Emotional community detection in social networks☆

Andreas Kanavos, Isidoros Perikos, Ioannis Hatzilygeroudis*,
Athanasios Tsakalidis

Computer Engineering and Informatics Department, University of Patras, Hellas, Greece

ABSTRACT

It is vastly acknowledged that analyzing social networks is a very challenging research area. Take as a striking example the organization of vertices in clusters, with many edges joining vertices of the same cluster and comparatively few edges joining vertices of different clusters. This comprises a fundamental aspect, which concerns the detection of user communities. In certain fields such as sociology and computer science where interactions and associations are often represented in the form of graphs, detecting communities is of vital importance. This paper addresses the need for an efficient and innovative methodology for community detection that will also leverage users' behavior on emotional level. Ekman emotional scale is the key point with which the methodology analyzes user's tweets in order to determine their emotional behavior. Consequently, the derived communities are estimated with the use of three different metrics, while the weighted version of a modularity community detection algorithm is utilized. There is substantial evidence indicating that our proposed methodology creates influential enough communities.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

Over the past few years, Twitter alongside other social media has had an increasing popularity. This in return, has generated an enormous research concernment as well as new opportunities for studying the reciprocal actions of different groups of people. Community detection and sentiment analysis are two such instances, comprising a popular topic in the investigation and better understanding of social networks. As for community detection on the one hand, it tries to analyze a social network with the capital objective of detecting clusters of associated and related users in it, while on the other hand sentiment analysis endeavors to settle upon the users' behavior on emotional level and consequently specify their attitude on a diverse number of topics, such as to recognize how individuals feel.

Being of the utmost fundamental aspects of the social network analysis, the determination of user behavior in each one of the arising communities as well as in the whole network is a vibrant concept of analyzing the exact way that users are associated for creating social communities. In order to explain social dynamics of interaction among groups of individuals, it is imperative that we study the community structure of a network. Based on this principal, there are research efforts in the literature that point to this direction [5]. A topic of extremely high research interest with wide range applications is the efficient analysis and the accurate specification of communities.

Community detection approaches could greatly assist in the economical and marketing domain. The accurate discovery of concrete communities and the analysis of each one could ameliorate the performance of advertising initiatives of the marketing industry by specifying and addressing the appropriate groups of users in a specific network. In this line, an effective way to offer to users recommendations adapted to their behavior and interests would be to utilize aspects of the structure of users' behavior and use the affiliations of the communities of the users [8].

People's decisions and attitudes as well as their social relationships can be affected and shaped by public stance in social networks, and since public opinion and emotions are essential to all aspects of human lives, their recognition is quite important [28]. Recognizing emotional status in user generated content in social networks is a very challenging and interesting topic in social networks and in microblogging area [3]. The analysis is necessary for understanding people status and for providing a number of indicative factors regarding public attitude towards different events. In this line, the accurate specification of the emotional content of users can potentially describe the emotional status of a community, a town or even a whole country [24].

Although users' emotional behavior as an important parameter for understanding public behavior most of the existing research efforts and methods for specifying structured communities in social networks do not utilize it. However, emotional aspects of people behavior and their opinions could provide indicative factors and assist in detecting more concrete and structured communities in terms of density.

In the present work, the main contributions concern the following aspects: Initially, a method for the analysis of user generated content and the specification of users' emotional behavior on Ekman's psychometric scale [6] is presented. Furthermore, a method for examining users' actions and their posts in social networks while calculating their influence based on their behavior and modeling a conversation as an emotional graph, is introduced. Finally, we present a method for specifying the most influential communities of users that are formulated based on each user's emotional behavior and their analytics profile in the network. Specifically, each user profile can be considered as the union of the two aforementioned characteristics, i.e. emotional profile and analytics profile.

The rest of this manuscript is structured as follows: Section 2 presents background topics in emotional detection and sentiment analysis as well as community detection. Section 3 presents our proposed methodology while in Section 4, details of the implementation are introduced. Furthermore, in Section 5, the evaluation study conducted and the results gathered regarding the community detection proposed schema are presented. Finally, Section 6 concludes the paper and provides directions for future research.

2. Related work

Community analysis in social networks has gained the interest of scientists in multiple areas such as graph theory, social network analysis, graph clustering algorithms, as well as web searching algorithms [9,15,16,20]. A community can be considered as a group of network nodes where links connecting these nodes are dense [30]. Moreover, another characteristic of a community is that it corresponds to a group of nodes on a graph or even a network that share common properties. Community detection is the problem of identifying structures of grouping of nodes which demonstrates high coupling and low cohesion. For a complete overview of wide used approaches and techniques, one can consider works presented in [21,23].

Concerning communities, the well known problem of graph partitioning draws a vast amount of attention. Concretely, in [10], an algorithm that is presented for identifying the edges lying between communities and their successive removal can be considered as a breakthrough in this area. In this algorithm, a procedure leads to the isolation of the communities after some iterations. One should also mention techniques that use the modularity metric, which designates the density of links inside communities against the density of links outside communities as proposed in [9,15,20]. The most popular approach regarding the modularity metric is the algorithm introduced in [1].

During the last years, sentiment analysis methods that can accurate recognize emotional status of users in social networks are of extreme interest [18] and many studies indicate the important role they can play in the analysis of users behavior and also the analysis of public attitude towards events and topics.

The automatic recognition of emotional presence in tweets can assist in creating more sophisticated social and personal applications as well as in the study of social relations [25]. In the work presented in [17], authors study emotion recognition in tweets; they developed a corpus of 5.553 tweets manually annotated with 28 emotion categories and in following they study the performance of machine learning algorithms in fine grained classification. The results indicate that classifiers can detect 28 emotion categories in text without a huge drop in performance compared to coarser - grained classification schemes. In addition, in [2], a knowledge based approach for recognizing emotions in text is examined and a tool is presented. This tool utilizes versions of SentiWordNet lexical resource, which is a subset of WordNet-Affect with some manual added words. Specifically, this corresponding tool detects contrasts between positive and negative words that shift emotion valence. Ferrada and Camarinha-Matos [7] present an approach for developing an emotions-oriented system for assisting in monitoring and managing the interactions among users in collaborative networks. The supervision of the emotional interactions within collaborative networks provides multi-modal emotional input for achieving awareness of the participants and also offers mechanism to promote the emotional health of the network.

Xu et al. [29] present two methods for detecting communities in social networks. In the first method, sentiment can be positive or negative, while the second method assumes that the range of sentiment is divided into intervals and then users are categorized into groups based on the differences in the ranges of sentiment values. Moreover, Deitrick et al. [5] utilize