Available online at www.elixirpublishers.com (Elixir International Journal)

Computer Engineering

Elixir Comp. Engg. 78 (2015) 29540-29543

ISSN: 2229-712X

User group model for mobile social based networks

Amol S Nalge and S.R. Durugkar
SND COE, 1,YEOLA, Nashik, India.

ARTICLE INFO

Article history:
Received: 24 November 2014;
Received in revised form:
21 December 2014;
Accepted: 1 January 2015;

Keywords
Overlapping community detection,
M2clustering,
HM2clustering,
Clustering.

ABSTRACT
Communities with different profiles shows the interests of community members. User and check-in venue details are used to cluster. Multimode multi-attribute edge-centric coclustering model is help to discover overlapping the communities. Overlapping communities is used to repair by replacing each edge with its vertices in edge clusters. Intermode and intra mode features are helped to us for making the process. Three intra mode features are used in the community detection process. M2 Clustering algorithm is used for community detection it is Edge clustering based on k-means and HM2 Clustering algorithm is used to detect overlapping communities of LBSNs and also called as Two-step hierarchical edge clustering. The overlapping community detection mechanism is enhanced with recommendation models. LBSNs, analyzing 1) the data source used, 2) the methodology community. 3) the objective of the community. The community are also classified with location. The system is enhanced with feature selection and feature fusion mechanism. This system also indicate the cluster of the two communities.

Introduction

The social network is a theoretically useful in the social science to study relationships between individuals, groups, or society. The term is used to describe a social structure determined by such interaction [4]. The study of the structures uses social network analysis to identify local pattern and examine it. We propose taxonomies that partition of the systems according to the properties.

Social networks and the analysis of them and divide it to three type that is social psychology, sociology and statistics. It is early structural theories in to sociology emphasizing. the Jacob Moreno is credited on web of group affiliations with developing the first paper in the 1930s to compare the relation. These approaches were mathematically formalized and theories in the 1950s .social phenomena should be primarily conceived and investigated through the properties of relations between themselves Social network analysis is now one of the major part of the sociology, and is also member of the social science. Together with other complex networks, it forms part of the related field of social network .this is need to create the components for social science. The social network is used to describe a social structure determined by such component as well as the interaction between themselves.

We further envision the rapid development of cross-space communities in recent years which gap between the human interactions and virtual world virtual world means merging social elements in online social networks with contexts in offline communities .significant means location-based social networks (LBSNs, facebook ,twitter). It is also work as clustering between the two communities .suppose we want the cluster of face-book user and the another of the twitter use people then find out the another cluster from the other than the face-book and the twitter user from the dataset using the foursquare API. find the online as well as offline communities then we can formulate it on the cluster.

Tele: 9225967081
E-mail address: nalgeamol97@gmail.com

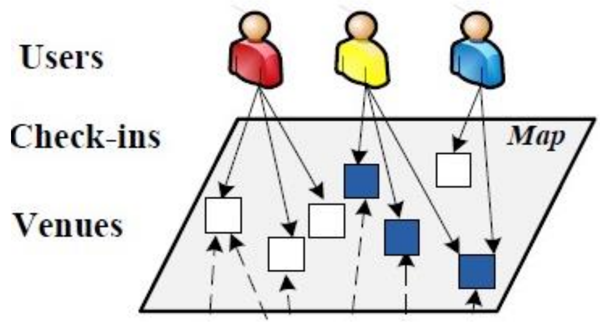


Fig. 1. User-venue check-in network example.

Related Work

In this section, we briefly review the related work that can be classified into three categories .

The first category contains the understanding the user behaviors based on LBSNs. [7],[8] analyzed the social, geographic and Geo-social properties of four social networks (Facebook, Foursquare, orcut,and Twitter). Noulas *et al.* [9] find out the user dynamics checkin and the find out the pattern of the Foursquare. Cheng *et al.* [10] studied the not stable patterns of Foursquare users and similar factor affect on the people. Vasconcelos *et al.* [11] analyzed how Foursquare users uses the tree features (i.e., tips, dones, and to-dos). We can studied the group of profiles in LBSNs. Li *et al.* [12] proposed two different clustering a find out the different communities. Noulas *et al.* [12] used a different clustering algorithm to group Foursquare data set users based on the above categories of venues they had checked in, aiming at find out the communities and find out the type of activity in each region of a city. Although we mentioned studies offer important of the properties of user interactions in LBSNs, Our work aims to form the cluster of communities in an overlapping manner.

The second category involves the work on community detection in the social network , [2]. In order to detect communities from a social network, one typically chooses an objective function based on the cluster is a collection of data with internal connectivity than external connectivity, and then applies cluster algorithms to extract clusters by optimizing the objective function. In general, community detection can be classified into two categories: overlapping and non-verlapping approaches. Some popular methods are modularity maximization [4], [5], Girvan .Newman algorithm [1], Louvain algorithm [6], clique percolation [7], link communities [8], etc. As users in LBSNs form the sociel network form communities, owe can not apply community detection based solely on the network links to generate the communities.

The third category focuses on community detection by considering communities, which are the closest to our work. Several existing works on community form the clustering fall into this category. The main idea is to design a cluster measure from soucre that combines both clustre information . Based on this measure, standard clustering algorithms such as *k*-means and edge clustering are then applied to form cluster . where the weight of each edge is defined as the number of attribute values shared by the two end nodes. The authors applied graph clustering algorithms on the constructed matrix to perform clustering. The state-of-the-art distance-based approach is the SA-cluster that defined a unified distance measure to combine structural and attribute similarities. Communties data and edges are added to the original graph to connect nodes that share the community, which is form the nearest cluster . Afterward, a clustering algorithm cluster is proposed based on the *k*-means method. However, all these works in the third and last category attempted to optimize the two communities; thus, the communities detected were not exactly. In this paper, we propose community link between users and which venue it is located venues, and find out the community structure. Specifically, we formulate the community detection problem into edge clustering issue, viewing both inter mode links and intra mode attributes for clustering. With this we can conclude that the community obtain for the using the clustering algorithms.

Problem Statement

In this paper, a community is defined as a cluster of edges (i.e., check-ins) with user and venue as two modes. We use $x = (x1, x2, \dots, xm)$ to represent the user set, and $y = (y1, y2, \dots, yn)$ to denote the venue category set. A community $Ci(1 \le i \le k)$ is a subset of users and venue categories, where k is the number of communities. On one hand, the check-in network between users and venue categories form a matrix M, where each entry Mij€[0.∞] corresponds to the number of check-ins that Ui has performed over vi . Therefore, each user can be represented as a vector of venue categories, and each venue category can be denoted as a vector of users. On the other hand, users and venue categories might have several independent attributes, denoted as $(ai1, ai2, \dots, ax)$, and $(bj1, bj2, \dots, biy)$, respectively. Normally, every attribute reveals a certain social aspect of users or venue categories. For instance, a user has a certain number of followers and followings in Foursquare, and a venue category has a common operating time. Therefore, both the user mode and the venue mode have two types of representations: an inter-mode representation and an intramode representation. Based on the above notations, the overlapping community detection in LBSNs can be formulated as a multimode and multi-attribute edge-centric co-clustering problem as follows.

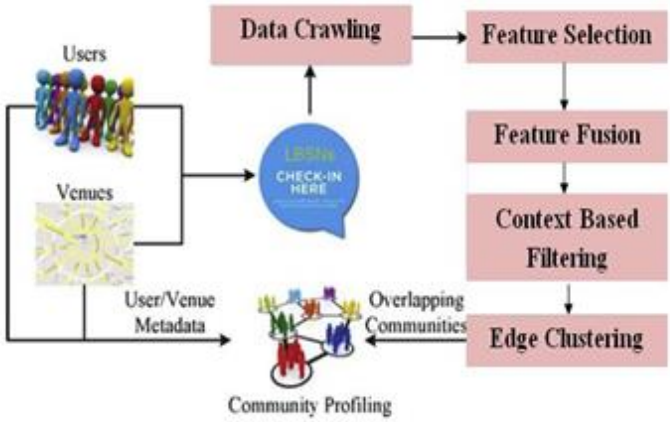


Fig. 1. Community discovering and profiling framework

***Input*:**

1) A check-in matrix $M(|x|\times|y|)$, where $|x|$ and $|y|$ are the numbers of users and venue categories, respectively.

2) A user attributes matrix $M(|x|\times|A|)$, where $|A|$ is the number of user attributes.

3)A venue category attributes matrix $M(|x|\times|B|)$, where $|B|$ is the number of venue category attributes.

4) The number of communities k, which is optional based on the clustering algorithm.

***Output*:**

k overlapping communities that consist of both users and venue categories.

Multimode & Multi-attribute Edge Clustering Framework

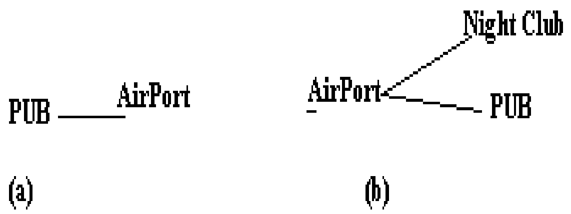
The key idea of the proposed community discovering and profiling framework is shown in Fig. 1. First, features are collected LBSNs data set and then feature normalization and fusion are performed. Second, the find out the community using edge clustering algorithm. Finally, by combining the detected communities together with user in meta data.

While the selecting of clustering (i.e., communities) has basic task of the finding network, few challenges to find out the community. Meanwhile, to find out the communities is very important to understand the features of each community. Thereby, a large systematic community profiling is used for community. With the recent surge of location-based social networks (LBSNs, e.g., Foursquare, Face book Places), huge amount of digital footprints about users' locations, profiles as well as their online social connections provide sufficient meta data for community profiling. Different from social networks (e.g., Flickr, Face book) which have explicit groups for users to subscribe or join, LBSNs usually have no explicit community structure. In order to capitalize on the large number of potential users, quality community detection and profiling approaches are needed so as to enable applications such as direct marketing, group tracking, etc. In this paper, based on the user-venue check-in relationship and user/venue attributes, we come out with a novel community profiling framework. Specifically, we first adopt edge-clustering to simultaneously group both users and venues into communities, and then based on the rich meta data of users and venues we put forward a quantitative community profiling mechanism to indicate the preferences, interests and habits of a community. The efficacy of our approach is validated by intensive empirical evaluations using the collected Foursquare data set of 266,838 users with 9,803,764 check-ins over 2,477,122 venues worldwide Graphs have been widely used to model relationships among data. For large graphs, excessive edge crossings make the display visually cluttered and thus difficult to explore. In this paper, we propose a novel geometry-based edge-clustering framework that can group edges into bundles to reduce the overall edge crossings.

Our method uses a control mesh to guide the edge-clustering process; edge bundles can be formed by forcing all edges to pass through some control points on the mesh. The control mesh can be generated at different levels of detail either manually or automatically based on underlying graph patterns. Users can further interact with the edge-clustering results through several advanced visualization techniques such as color and opacity enhancement. Compared with other edge-clustering methods, our approach is intuitive, flexible, and efficient. The experiments on some large graphs demonstrate the effectiveness of our method. Using the sim form we can also find the using the [5].sim formula.

A. Feature Description

The intermode is describes the structure community between a check- ins user and the other users.. According to [5], we adopt two intermode features (i.e., user-venue and venue-user) in this paper, where each user is equal to vector of venue categories and each venue is denoted as a vector of users. The intramode feature depicts similar attributes . Where attribute is belong to the user venue or check ins.. We select three



intramode based on Foursquare data .

Fig. 2. Tag clouds of two Foursquare users from London. (a) Tag cloud of user A. (b) Tag cloud of user B.

1.Intermode Feature User-Venue Similarity:-

a) from figure 2 we can describe as Foursquare classifies venues into sub-categories .

b)we adopt cosine similarity to calculate the user-venue

2)Intermode Feature Venue-User Similarity:-

a) venue category of Foursquare can be represent as a vector by treating users as features as well .

b)Venue temporal similarity

3)Feature Normalization and Fusion

a) we simply normalize each similarity measure sim x into the interval [0, 1] .

b) 1.Intramode Feature: User Social-Influence Similarity:

2.Intramode Feature: User Geo-Span Similarity

3.Intramode Feature: Venue Temporal Similarity

Clustering Algorithm

on the above we can describe as the multimode edge clustering problem is converted into an ordinary cluster, which can be handled by using different type of clustering algorithms.

There are two different type of the clustering algorithm : 1)M^2 edge clustering algorithm:-

Input:

• *E*, an edge list $\{e_i | 1 \le i \le n\}$

• *k*, the number of communities

• *Mu*, the user-user similarity matrix

• *Mv*, the venue-venue similarity matrix

Output:

• *C*, a set of detected communities

1: *k* edges are randomly selected .using the sine algorithm and clusterin form algorithm .

1)HM^2 edge clustering algorithm:-

Input:

• *E*, an edge list .

• *K*, a large number which is *k*

• *MX* the user–user similarity matrix

• *MY* the venue-venue similarity matrix

Output:

• *D*, an edge dendrogram

Performance Evaluation

Data Collection

Foursquare API provides less action taken on the check-in information; therefore, we used o Twitter to form of clusters, dataset of Foursquare choose users from that to obtain the cluster and publish it . Our data collection started from October 24th, 2011 and lasted for eight weeks, which results in a raw dataset of more than 12 million checkins performed by more than 700 000 users over 3 million venues. In the meantime, we also search and fin the relevent data to users and venues, to show the information of each venue.

Experiment Setup

To increase the performance of the proposed framework, we chose three large cities (i.e., London, Los Angeles, and New York). Then we formulated it on the new cluster. First, we delete or ignore the check-ins that are frorm the other venue, where other venue is used from the datset using the Foursquare API. Second, use the only one for the create the another cluster.

Table 1

Different Feature Sets Evaluated in the Experiments

FEATURE SET	USED FEATURE
I	USER VENUE WHICH ARE IN THE SAME EDGE CLUTERING
II	VENUE THAT ARE TEMPORARY SIMILAR
III	USER SOCIAL INFLUENCE GROUP
IV	USER GEO SPAM SIMILARITY
V	FIND THE FINAL CLUSTER.

Check-ins means find out the region of less clustering form, which means not active users come with their region are sort out . Finally, users who used this algorithm are to form the large datad-set (checkin speed faster than than 1200 km/h, which is the common airplane speed) and check-ins from these users are eliminated as well.

Benchmark

In this paper, we conducted a series of experiments to evaluate the performance of two different algorithms (i.e., M2 Clustering and HM2 Clustering) shown in Table I

Quality of the Detected Communities

From the above figure we can denote as the find the new community pool area from the bar .using the M2 Clustering and

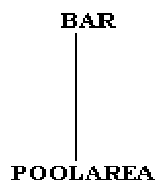


Fig4. Tag cloud of one topic in the constructed topic model

HM2 Clustering algorithms.

Conclusion

In this paper, from the user-venue check-in network and user attributes, we proposed a multimode multi-attribute edge clustering framework to detect overlapping communities for LBSNs users.

Experimental results showed that the proposed framework was able to discover the communities from different user venue and at several community, which can be used to different applications, such as more than one clustering we reported several cluster to findings the community. The fundamental studied suggested several interesting problems by exploring this providing a framework to guide the selection and fusion from different cluster features is one direction to work on. The proposed community detection community can also help the study of friend place also .

Acknowledgment

The authors would like to thank all of their colleagues for their discussions and suggestions**.**

References

[1] M. E. J. Newman and M. Girvan, “Finding and evaluating community structure in networks,” *Phys. Rev. E*, vol. 69, no. 2, pp. 26 113–26 127, 2004.[

2] S. Fortunato, “Community detection in graphs,” *Phys. Rep.*, vol. 486, nos. 3–5, pp. 75–174, 2010.

[3] Y.-Y. Ahn, J. P. Bagrow, and S. Lehmann, “Link communities reveal multiscale complexity in networks,” *Nature*, vol. 466, no. 7307, pp. 761–764, 2010.

[4] J. D. Cruz, C. Bothorel, and F. Poulet, “Entropy based community detection in augmented social networks,” in *Proc. IEEE CASoN*, 2011, pp. 163–168.

[5] X. Wang, L. Tang, H. Gao, and H. Liu, “Discovering overlapping groups in social media,” in *Proc. ICDM*, 2010, pp. 569–578.

[6] I. S. Dhillon, “Coclustering documents and words using bipartite spectral graph partitioning,” in *Proc. KDD*, 2001, pp. 269–274.

[7] S. Scellato, C. Mascolo, M. Musolesi, and V. Latora, “Distance matters: Geo-social metrics for online social networks,” in *Proc. WOSN*, 2010, p. 8.

[8] S. Scellato, A. Noulas, R. Lambiotte, and C. Mascolo, “Socio-spatial properties of online location-based social networks,” in *Proc. ICWSM*, 2011, pp. 329–336.

[9] A. Noulas, S. Scellato, C. Mascolo, and M. Pontil, “An empirical study of geographic user activity patterns in Foursquare,” in *Proc. ICWSM*, 2011, pp. 570–573.

[10] Z. Cheng, J. Caverlee, K. Lee, and D. Z. Sui, “Exploring millions of footprints in location sharing services,” in *Proc. ICWSM*, 2011, pp. 81–88.

[11] M. A. Vasconcelos, S. Ricci, J. Almeida, F. Benevenuto, and V. Almeida, “Tips, dones and todos: Uncovering user profiles in Foursquare,” in *Proc. WSDM*, 2012, pp. 653–662.

[12] N. Li and G. Chen, “Analysis of a location-based social network,” in *Proc. CSE*, 2009, pp. 263–270.