and E.Fleury, Understanding community evolution in Complex systems science, 1st International Workshop on Dynamicity, December 12, Collocated with OPODIS 2011, Toulouse, France