International Journal of Information Processing, 10(3), 45-54, 2016
ISSN : 0973-8215
IK International Publishing House Pvt. Ltd., New Delhi, India

A Survey on Modularity based Overlapping Community Detection Techniques in Social Networks

Chitra Devi J[a], Poovammal E[a]

[a]Department of Computer Science and Engineering, SRM University, Kattankulathur, Chennai 603 204, India, Contact: chitradevi21@gmail.com

Social Network is widely popular data mining domain in the field of research. Social network analysis combines the various fields namely sociology, anthropology, social psychology, sociolinguistics etc. Some characteristics that differentiates social network from real network are user-based, interactive, community-driven, relationships and emotion expression. Most of social networks exhibits the community-driven characteristics within the network. Social network is treated as a graph with nodes and edges connecting it. Finding the similar nodes based on strong connectivity gives the concept namely community detection. In community detection, a node of social network graph involves overlapping structure. This paper gives a detailed survey on various modularity based approaches on detecting the overlapping communities in the social networks. It also summarizes the characteristics and limitations of modularity based overlapping community detection algorithms.

Keywords : Community Detection, Modularity, Overlapping Community, Social Network, Social Network Mining.

1. INTRODUCTION

As real world scenario is dynamic and evolving, Social network is used intensively in wide range of applications and represented as a network graph with nodes and edges. Nodes represent the individual users/actors/items/resources whereas edge represents the link/flow of interaction/relationship among the users. The arrangement of nodes and edges in a graph is coined as topology. Some common techniques involved in social network includes Recommendation Systems, Link Analysis, Expert Identification, Influence Propagation, Trust and Distrust Relationship Prediction, Opinion Mining, Mood Analysis and Community Detection[1–3].

Recommendation systems analyze the available data and suggest something the actor might be interested in. As a result a new link is introduced in the network. Recommending friends or recommending resources happens through collaborative filtering and Content-based filtering. Each node represents entity and the analyzes of links among the nodes gives the behavior pattern of different activities. Link analysis helps in finding this behavior pattern of a social network. Strategy of finding an expert in a required domain by analyzing the social network is coined as Expert Identification. With the help of previous log information, Influence Propagation between the nodes is identified. Two models are used in Influence Propagation namely Independent Cascade Model and Linear Threshold Model. Based on attributes of a node, a relationship or a link in a social network is identified as Trustable or Not. This study is termed as Prediction of Trust or Distrust relationship.

Opinion mining involves building a system to collect and categorize opinions about a product or a person. It uses ideas of machine learning, text mining and natural language processing. The levels of opinion mining are document level, sentence level and phrase level. To track the mood of public about a particular product, mood analysis is performed. Set of actors interacting with each other frequently

Table 2
Overview of Modularity Adaptation based Community Detection

Author Name/ Algorithm	Observations
Girvan-Newman[6]	The algorithm runs slowly, taking time complexity as $O(m^2n)$ on a network of n vertices and m edges.
Leicht *et al.,*[7]	Communities are identified well. But not applicable for weighted or overlapped communities.
Nicosia *et al.,*[8]	This belonging factor is defined for both indegree as well as outdegree edges. Has little computational overhead because of this. Not applicable for weighted graph.
Nepusz *et al.,*[9]	Simple dot product does not holds good. Since factor value ranges from 0 to 1, dot product results in minimal value. Not applicable for weighted graph.
Chen *et al.,*[10]	Not applicable for directed or unweighted graph
Shen *et al.,*[11]	The strength of node u in the number of communities are considered equal but practical scenario has variant in intensity of node participation.

score for the same scenario. But his work can be preferred in a sparse network environment.

5. MODULARITY APPROACHES IN COMMUNITY DETECTION

Various algorithms of community detection are analyzed based on modularity. The summarized view of the algorithms which have been discussed in Section 3 is depicted in Tables 2 and 3. It gives the review of modularity adaptation based community detection Algorithms.

Table 3
Overview of Modularity Adaptation based Community Detection

Author Name/ Algorithm	Observations
EAGLE algorithm[12]	Only applied to undirected and unweighted graph.
Chen *et al.,*[13]	Sparse graph if taken for computation cannot be solved.
Feng jiao *et al.,*[14]	This algorithm gives the measure of closeness of a node in a community.
Lee *et al.,*[15]	Not applicable for weighted network.
OCDLCE[17]	Its time complexity is O(m) with m-number of edges.
Li *et al.,*[18]	It isnot applicable to weighted graph too. And it limits the size of the social graph.
Lu *et al.,*[19]	It is the modularity approach which works for a weighted graph.
Partha[28]	It deals with edge betweeness centrality of the graph
Erika[29]	It is a heurisitic greedy approach applicable for undirected and unweighted graph.
Mingming[30]	Modularity density is computed and applicable well for directed and weighted graph.

6. CONCLUSIONS

Detecting community structure in a social network is a quite challenging problem. In social network, most of the communities are overlapped to some extent. In recent research, several algorithms are evolved to find the overlapped community in the social network. In this work, several state-of-the-art modularity based community detection algorithms with disjoint and overlapped communities are analyzed. Quantitative analysis is performed with the modularity score to infer the best available method. Chen *et al.,*[10] algorithm proves to be the best with high modularity score. But computational overhead is high since it requires belonging coefficient of each node. Newman *et al.,* modularity score is equivalent to Chen *et al.,*[10] with a limitation of detecting only disjoint communities.

Shen *et al.,*[11]'s computation of modularity is simple compared to Chen *et al.,*[10]. Few algorithms cannot be applied to directed graphs namely Newman *et al.,*[6], Chen *et al.,*[10], Shen *et al.,*[11], EAGLE algorithm[12]. Lu *et al.,*[19] have worked on graph with weights in it. This analysis can be extended for the approaches involving machine learning techniques in identifying communities. Also the analysis can be made on how these kinds of algorithms behave in a dynamic social network. Edge weights have a major role in determining the strength of node in a community. Few researchers work on considering this edge weight as a key role in community detection in the social network field.

REFERENCES

1. Ting Yuan, Jian Cheng, Xi Zhang, Qingshan Liu and Hanqing Lu. How Friends Affect User Behaviors? An Exploration of Social Relation Analysis for Recommendation, *Knowledge Based Systems*, 88:70-84, 2015.
2. Malhar Anjaria, Ram Mohana and Reddy Guddeti. A Novel Sentiment Analysis of Social Networks using Supervised Learning, *Social Network Analysis Mining*, 4(181), March-2014.
3. Jing Zhang, Zhanpeng Fang, Wei Chen and Jie Tang. Diffusion of Following Links in Microblogging Network, *IEEE Transactions on Knowledge and Data Engineering*, 27(8):2093-2106, August-2015.
4. Palla I, Derenyi I, Farkas and Vicsek T. Uncovering the Overlapping Community Structure of Complex Networks in Nature and Society, *Nature*, 435:814-818, 2005.
5. Ahn Y Y, Bagrow J P and Lehmann S. Link Communities Reveal Multiscale Complexity in Networks, *Nature*, 466:761764, 2010.
6. Newman M E J, GirvanJing Zhang M, Zhanpeng Fang, Wei Chen and Jie Tang. Finding and Evaluating Community Structure in Networks, *Physical Review E*, 69:026113, Feb 2004.
7. Leicht E A and Newman M E J. Community Structure in Directed Networks, *Physical Review Letter*, 100(11), March 2008.
8. Nicosia V, Margioni G, Carchiolo V and Malgeri M. Extending the Definition of Modularity to Directed Graphs with Overlapping Communities, *Journal of statistical mechanics: Theory and experiment*, 2009(3):03204, 2004.
9. Nepusz T, Petroczi A, Negyessy L and Bazso F. Fuzzy Communities and the Concept of Bridgeness in Complex Networks, *Physical Review E*, 77:96–107, Jan 2008.
10. Chen D, Shang M, Lu Z and Fu Y. Detecting Overlapping Communities of Weighted Networks *via* a Local Algorithm, *Physica A: Statistical Mechanics and its Applications*, 389(19):4177-4187, 2010.
11. Shen H W, Cheng X Q and Guo J F. Quantifying and Identifying the Overlapping Community Structure in Networks, *Journal of Statistical Mechanics: Theory and Experiment*, 2009(7):42–70, 2009.
12. Shen H, Cheng X, Cai K and Hu M B. Detect Overlapping Hierarchical Community Structure in Networks, *Physica A: Statistical Mechanics and its Applications*, 388(8):1706-1712, March 2014.
13. Mingming Chen, Konstantin kuzmin, Bole slaw K and Szymanski. *IEEE/ACM 4th Social Network Analysis and Applications(SNAA) workshop at ASONAM*, Beijing, China, 2014.
14. Feng jiao Chen and Kan Li. Detecting Hierarchical Structure of Community Members in Social Networks, *Knowledge based Systems*, 87:3-15, Oct 2015.
15. Conrad Lee, Fergal Reid, Aaron McDaid and Neil Hurley. Detecting Highly Overlapping Community Structure by Greedy Clique Expansion, *4th SNA-KDD workshop*, 2010.
16. Newman M E J, GirvanJing Zhang M, Zhanpeng Fang, Wei Chen and Jie Tang. Finding and Evaluating Community Structure in Networks, *Springer-Verlag Berlin Thesis*, 2013.
17. Yan Xing, Fanrong Meng, Yong Zhou, Ranran Zhou and Zhixiao Wang. Overlapping Community Detection by Local Community Expansion, *Journal Of Information Science And Engineering*, 31:1213-1232, 2015.

18. Yixuan Li, Kun He, David Bindel and John E Hopcroft. Uncovering the Small Community Structure in Large Networks: A Local Spectral Approach, *International World Wide Web Conference Committe*, 2015.
19. Zong Qing Lu, Xiao Sun, Yong gang Wen, Guo hong Cao and Thomas La Porta. Algorithms and Applications for Community Detection in Weighted Networks, *IEEE Transactions on Parallel and Distributed Systems*, 26(11): Nov 2015.
20. Zhang S, Wang R S and Zhang X S. Identification of Overlapping Community Structure in Complex Networks using Fuzzy C-means Clustering, *Physica A: Statistical Mechanics and its Applications*, 374:483–490, 2007.
21. Nepusz Y, Petrocz A, Negyessy L and Bazso F. Fuzzy Communities and the Concept of Bridgeness in Complex Networks, *Physica Review E 77*, pages 16–30, 2008.
22. Psorakis I, Roberts S and Ebden Mark. Overlapping Community Detection using Bayesian Non-Negative Matrix Factorization, *Physica Review E* 83:66–74, 2011.
23. Raghavan U N, Albert R and Kumara S. Near Linear Time Algorithm to Detect Community Structures in Large-Scale Networks, *Physica Review E*, 76(3):36-56, August 2015.
24. Lancichinetti A, Fortunato S and Kertesz J. Detecting the Overlapping and Hierarchical Community Structure of Complex Networks, *New Journal of Physics*, 11:15–33, 2009.
25. Qiguo Dai, MaoZu Guo, Yang Liu, Xiaoyan Liu and Ling Chen. MLPA - Detecting Overlapping Communities by Multi-Label Proapagation Approach, *IEEE Congress on Evolutionary Computation*, June 2013.
26. Ting Wang, Xu Qian and Xiaomeng Wang. A Hybrid Label Propagation Algorithm to Find Communities in Large-Scale Networks, *IEEE Conference on Awareness Science and Technology*, 2015
27. Xie J, Kelley S and Szymanski B K. Overlapping Community Detection in Networks: The state-of the-Art and Comparative Study, *ACM Computing Survey*, 4(43), August 2013.
28. Partha Basuchowdhuri, Satyaki Sikdar, Sonu Shreshtha and Subhashis Majumder. Detecting Community Structures in Social Networks by Graph Sparsification, *CODS 2016*, March 13-16, 2016, Pune, India.
29. Erika Griechisch and Andras Pluhar. Community Detection by using the Extended Modularity, *Acta Cybernetica archive*, 20(1):69-85, January 2011.
30. Mingming Chen, Konstantin Kuzmin and Boleslaw K Szymanski. Community Detection *via* Maximization of Modularity and Its Variants, *IEEE Transaction on Computation Social System*, 1(1):46-65, March 2014.

Chitra Devi J is currently a Research Scholar, Department of Computer Science and Engineering, SRM University, Kattankulathur, Chennai-603203. She obtained her Bachelor of Engineering from Siva Subramaniya Nadar College of Engineering and her Master of Engineering from Anna University, MIT Campus, Chennai.

E Poovammal is a Professor, Department of Computer Science and Engineering, SRM University, Kattankulathur, Chennai. She completed her Bachelor of Engineering from Madurai Kamaraj University and her Master of Engineering from Madras University. She obtained her Ph.D from SRM University. She has 20 years of experience in teaching. She has good publications in referred International Journals and Conference Proceedings.