Community Detection Approaches in Real World Networks: A Survey and Classification

Pooja Wadhwa, Computer Engineering Department, Netaji Subhas Institute of Technology, New Delhi, India

M.P.S Bhatia, Computer Engineering Department, Netaji Subhas Institute of Technology, New Delhi, India

ABSTRACT

Online social networks have been continuously evolving and one of their prominent features is the evolution of communities which can be characterized as a group of people who share a common relationship among themselves. Earlier studies on social network analysis focused on static network structures rather than dynamic processes; however, with the passage of time, the networks have also evolved and the researchers have started to focus on the aspect of studying dynamic behavior of networks. This paper aims to present an overview of community detection approaches graduating from static community detection methods towards the methods to identify dynamic communities in networks. The authors also present a classification of the existing dynamic community detection algorithms along the dimension of studying the evolution as either a two-step approach comprising of community detection via static methods and then applying temporal dynamics or a unified approach which comprises of dynamic detection of communities along with their evolutionary characteristics.

Keywords: Community Detection, Dynamic Communities, Social Network Analysis, Static, Virtual Communities, World Wide Web

INTRODUCTION

In the world of rapidly evolving and emerging technologies, there has been a remarkable growth in online social networks. An online social network is a virtual community consisting of social structures made up of nodes/vertices that represent individuals and links that represent the relationships between them. The network paradigm provides a formal way of representing data and its associations, and forms the basis of any analysis. Today almost anything can be represented in the form of networks. The plethora of different problems, ranging from social interactions to World Wide Web have lead to an interdisciplinary approach

DOI: 10.4018/ijvcsn.2014010103
to these problems using tools from statistical mechanics, computer science to behavioral sciences and sociology. Social Network Analysis (SNA) (Wasserman & Faust, 1994) is a set of powerful techniques to identify social roles, important groups and hidden structures in organizations and groups. SNA has a long and successful history within sociology; networks are everywhere and related methodology can be used to analyze a wide variety of problems.

Community detection has been an active area of research and has proved to be very useful in understanding behaviour of entities involved in networks. Nodes belonging to a tight-knit community are more likely to have some properties in common. For example in the World Wide Web, community analysis has uncovered thematic clusters, in biochemical networks community detection may result in identification of functional groups. Community detection aims to identify modules and their boundaries which further results in the classification of vertices, according to their structural position in the modules. This further results in identification of the most ‘central’ node in a cluster i.e. the one which shares a large number of edges with the other partners in a group; vertices lying at the boundaries between different clusters act as a ‘bridge’ between different clusters and can be said to lead the relationships and exchanges between different communities.

Another important aspect related to community structure is the hierarchical organization of most of the real world networks. Real networks usually consist of many communities which are in turn composed of many smaller communities each having distinct characteristics. In many real world networks, one may observe that an edge may belong to more than one group. In this case, the problem of community detection leads to the detection of overlapping communities.

The dynamic nature of networks gives rise to evolving communities and a new area of research known as dynamic community detection emerges. We may observe that at a given instance of time a network may have no community structure at all, however, later on one may observe the emergence of a community structure which may mature after some time. So, the study of this dynamism provides an insight into the network evolution with time. Conversation networks like Twitter, Facebook may result in such evolving community networks where one may observe maturing of a community only when a topic propagates to many users and they become inspired or excited and participate in related discussions over a passage of time.

This paper explores the aspect of community detection in online social networks emanating its roots from the traditional methods such as graph partitioning, hierarchical clustering, partitional clustering and the advancements made with time. The aspect of dynamic community detection and the issues related to dynamic community detection will be the focus of the paper. We present a classification of the available research on the basis of the approaches followed in static as well as dynamic community detection which will help users to choose effective techniques for community detection.

SOCIAL NETWORK ANALYSIS

The main aim of Social Network Analysis is to understand the relationship between the actors involved in an interaction. Social networks are said to have been evolved in sixties, with the history deep rooted in social sciences, when Travers and Milgram (1969) proposed “The Small World Problem”. Social Networks are characterized by three prominent key features (i) “Small World Effect” (ii) “Scale Free Distributions” (iii) “Emergence of a community structure” (Ferrara & Fiumara, 2011). “Small World Effect” puts into evidence that there exists a relatively short path connecting any pair of nodes within the network (Travers & Milgram, 1969). Another important feature of Social Networks is the degree distribution of nodes within the network which specifies how the nodes within a network are interconnected. In particular, many large scale networks like World Wide Web (Albert, Jeong & Barabasi, 1999) follow power law and are free from any
Related Content

The Impact of Online News Consumption on Young People’s Political Participation
[www.irma-international.org/article/the-impact-of-online-news-consumption-on-young-peoples-political-participation/112100/](www.irma-international.org/article/the-impact-of-online-news-consumption-on-young-peoples-political-participation/112100/)

Virtual Social Networks: Toward A Research Agenda
[www.irma-international.org/chapter/virtual-social-networks/39849/](www.irma-international.org/chapter/virtual-social-networks/39849/)

Facebook as a Marketing Tool: An Analysis of the 100 Top-Ranked Global Brands
[www.irma-international.org/article/facebook-as-a-marketing-tool/132866/](www.irma-international.org/article/facebook-as-a-marketing-tool/132866/)

Viewing Cybercommunities through the Lens of Modernity: The Case of Second Life
[www.irma-international.org/article/viewing-cybercommunities-through-the-lens-of-modernity/85358/](www.irma-international.org/article/viewing-cybercommunities-through-the-lens-of-modernity/85358/)

Use Cases and Application Purposes of Social Media in Healthcare
[www.irma-international.org/chapter/use-cases-and-application-purposes-of-social-media-in-healthcare/130407/](www.irma-international.org/chapter/use-cases-and-application-purposes-of-social-media-in-healthcare/130407/)