

Detection Of Communities On Social Network Using Frequnt Pattern Mining On Incremental Data

¹Ms.Suchita Aniruddha.Chavan, ² Dr.Nandini Mukund Choudhari

¹ Assistant Professor, ² Professor ¹Information Technology, ²Computer Engineering ¹Late.Sau.SNJB's College of Engineering, Nashik, ²J.T.Mahajan College of Engineering, Jalgaon, INDIA

Abstract

Billion's of User are using online Social Networks like Facebook, Twitter, Blogs, Apps, etc and this creates a Huge amount of Information. As per market concern for Advertisement, sales, purchase, to target specific community from this social network gives high benefits. Identification of the people's behavior from their usage of internet have become a challenging task to recognized specific community from the social network. So to detect the communities, users from Social Network can be grouped as per their interest & activity traced over the network. In this paper, many methods of community detection has been covered.

Index Terms – Community Detection, frequency analysis, pattern mining.

INTRODUCTION

A huge repository are growing day by day from various Social networks which becomes a rising field for many researchers. Social media users forms their groups for communication, as per their interest of buying or sailing online products, for creating social communities etc. So the users on social networks are using many media by connecting different groups. The content of web are rapidly generated and communicated over various forums and blogs which shows continuous flow of data. This creates an opportunity by analysing the similarity on social network which result in detection of Community

Community detection has a huge importance in business such as target completion of online product's sailing, advertisement, ecommerce transaction.

I. COMMUNITY DETECTION

On Social networks, communities are of 2 types, explicit community & implicit community. Explicit community is always in the form of groups. Group based communities are recognized by characteristics such as robust, balanced, hierarchical, modular or dense. Implicit community is Single user who writes topics, articles or blogs. Most probably, algorithms of community detection shows a graph where individual is represent by node & relation between individuals are represented by edges. Member based communities are detected using graph specification i.e. Node Degree, Node Reachability, Node Similarity [1].



Fig.1 Different types of Community Detection [1]



II.LITERATURE SURVEY

2.1 Pattern Mining For Community Discovery On Network

Here, input is provided in CSV format traced by software on Social Network. This data is filter to achieve anonymous logs which gives negative affection for the execution of algorithm. On the basis of time required for access as well as usage & base on different category, all logs are analyzed. Common patterns are recognized after all possible combinations. Frequency analysis is perform on all motifs & common one mapped motif is identified. Groups are detected by checking the high score of motif and frequency analysis. Frequency of these groups are again analyzed to determine suitable one. The drawback of this method is that, experiement has been performed on limited test data, So incorrect category detection may be possible as well as data loss can be occur[2].

2.2 CNM Algorithm on Virtual Network.

In this method, virtual network is constructed from original social network. Dataset is preprocessed & virtual connection is form on the basis of similarity of nodes. On unweighted virtual network, node similarity technique is applied by CNM algorithm for the detection of communities .By this method, time required to detect communities from social network is reduce [3].

2.3 Frequent Pattern Mining

- Detection of communities using frequent pattern mining method contains 4 steps.
- 1. Preprocessing of Data: Input to the system contain neighbor table and operation user table .Then 3 threshold values are evaluated from total number of nodes, size of network, number of edges & count of operations in the network
- 2. .Mining of Frequent pattern: parallel pattern mining algorithm is used to mine a pattern which frequently occurs in a given dataset, this reduces time complexity. Homogeneous group i.e. Sequence of users are mined by applying frequent pattern mining algorithm. Each Group perform similar actions which is equal or above threshold.
- 3. Substantiation of small communities from Groups: Shortest distance is search to confirm the small groups from Communities. Group is sustainable, if shortest path distance is less.
- 4. Expansion of small community: Here, influence of groups is checked. Neighbor has more influence means similar operation perform by more neighbor of one node. In this 2 thresholds are used to expand the specific groups. Result of this method evaluated on datasets (Blogcatalog and Flickr).[4]

2.4 Social objects and link analysis

On the basis of social objects clustering and link analysis, communities are detected. Topic oriented community detection is divided into 4 module.

- 1. Data modeling of Social Network: It creates a structure of formal models from Datasets for operation.
- 2. Clustering of Social Objects: Here, clustering is done on topic. Each cluster represent topic.
- 3. Partitioning of Social Members: Members & Social objects are associated with each other. On this basis, partitioning of member is perform into topical cluster.
- 4. Analysis of Link: Link is Analyzed to detect the community in every topical cluster.
- The number of community is determined by use of modularity function.

In this approach, Experiment is performed on three real data sets namely Enron dataset, Cora dataset, Political Blogs Dataset[5].

2.5 FP-Growth Algorithm on Incremental Data

Many users are using network in their day to day life, due to this data are continuously increases in size .As new transactions are increases, it should be update with time. With respective database, frequent items mining on incremental data is required. In this paper, FP-Growth algorithm is used for incremental data mining using heap tree data structure. This system works on following basis.

- 1. Updating the Counting Table.: After Scanning the incremental database, Counting Table is updated.
- 2. Re-sequence: Counting Table is Utilize to re-arrange the sequence of items in the Counting Table, and keep the record of the re-sequenced items.
- 3. Updating the HFP-tree –Node V is Split into two nodes using linking node from counting table, then those 2 nodes are exchange and merge .Finally new transaction is added into HFP-tree[6].

Other than this, Graph mining approach can be used to extract the knowledge from community graph. It also helps to detect the communities[7].

PROPOSED SYSTEM:

As per the scenario, many methods has been developed to detect communities on social Network. All methods vary with parameters such as accuracy of the system, execution time required to detect community & its modularity.





Fig 2. Proposed System for detection of community

So, in this proposed system, Incremental data will be the Input which will filter to remove unwanted data such as name, age. Filtration helps to reduce time complexity of the system. After filtration, user will consider as node & relation among users represent by links. Graph $G=\{V,E\}$ shows the structure on which frequent patterns are identified by applying a threshold value. To recognize the frequent pattern i.e. similar operation perform by many users, modified FP-growth algorithm can be apply. This will result in specific groups. On the basis of modularity features, appropriate communities will detect.

CONCLUSION

Now a days, Community detection is become a research topic for many people. Here, various types of community detection have been cover. In Literature Survey, Different Techniques of detecting communities on social network are seen. To reduce the time complexity & achieve the modularity feature, new system is proposed to detect the communities.

References

- [1] Reza Zafarani, Mohammad Ali Abbasi, Huan Liu. Social Media Mining, CAMBRIDGE UNIVERSITY PRESS. April 20, 2014
- [2] Hafzullah, Ahmet Anl MÜNGEN, Taner TUNCER, Mehmet KAYA. Frequent Pattern Mining for Community Dedection in Web Logs, 978-1-5386-1880-6/17/\$31.00 © 2017 IEEE

[3] Kanna Alfalahi, Yacine Atif, Saad Harou.Community Detection In Social Networks through Similarity Virtual Networks,2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining

[4] Seyed Ahmad Moosavi, Mehrdad Jalali1 Negin Misaghian, Shahaboddin Shamshirband, Mohammad Hossein Anisi. Community detection in social networks using user frequent pattern mining. DOI 10.1007/s10115-016-0970-8, Springer- Verlag London 2016

[5] Zhongying Zhao, Shengzhong Feng, Qiang Wang, Joshua Zhexue Huang, Graham J. Williams, Jianping Fan. Topic oriented community detection through social objects and link analysis in social networks, doi:10.1016/j.knosys.2011.07.017

[6] Hong-Yi Chang, Jia-Chi Lin, Mei-Li Cheng, Shih-Chang Huang. A Novel Incremental Data Mining Algorithm based on FP-Growth for Big Data, 978-1-4673-9803-9/15 \$31.00 © 2015 IEEE

[7]Bapuji Rao, Anirban Mitra.A New Approach for Detection of Common Communities in a Social Network using Graph Mining Techniques. 978-1-4799-5958-7/14/\$31.00 ©2014 IEEE