Computers and Electrical Engineering 39 (2013) 2521–2530

Contents lists available at ScienceDirect

Computers and Electrical Engineering

journal homepage: www.elsevier.com/locate/compeleceng

The learning system of collective behavior in students' social network ☆

Ying Xie, Bin Luo, Rongbin Xu *

Computers Studies Department, Anhui University, Hefei 230601, China
The School of Computer Science and Technology, Anhui University, Hefei 230601, China

ARTICLE INFO

Article history:
Received 27 February 2013
Received in revised form 30 September 2013
Accepted 2 October 2013
Available online 19 October 2013

Keywords:
Collective behavior
Social network
Dimensionality refinement
Feature selection

ABSTRACT

The rapid development of social networking sites brings about many data mining tasks and novel challenges. We focus on classification tasks with students' interaction information in a social network. To mitigate the difficulties of developing a learning system, this study proposes a new computing paradigm: spectral clustering as a service, providing a service to enable exacting social dimensionality on demand. Spectral clustering has been developed in a social network dimensionality refinement model as a kernel middleware, namely SNDR. The SNDR service can process the sparse information, explore the network's topology and finally exact suitable features. Experimental results justify the design of Collective Behavior Learning System and the implementation of the Social Network Dimensionality Refinement model's service. Our system makes better performance than baseline methods.
© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

Today's computation and communication technologies enable students to share information, express their thoughts and connect with each other in innovative ways. Social networking sites empower students of different grades and backgrounds with novel forms of collaboration, communication, and collective behavior [1]. Great numbers of online students collaboratively do class articles at a large scope and scale; online market suppliers recommend products by investigating students' shopping behavior and interactions; and students' union activities also exploit new shapes of collective action. During this procedure, social media provide ample opportunities to study students' interactions and collective behavior on a large scale. Given a social network of students with their interaction information, we can investigate how to use this rich interaction information for classification tasks in social network [2]. In particular, given the behavior of some students understand behavioral patterns in social media. We study how networks in social media can help to predict other students' behaviors and information.

There is an example in Fig. 1. Student *a* connects to student *b* because they study in the same class, and to student *c* because they often collaborate in Tennis Club. Given the label information that student *a* is interested in C++ programming and Tennis, we can infer the labels of student *b* and *c*. Treating these two connections homogeneously, we guess that both student *b* and *c* are also interested in programming and tennis. Furthermore, if we know how student *a* connects to other students, it is more reasonable to conclude that student *b* is more interested in programming, and student *c* likes tennis.

Furthermore, let's discover the network example with two groups in Fig. 2. There is no zero degree entity. A student associates with one association if one of his connections is assigned to that association. As shown in Fig. 2, two communities can

☆ Reviews processed and approved for publication by Editor-in-Chief Dr. Manu Malek.

* Corresponding author at: The School of Computer Science and Technology, Anhui University, Hefei 230601, China. Tel.: +86 13866747580.
E-mail address: xurongbin910@gmail.com (R. Xu).

0045-7906/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.compeleceng.2013.10.001

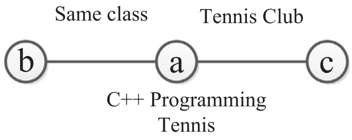

Fig. 1. A small students' connections example.

be represented by two edge sets. The disconnected edge clusters can be converted into the description of social dimensions, where an entry is 1 means the student is involved in that corresponding social dimension while 0 is not. Node *a* is associated with both groups because it has edges in both group.

For the number of one student's association is no more than that of his connections, the extracted social dimensions following edge partition are sparse [7]. In our network with *m* edges (links) and *n* nodes, each node v_i has no more than $min(d_i,k)$ non-zero entries in its social dimensions. Generally speaking, the connections in social media networks are not homogeneous. Different connections are contacted with different relations.

In this paper, our contributions are: (1) We present a learning system of collective behavior based on students' social dimensions. This system is shown to be effective in solving this problem. (2) The framework suggests a novel network semi-supervised learning model named Social Network Dimensionality Refinement. (3) This learning system can capture the latent associations of students by extracting social dimensions based on network similarity connections, then apply discriminative techniques to classification based on the extracted dimensions. (4) This system is based on sparse social dimensions, without sacrificing the prediction performance, can efficiently handle the students' networks of thousands of individuals.

This paper is organized as follows. We present the related work in Section 2. Then we formally state the learning system of collective behavior with students' network data in Section 3, and analysis the important parts of our system in Section 4. Section 5 shows the plentiful experiments, empirical results on networked data which are presented and discussed. Section 6 points out the conclusion and the promising directions for future research by using the proposed framework.

2. Related Work

2.1. Clustering methods on graph

At this time, many approaches developed for clustering on graphs for the purpose of social dimension extraction, including modularity maximization [10], latent space models [11,12], block models [13,14] and spectral clustering [5]. Probabilistic methods are also developed [15,13]. Social networking federation [16] proposes the concept as a paradigm where information on various social network systems can be seamlessly integrated. One paper [17] introduces the Automated Two-Dimensional *k*-means algorithm, a novel unsupervised clustering technique. It incorporates local and spatial information of the data into the clustering analysis.

Spectral clustering has been shown to work reasonably well in various domains, including graphs, images and text. It is also proved [18] to be equivalent to a soft version of the classical *k*-means algorithm for clustering. We adopt spectral clustering to extract social dimensions because the effectiveness in various domains and the availability of a huge number of existing linear algebra methods to help solving this problem. There are many other methods to choose. We consider this is one feature of our learning system.

2.2. Similarity construction graphs

There are many popular constructions to build similarity graphs, they can be used to transform a given set $x_1,\cdots,x_n$ of data points with pairwise similarities w_{ij} or pairwise distances d_{ij} into a graph. We construct similarity graphs to model the local neighborhood relationships between the data points.

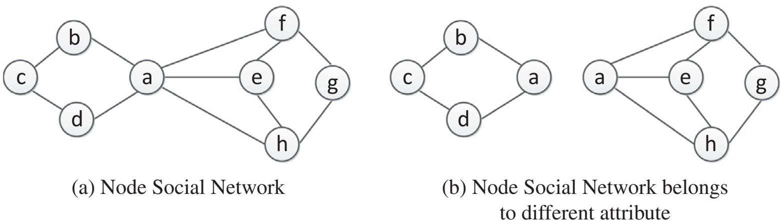

Fig. 2. Small social network example.

Download English Version:

https://daneshyari.com/en/article/455649

Download Persian Version:

https://daneshyari.com/article/455649

Daneshyari.com