Physica A 457 (2016) 270–279

Contents lists available at ScienceDirect

Physica A

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

Modeling one-mode projection of bipartite networks by tagging vertex information

Jian Qiao [a,*], Ying-Ying Meng [a], Hsinchun Chen [b,c], Hong-Qiao Huang [a], Guo-Ying Li [a]

[a] School of Management, Northwestern Polytechnical University, Xi'an 710072, China
[b] Artificial Intelligence Lab, University of Arizona, Tucson, AZ 85721, USA
[c] Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China

HIGHLIGHTS

- We propose a new method for modeling one-mode projection of bipartite networks.
- Our modeling method breaks through the limitation of traditional methods.
- Our one-mode collaboration network model outperforms available models.
- We find that five mechanisms are common and crucial to collaboration networks.

ARTICLE INFO

Article history:
Received 9 November 2015
Received in revised form 3 February 2016
Available online 4 April 2016

Keywords:
Bipartite networks
Collaboration networks
One-mode projection
Collaboration network model

ABSTRACT

Traditional one-mode projection models are less informative than their original bipartite networks. Hence, using such models cannot control the projection's structure freely. We proposed a new method for modeling the one-mode projection of bipartite networks, which thoroughly breaks through the limitations of the available one-mode projecting methods by tagging the vertex information of bipartite networks in their one-mode projections. We designed a one-mode collaboration network model by using the method presented in this paper. The simulation results show that our model matches three real networks very well and outperforms the available collaboration network models significantly, which reflects the idea that our method is ideal for modeling one-mode projection models of bipartite graphs and that our one-mode collaboration network model captures the crucial mechanisms of the three real systems. Our study reveals that size growth, individual aging, random collaboration, preferential collaboration, transitivity collaboration and multi-round collaboration are the crucial mechanisms of collaboration networks, and the lack of some of the crucial mechanisms is the main reason that the other available models do not perform as well as ours.

1. Introduction

Bipartite networks are an important class of complex networks. A bipartite network is composed of two types of vertices and the edges running only between the vertices of unlike types. Many natural, social, and technical systems can

* Corresponding author.
E-mail address: qiaojian@nwpu.edu.cn (J. Qiao).

http://dx.doi.org/10.1016/j.physa.2016.03.106

be represented as bipartite networks, such as a coauthors network [1], movie actors network [2], directors network [3], recommendation system [4], and so on [5–8].

In the past several decades, the characteristics of extensive bipartite networks have been analyzed empirically. For example, Newman [1,9–11] analyzed many statistical properties of scientific collaboration networks in the realms of physics, mathematics, biomedicine and computer science. Lambiotte and Ausloos [12] analyzed the properties of a bipartite network about people sharing their music library. Shang et al. [13] reported the empirical analysis of two large-scale web sites, in which users are connected by music groups and bookmarks, respectively. Zhang et al. [14] presented an empirical study on the Bus Route Networks of Beijing and Yangzhou, Travel Route Network of China, Huai–Yang recipes of Chinese cooked food, and collaboration network of Hollywood actors.

In addition, many new statistical indices for bipartite networks have been proposed in recent years, such as a clustering coefficient based on the fraction of cycles with size four defined by Lind et al. [15], two edge clustering coefficients based on squares and triples, respectively, proposed by Zhang et al. [16], and an index called "collaborative similarity to quantify the diversity of tastes based on the collaborative selection" proposed by Shang et al. [13].

To gain insight into the evolution of bipartite systems, many two-mode network models, such as the sexual contact network model [17], collaboration network model [18], plant–animal mutualistic network model [19], ecological and organizational network model [7], general bipartite network model [20], online bipartite network model [21], and a model for the self-assembly of creative teams [22], have been developed. Two-mode network models are natural in form and can hold information about the complete structure. Another modeling tool, the hyper graph [23], can also hold the complete structure information of bipartite networks [24,25] because each of its edges, also known as hyper edges, can relate groups of more than two vertices.

Yet, people perhaps are more interested in the one-mode projection of bipartite networks in many scenarios. For instance, for scientific collaboration networks, we usually are more interested in the relationships between scientists rather than the relationships between scientists and their publications. Because of this reason, many unweighted one-mode collaboration models have been proposed. Barabási et al. [26] proposed a model for capturing the temporal evolution of collaboration networks. A one-mode collaboration network model developed by Zhou et al. [27] interpolates between the networks that follow a power-law and an exponential degree distribution. Guimerà et al. [22] designed a model about the self-assembly of creative teams. Zhang et al. [14] suggested a model to understand the evolutionary mechanisms of four non-social systems and a social system. Because many of the informative structures have to be ignored, unweighted one-mode models could not hold the complete structure information of bipartite networks [28,29]. For instance, from the unweighted projection of a scientist–paper network, we know who are the collaborators of each scientist, but we cannot accurately tell who are the authors of each paper.

Recently, many weighted one-mode models were proposed to contain the structure information of bipartite networks more completely. In the scientist network model designed by Ramasco and Morris [30], the edges of the model are weighted by the times of collaboration. Ke and Ahn [31] proposed a weighted model for reproducing the observed pattern in scientist networks—local clusters consist of dense, weak ties and are interconnected by sparse, strong ties. Zhou et al. [32] proposed a weighting method using asymmetrical weights and self-connection to mimic the information that coauthors might assign a specific paper with different weights. Apparently, weighted one-mode models are more informative than unweighted ones, but such models still have some limitations. First, any weighting method could not exclude subjective factors completely. In some methods, the information of those vertices with one degree is even lost in the projection [32]. Second, the information contained in the models by the static weights of the links is unreliable because the link weight remains constant after being assigned to an edge [33]. Third, the evolving weighted models are time-consuming because a great number of link weights have to be updated instantly at each time step.

In this paper, we propose a new method for modeling a one-mode projection of bipartite networks that overcomes the deficiency of traditional methods. Modeling a one-mode projection of bipartite networks with our method maintains the complete structure information of the original systems, and more importantly, the projection's topological structure can be controlled very flexibly. As an application example, we design a one-mode collaboration network model with our modeling method to verify the method's feasibility and explore the crucial evolving mechanisms of some real collaboration networks.

The rest of this paper is arranged as follows. Section 2 is the introduction of our method for modeling the one-mode projection of bipartite networks. In Section 3, the statistical indicators used to measure the local and global properties of the one-mode projection of bipartite networks will be introduced. We will infer the possible evolutionary mechanisms of collaboration networks in Section 4 by observing and analyzing the local and global statistical properties of some real data to provide the design of our collaboration network model a strong basis. In Section 5, we will describe and interpret our one-mode collaboration network model in detail. The feasibility of our modeling method and the performance of our collaboration network model will be examined and discussed through numerical experiments in Section 6. Finally, we will summarize our research results in Section 7.

2. Method for modeling a one-mode projection of bipartite networks

Traditional unweighted and weighted projecting methods cannot completely contain the structure information of a bipartite network; therefore, if such methods are used to model a one-mode projection of the bipartite network, it is