Review Article **ISSN: 2393 – 9532**

International Journal of Arts and Science Research

Journal home page: www.ijasrjournal.com

https://doi.org/10.36673/IJASR.2020.v07.i02.A06

A STUDY ON SOCIAL NETWORK ANALYSIS AND DATA MINING

G. Inbasakaran*[1] and V. Shoba[1]

[1*]Department of Computer Science, SRM Arts and Science College, Kattankulathur, Chengalpattu-603 203, Tamil Nadu, India.

ABSTRACT

The conceptualization of Online Social Networks (OSNs) has been one of the most stimulating events in this century. Asocial network can be considered to be any website or web application which allows for social experience in the form of user interactions. Nowadays, such online social interactions have led to several amusing online activities including posting photos, chatting, tweeting, online shopping, etc. Many popular OSNs such as Facebook, Orkut, Twitter, Linked In and YouTube have become more popular day-by-day. Statistics reveal that currently over 75% of all Internet users are also profoundly engaged in using social media sites. Growth in usage of social media emerged mainly due to the heavy increase in the usage of smart phones, tablets, laptops, and other similar gadgets, which has become one of the defining factors in advancement of human life.

KEYWORDS

OSN, Facebook, Orkut, Twitter and YouTube.

Author for Correspondence:

Inbasakaran G,

Department of Computer Science,

SRM Arts and Science College,

Kattankulathur, Chengalpattu-603 203,

Tamil Nadu, India.

Email: murugan10021981@gmail.com

INTRODUCTION

In this new era, where social networking has gained huge popularity, social media marketing has also become a major business strategy. Social media marketing is a form of marketing that makes use of the social networking sites as a marketing tool to increase brand exposure and to expand customer reach. As a result, the vast amount of content-rich data retrieved from such OSNs is considered to be a wealth of information that is analyzed using data mining techniques for various analytical researches. This social data is generated as a result of various implicit and/or explicit user activities such as

purchasing of products, rating an item, liking an image, chatting with others, posting tweets, and so on. However, the real challenge lies in how to accurately and efficiently analyze and process such data to extract meaningful and valuable patterns or knowledge.

SOCIAL NETWORK ANALYSIS AND MINING

In this new era, where social networking has gained huge popularity, social media marketing has also become a major business strategy. Social media marketing is a form of marketing that makes use of the social networking sites as a marketing tool to increase brand exposure and to expand customer reach. As a result, the vast amount of content-rich data retrieved from such OSNs is considered to be a wealth of information that is analyzed using data mining techniques for various analytical researches. This social data is generated as a result of various implicit and/or explicit user activities such as purchasing of products, rating an item, liking an image, chatting with others, posting tweets, and so on. However, the real challenge lies in how to accurately and efficiently analyze and process such data to extract meaningful and valuable patterns or knowledge.

It is also important to note that OSN data used in various researches is noisy, enormous, scattered and dynamic. To analyze such vast, complex, and regularly changing raw data requires proper data mining techniques, commonly known as *social network mining*. Social Network Analysis and Mining (SNAM) is one of the hot topics in the area of research in data mining and it is still in its infancy. One primary issue to consider in SNAM is whether to carry out *"static structure mining"* or *"dynamic structure mining"*. In case of static structure mining, work is done using snapshots of a social network, which is reflected at a single point of time. Thus, the analysis is focused on the structural regularities in the graph structure of a static network.

INFLUENCE MAXIMIZATION

In this day and age, as OSNs have become a daily activity for majority of the Internet users, the latter to a great extent rely on making decisions based on the influence of such sites. For instance, influence propagation helps a social network user to consciously or unconsciously decide in making a choice on which product to buy, which movie to watch, which community to join, which programme to participate in, and so on. Patterns of influence can also be studied as an indication of user trust and utilized for computing trust propagation. For instance, a user "*X*" writes in her timeline about a movie that she recently watched and liked. This update gets communicated to many of her online friends who read her update. When few such friends give a comment on that update, the information gets passed on to their friends, and this becomes a chain process.

A major goal of influence maximization is *"viral marketing"* or *"viral advertising"*, which is nowadays considered a major marketing technique being adopted by every company and organization to promote their products or brands, or to create awareness about their organization. The basic principle of *viral marketing* is to initially find a set of few influential users of a particular social network, and convince those users about the goodness of a product (one way of doing so is to supply free samples of the product), so that it can generate a cascade of influence among the users' friends who will, in turn, recommend the same product to other friends, and this will result in a grand promotion of the product. But the major issue lies in initially finding the best influential users in the social network, who are considered as the *"seed set"* of the network. A similar strategy of viral marketing via social networks is also adopted in political campaigns or personalized recommendations.

LINK PREDICTION

Another common research area in SNAM is the 'link prediction' problem which considers a static snapshot of the network at a given time 't1' and predicts future links or edges of the network at a future time 't2'. The link prediction problem is a means to predict the user as to 'who can be a possible friend' of that particular user and accordingly provide suggestions for the same.

RECOMMENDER SYSTEMS

Recommender Systems (RS) provide recommendations to users about a set of articles or services they might be interested in. This facility in OSNs has become very popular due to the easy access of information on the Internet. A few important applications of RS are its use in several e-commerce sites, such as Amazon, Flipkart and Firstcry, for recommendation of items such as movies, books, gadgets, and jewellery. The data required for providing recommendations can be obtained explicitly based on users' ratings or comments, or implicitly by monitoring user activities, such as items checked, books bought, audios heard, web sites visited, and so on. RS may also use demographic information of users like occupation, gender, or age for clustering group of users that may have similar likings.

It is a challenge for every e-commerce site to correlate consumers with the most suitable products which, in turn, enhance customer satisfaction and loyalty. Hence, the majority of the e-commerce sites have become interested in RS, which provide personalized recommendations to each visited user of a site by examining patterns of user interests in products that suggests a user's taste.

COMMUNITY OR GROUP DETECTION

In general, community or group detection in OSNs is based on studying the social network structure to find individual nodes in the network that correlate more with each other than with other related group of users.

Discovery of such communities lead to intra-communities (users or nodes belonging to same community) that are more likely to be connected or associated compared to inter-communities (users or nodes belonging to different community). Such kind of clustering in groups helps to further make estimation about the users in the network, regarding his/her likes and interests, tastes and future activities. This is turn, will help in assessing the probability of which products he/she would buy, which songs or movies he/she would watch, which services he/she may be interested in, and so on.

EXPERT FINDING

OSNs consist of a pool of experts for different related tasks and there exists social relationships among these experts. For a given task, the expert finding problem tries to capture the list of all related experts for the task T from the entire social network. Thus, the main role of expert finding is to undertake the task of retrieving a ranked list of experts who are well conversant on a given topic. Experts from OSNs are required for a diversity of reasons such as, questions answering and problem solving. This task of expert finding has gained a great deal of research interests for the information retrieval community over the past few years.

Figure No.1: Key research issues in online social network analysis

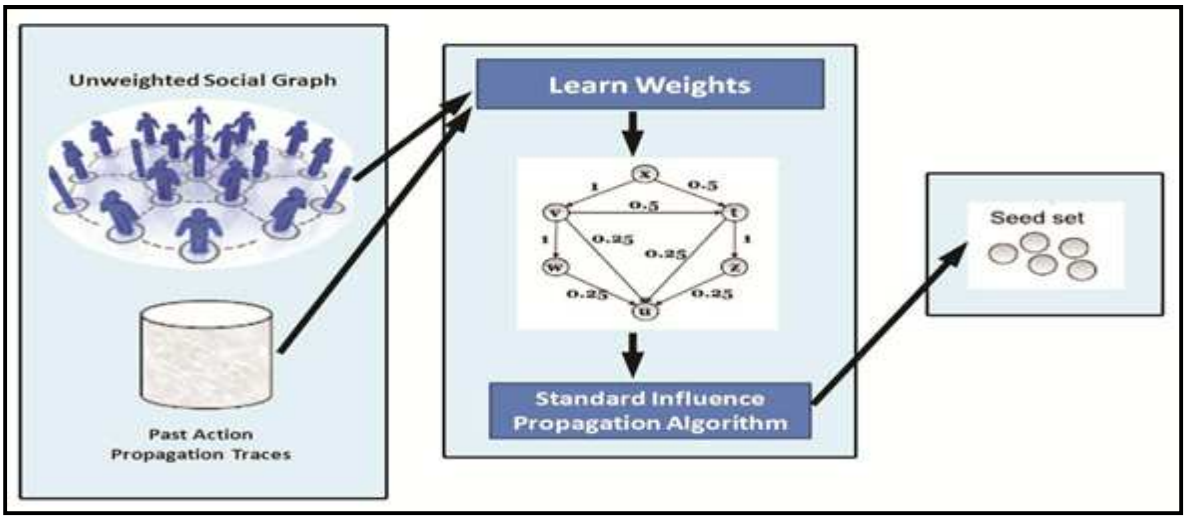

Figure No.2: The generic influence maximization framework

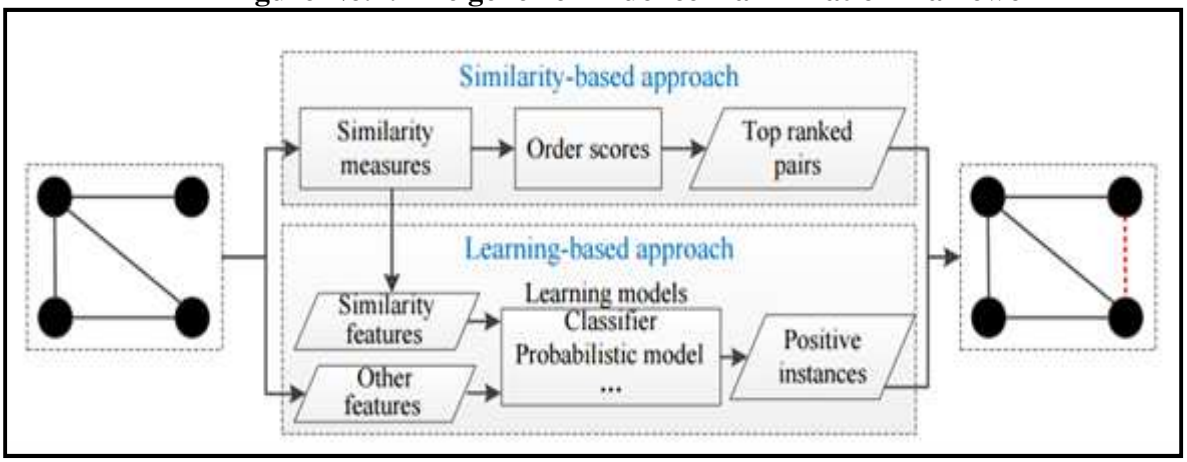

Figure No.3: The generic link prediction framework

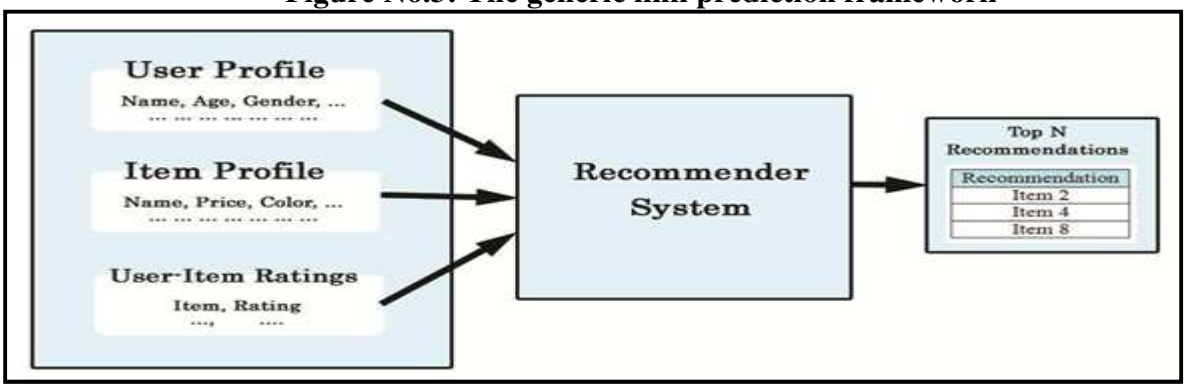

Figure No.4: The generic recommender system framework

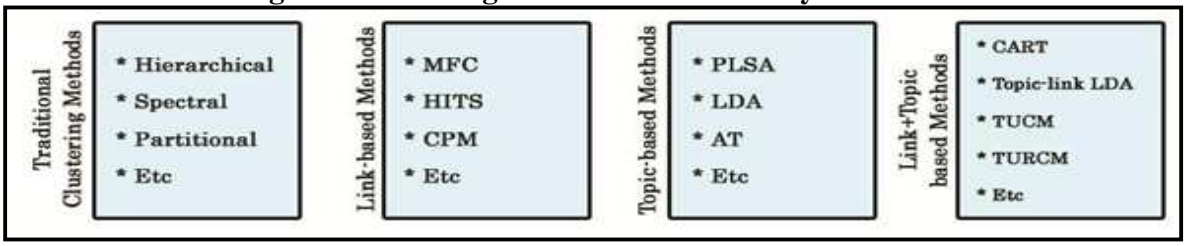

Figure No.5: various categories of community detection methods

CONCLUSION

Hence, the more effortlessly we can handle a complex social network, the better we can analyze the network, keeping in mind the dynamic growth of the network in terms of linkage structure and contextual information, that provides as input to researchers in this field of social network mining to carry out extensive research work for further development.

ACKNOWLEDGEMENT

The authors which to express their sincere gratitude to Department of Computer Science, SRM Arts and Science College, Kattankulathur, Chengalpattu-603 203, Tamil Nadu, India for providing necessary facilities to carry out this review work.

CONFLICT OF INTEREST

We declare that we have no conflict of interest.

BIBLIOGRAPHY

1. Diego García-Saiz, Camilo Palazuelos, Marta Zorrilla. Data mining and social network analysis in the educational field: an application for non-expert users, *Springer International Publishing, Educational Data Mining: Applications and Trends,* 2014, 411-439. www.datamining and social network.com.in.
2. www.networkanalysis.in.
3. www.google.com.
4. Umair Sajid Hashmi, Arsalan Darbandi, Ali Imran. Enabling proactive self-healing by data mining network failure logs, *2017 International Conference on Computing, Networking and Communications (ICNC),* 2017. www.dataminingnetwork.

Please cite this article in press as: Inbasakaran G and Shoba V. A study on social network analysis and data mining, *International Journal of Arts and Science Research,* 7(2), 2020, 50-54.