Numer. Math. Theor. Meth. Appl.
doi: 10.4208/nmtma.2012.m1108

Vol. **5**, No. 3, pp. 384-402
August 2012

On Multivariate Markov Chains for Common and Non-Common Objects in Multiple Networks

Xutao Li[1], Wen Li[2], Michael K. Ng[3] and Yunming Ye[1,*]

[1] *Department of Computer Science, Shenzhen Graduate School, Harbin Institute of Technology, Shenzhen, 518055, China.*
[2] *School of Mathematical Sciences, South China Normal University, Guangzhou, 510631, China*
[3] *Department of Mathematics, Hong Kong Baptist University, Kowloon Tong, Hong Kong.*

Received 15 April 2011; Accepted (in revised version) 27 October 2011

Available online 3 July 2012

Abstract. Node importance or centrality evaluation is an important methodology for network analysis. In this paper, we are interested in the study of objects appearing in several networks. Such common objects are important in network-network interactions via object-object interactions. The main contribution of this paper is to model multiple networks where there are some common objects in a multivariate Markov chain framework, and to develop a method for solving common and non-common objects' stationary probability distributions in the networks. The stationary probability distributions can be used to evaluate the importance of common and non-common objects via network-network interactions. Our experimental results based on examples of co-authorship of researchers in different conferences and paper citations in different categories have shown that the proposed model can provide useful information for researcher-researcher interactions in networks of different conferences and for paper-paper interactions in networks of different categories.

AMS subject classifications: 60J10, 91D30, 05C81

Key words: transition probability, multivariate Markov chains, stationary probability distribution, irreducible, multiple networks.

1. Introduction

Node importance evaluation is an important methodology for network analysis that can assist in the tasks of ranking query results of search engine, extracting communities of social networks and studying communities evolutions of dynamic networks. In the literature, there are many approaches to evaluating node importance or centrality

*Corresponding author. *Email addresses:* xutaolee08@gmail.com (X. Li), liwen@scnu.edu.cn (W. Li), mng@math.hkbu.edu.hk (M. Ng), yeyunming@hit.edu.cn (Y. Ye)

[3, 10, 11, 13, 15, 19]. Among them, PageRank [19] and HITS [13] are the most well-known and have been successfully applied to determine the popularity of different webpages. PageRank algorithm considers that webpages are visited randomly in a network and their limiting probabilities are used to evaluate webpages. Different from PageRank, HITS defines two evaluation scores for a node, i.e., authoritativeness score and hubness score, and computes them in a mutually reinforcing way. There are many variants of both methods for different purposes or different applications, see [6, 8, 12, 14, 18, 21].

On the other hand, there are many centrality measures that have been developed. Puzis *et al.* proposed a method for rapid computation of the group betweenness centrality [20]. Newman proposed a betweenness centrality measure based on random walks, which is computed through counting how often a node is traversed by a random walk between two nodes [17]. In [4], Brandes discussed several variants of shortest path betweenness centrality and studied their computational algorithms. Ercsey-Ravasz and Toroczka [9] studied the property of betweenness centrality in a large network.

In many scenarios, objects are involved in multiple networks rather than in a single network. For example, people are involved in multiple communication networks characterized by different communication tools, and researchers are involved in multiple collaboration networks characterized by different conferences. Common objects in multiple networks result in a network-network interactions via object-object interactions. It is more interesting to analyze object-object interactions over multiple networks and find out useful information across networks.

The main contribution of this paper is to model multiple networks where there are some common objects across them, in a multivariate Markov chain framework, and to develop a method for solving common and non-common objects' stationary probability distribution in these networks. The stationary probability distribution can be used to evaluate the importance of common and non-common objects via network-network interactions. Our model is able to handle both directed networks and undirected networks. Experimental results based on examples of co-authorship of researchers in different conferences and paper citations in different categories have shown that the proposed model can provide useful information for researcher-researcher interactions in networks of different conferences or for paper-paper interactions in networks of different categories. The proposed method is also very efficient to compute the stationary probability distribution for network analysis purpose.

In [5], Ching *et al.* have studied multivariate Markov chain models, and showed under some assumption that the existence and uniqueness of blockwise stationary probability vector of a multivariate Markov chain. The main differences between this paper and [5] are that (i) we relax the assumption to show the existence and uniqueness results, and (ii) we are interested in analysis of common and non-common objects in multiple networks via multivariate Markov chain models, where such analysis is not studied in [5].

Recently, Bini *et al.* [1, 2] and Del Corso *et al.* [7] study some integrated models for ranking scientific publications together with authors and journals. Their models rely on certain adjacency matrices obtained from the relations of citation, authorship and publication, which combine to form a suitable irreducible stochastic matrix whose Perron vector