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 provides the ranking. They design two models. One is to scale the rows to obtain a row-stochastic matrix. The other one is to partition authors, papers and journals in block, and scale each block to be row-stochastic matrix. The latter model is similar to multivariate Markov chain models. However, in their papers, they only study multivariate Markov chains of three-by-three block, while we give a general existence and uniqueness of block-wise stationary probability vector of a multivariate Markov chain. In addition, their analysis of multivariate Markov chains of three-by-three block is based on the coupling theorem involving Schur complement in [16], and these results are also different from our results in Section 2.

The rest of the paper is organized as follows. In Section 2, we present and study the model and the method. In Section 3, we present the experimental results for multiple author and paper networks arising from international conferences. Finally, some concluding remarks are given in Section 4.

2. The Proposed Model

In this paper, we are interested to model multiple networks where there are some common objects in a multivariate Markov chain framework. Assume that there are m networks and let $\mathcal{N}_k$ denote the k-th network. The number of objects in k-th network is given by n_k. For simplicity, we set $n = \sum_{k=1}^{m} n_k$ and assume there are c common objects among m networks,

Let $A^{(k,k)}$ represent the adjacency matrix of $\mathcal{N}_k$, where $A^{(k,k)}_{i,j} = 1$ if there is a directed edge from the j-th object to the i-th object in $\mathcal{N}_k$, i.e., the j-th object and the i-th object have a relationship in $\mathcal{N}_k$, otherwise $A^{(k,k)}_{i,j} = 0$. If the relationship is also reciprocal, then there is a directed edge from the i-th object to the j-th object, and therefore $A^{(k,k)}$ is symmetric. It is clear that $A^{(k,k)}$ is an n_k-by-n_k matrix.

Let $A^{(k_1,k_2)}$ represent the adjacency matrix for the objects in $\mathcal{N}_{k_2}$ interacting to the objects in $\mathcal{N}_{k_1}$. In this case, $A^{(k_1,k_2)}$ is an n_{k_1}-by-n_{k_2} matrix. The objects in $\mathcal{N}_{k_2}$ can interact with the objects in $\mathcal{N}_{k_1}$ via the common objects in the two networks. If the j-th object in $\mathcal{N}_{k_2}$ and the i-th object in $\mathcal{N}_{k_1}$ are indeed the same, i.e., the common object in the two networks, we set the j-th column of $[A^{(k_1,k_2)}]_{\cdot,j}$ equal to the i-th column of $[A^{(k_1,k_1)}]_{\cdot,i}$. In our setting, $A^{(k_1,k_2)}$ is not necessarily to be symmetric, and $A^{(k_1,k_2)}$ is not necessarily the same as $A^{(k_2,k_1)}$.

We remark that our motivation is to study the network-network interactions via common object interactions. Instead of formulating a combined single network via common objects, we keep the structure of each network and model the influence of networks with the others so that we can understand the importance and behavior of common and non-common objects in each network.

2.1. Multivariate Markov Chains

In order to evaluate the importance of common objects and non-common objects in multiple networks, we would like to compute their stationary probabilities based on a multivariate Markov chain framework. The stationary probability of an object in a network can be interpreted as the expectation of the number of random interactions required to get from the object back to itself. It is interesting to note that the random interactions include both object-object interactions within a network and across networks via common objects.

From $A^{(k_1,k_2)}$, one can compute the transition probability $[P^{(k_1,k_2)}]_{i,j}$ that the j-th object in $\mathcal{N}_{k_2}$ will interact with the i-th object in $\mathcal{N}_{k_1}$ given that currently the j-th object $\mathcal{N}_{k_2}$ is considered. Clearly, one has

$$[P^{(k_1,k_2)}]_{i,j} = \frac{[A^{(k_1,k_2)}]_{i,j}}{\sum_{l=1}^{n_{k_1}}[A^{(k_1,k_2)}]_{l,j}}, \quad i = 1,2,\cdots,n_{k_1},$$

if $[A^{(k_1,k_2)}]_{\cdot,j}$ is a non-zero vector. We note in our setting that there are some columns $[A^{(k_1,k_2)}]_{\cdot,j}$ that can be zero as the j-th object in $\mathcal{N}_{k_2}$ is not a common object with $\mathcal{N}_{k_1}$. In this case, we consider the j-th object in $\mathcal{N}_{k_2}$ interacts with objects in $\mathcal{N}_{k_1}$ with equal chance, i.e., we set

$$[P^{(k_1,k_2)}]_{i,j} = \frac{1}{n_{k_1}}, \quad i = 1,2,\cdots,n_{k_1}.$$

The whole matrix $[P^{(k_1,k_2)}]$ is called the one-step transition probability matrix from $\mathcal{N}_{k_2}$ to $\mathcal{N}_{k_1}$ because

$$[P^{(k_1,k_2)}]_{i,j} \geq 0, \ \sum_{i=1}^{n_{k_1}}[P^{(k_1,k_2)}]_{i,j} = 1, \quad j = 1,2,\cdots,n_{k_2}.$$

Let $\mathbf{x}_t^{(k)}$ and $\mathbf{x}_{t+1}^{(k)}$ $(k = 1,2,\cdots,m)$ be the probability distributions of objects being considered in the k-th network at time t and $t+1$ respectively. In the multivariate Markov chain model, we assume the following relationship:

$$\mathbf{x}_{t+1}^{(k_1)} = \sum_{k_2=1}^{m} \lambda_{k_1,k_2} P^{(k_1,k_2)} \mathbf{x}_t^{(k_2)}, \quad k_1 = 1,2,\cdots,m,$$

where the parameter λ_{k_1,k_2} satisfies the following properties:

$$\lambda_{k_1,k_2} \geq 0, \quad 1 \leq k_1,k_2 \leq m, \tag{2.1}$$

and

$$\sum_{k_2=1}^{m} \lambda_{k_1,k_2} = 1, \quad k_1 = 1,2,\cdots,m. \tag{2.2}$$

Here the parameters λ_{k_1,k_2} is the influence of the k_2-th network to the k_1-th network. In [5], it is required to set

$$\lambda_{k_1,k_2} > 0, \quad 1 \le k_1, k_2 \le m,$$

in order to obtain stationary probability distribution of a multivariate Markov chain. In other words, the proposed model is less restricted.

The object probability distribution of the k_1-th network at time $(t+1)$ depends on the weighted average of $P^{(k_1,k_2)}\mathbf{x}_t^{(k_2)}$. We note that the parameters λ_{k_1,k_2} are the weights referring the influence of the k_2-th network to the k_1-th network. In matrix form we can write

$$\mathbf{x}_{t+1} \equiv \begin{pmatrix} \mathbf{x}_{t+1}^{(1)} \\ \mathbf{x}_{t+1}^{(2)} \\ \vdots \\ \mathbf{x}_{t+1}^{(m)} \end{pmatrix} = \begin{pmatrix} \lambda_{1,1}P^{(1,1)} & \lambda_{1,2}P^{(1,2)} & \cdots & \lambda_{1,m}P^{(1,m)} \\ \lambda_{2,1}P^{(2,1)} & \lambda_{2,2}P^{(2,2)} & \cdots & \lambda_{2,m}P^{(2,m)} \\ \vdots & \vdots & \vdots & \vdots \\ \lambda_{1,1}P^{(m,1)} & \lambda_{m,2}P^{(m,2)} & \cdots & \lambda_{m,m}P^{(m,m)} \end{pmatrix} \begin{pmatrix} \mathbf{x}_t^{(1)} \\ \mathbf{x}_t^{(2)} \\ \vdots \\ \mathbf{x}_t^{(m)} \end{pmatrix} \equiv P\mathbf{x}_t.$$

The requirement in (2.2) makes sure that when $\mathbf{x}_t^{(i)}$ is a probability vector with its sum being equal to 1, then $\mathbf{x}_{t+1}^{(i)}$ has the same property. Although the column sum of P is not equal to one (the column sum of $P^{(k_1,k_2)}$ is equal to one), we still have the following results.

Lemma 2.1. *Let $\lambda_{k,k}P^{(k,k)}$ be irreducible, $k = 1, 2, \cdots, m$. Then P is irreducible if and only if the matrix $\Lambda = [\lambda_{k_1,k_2}]_{k_1,k_2=1}^{k_1,k_2=m}$ is irreducible.*

Proof. Let us define $c_1 = 0$ and $c_k = \sum_{u=1}^{k-1} n_u$ for $k = 2, 3, \cdots, m$, and $\bar{k} = \{c_k + 1, c_k + 2, \cdots, c_k + n_k\}$ for $k = 1, 2, \cdots, m$. For i and j in between 1 and n, if i and j are in $\bar{k}$ for some k, there is a path connecting i and j as $\lambda_{k,k}P^{(k,k)}$ is irreducible.

Now, let us consider $i \in \bar{k}$ and $j \in \bar{r}$ where $k \neq r$. Since Λ is irreducible, there is a path connecting k and r, i.e., λ_{k,v_1}, λ_{v_1,v_2}, $\cdots$, $\lambda_{v_t,r}$ are not equal to zero. It is noted that since $P^{(\cdot,\cdot)}$ is a transition probability matrix, i.e., $P^{(k,v_1)} \neq 0, \cdots, P^{(v_t,r)} \neq 0$, and $\lambda_{j,j}P^{(j,j)}$ for $j = 1, \cdots, m$ are irreducible, we can deduce that there is a path connecting i and j.

Conversely, if Λ is reducible, then there exist two subsets S_1 and S_2 such that $S_1 \cup S_2 = \{1, 2, \cdots, m\}$, $S_1 \cap S_2 = \phi$, and $\lambda_{i,j} = 0$ for any $i \in S_1$ and $j \in S_2$. We set $\bar{k_1} = \cup_{i \in S_1} \bar{i}$ and $\bar{k_2} = \cup_{j \in S_2} \bar{j}$. Then there is no path connecting from $i \in \bar{k_1}$ to $j \in \bar{k_2}$ which contradicts the assumption. The result follows. ☐

Remark: The result of the above lemma is less restricted than that given in [5] where all λ_{k_1,k_2} must be positive. The above lemma does not hold without the assumption that the irreducibility of $\lambda_{j,j}P^{(j,j)}$ for $j = 1, \cdots, m$. For example, we consider $\lambda_{1,1}$ and $\lambda_{2,2}$ are equal to 0, $\lambda_{2,1}P^{(2,1)}$ and $\lambda_{1,2}P^{(1,2)}$ are the 2-by-2 identity matrix, and $P^{(11)}$ and $P^{(22)}$ are irreducible matrices. Then we have

$$P = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

Clearly, P is reducible.

Lemma 2.2. *If Λ is irreducible, then the matrix P has an eigenvalue equal to one.*

Proof. It is easy to see that 1 is a maximal eigenvalue in modulus of Λ. Then by Perron-Frobenius Theorem, there exists a vector $\mathbf{y} = (y_1, y_2, \cdots, y_m)^T$ such that $\mathbf{y}^T\Lambda = \mathbf{y}^T$ where $\cdot^T$ denotes the transpose operation. We note that $\mathbf{e}^T P^{(i,j)} = \mathbf{e}^T$ for $1 \le i, j \le m$ where $\mathbf{e}$ is the vector of all ones. Then it is easy to show that

$$(y_1\mathbf{e}^T, y_2\mathbf{e}^T, \cdots, y_m\mathbf{e}^T)P = (y_1\mathbf{e}^T, y_2\mathbf{e}^T, \cdots, y_m\mathbf{e}^T),$$

and hence 1 is an eigenvalue of P. □

Lemma 2.3. *Let $\lambda_{k,k}P^{(k,k)}$ be irreducible, $k = 1, \cdots, m$, and $\Lambda = [\lambda_{k_1,k_2}]_{k_1,k_2=1}^{k_1,k_2=m}$ be irreducible. Then 1 is the maximal eigenvalue of P in magnitude, and P is convergent if and only if P is primitive*†*, in this case,*

$$\lim_{t\to\infty}(P)^t = \mathbf{u}\mathbf{v}^T,$$

where $\mathbf{u}$ *and* $\mathbf{v}$ *are positive n-by-1 vectors.*

Proof. By the proof of Lemma 2, P has a positive eigenvector corresponding to the eigenvalue 1, i.e., $\mathbf{y}^T P = \mathbf{y}^T$. Now we show that 1 is the spectral radius of P. Let μ be the maximal eigenvalue of P in magnitude. By Lemma 1, P is irreducible. Hence by Perron-Frobenius theorem μ is positive and there is a positive eigenvector $\mathbf{x}$ corresponding to μ, i.e., $P\mathbf{x} = \mu\mathbf{x}$. This implies that

$$1 \cdot \mathbf{y}^T\mathbf{x} = \mathbf{y}^T P\mathbf{x} = \mathbf{y}^T \mu\mathbf{x}.$$

Because $\mathbf{y}^T\mathbf{x}$ is a scalar and it must be positive, we obtain $\mu = 1$.

Now we know that P is primitive if and only if there is nonsingular matrix S such that

$$P = S\begin{pmatrix} 1 & 0 \\ 0 & P' \end{pmatrix}S^{-1},$$

where the spectral radius of P' is less than 1. It follows that the first column $\mathbf{u}$ of S and the first row $\mathbf{v}^T$ of S^{-1} are the eigenvectors corresponding to the eigenvalue 1. They can be chosen to be positive vectors. Hence we have

$$\lim_{t\to\infty} P^t = S\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}S^{-1} = \mathbf{u}\mathbf{v}^T.$$

The result follows. □

†Let A be a nonnegative and irreducible matrix with maximal eigenvalue r. If there is only one eigenvalue of modulus r, then A is said to be primitive.

Remark: In the above lemma, we cannot omit the primitivity of P. For example, we consider

$$P = \frac{1}{2}\begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix},$$

where $\lambda_{1,1} = \lambda_{2,2} = \lambda_{1,2} = \lambda_{2,1} = 1/2$, and

$$P^{(2,1)} = P^{(1,2)} = P^{(1,1)} = P^{(2,2)} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

are irreducible. However, P is not convergent since 1 and -1 are eigenvalues of P. We should emphasize that this requirement is not discussed in [5].

Lemma 2.4. *Let $\lambda_{k,k}P^{(k,k)}$ be irreducible, $k = 1, \cdots, m$, and $\Lambda = [\lambda_{k_1,k_2}]_{k_1,k_2=1}^{k_1,k_2=m}$ be irreducible. If there is an index j such that $P^{(j,j)}$ is primitive, then P is convergent.*

Proof. By the assumption that $P^{(j,j)}$ is primitive, it is equivalent to the fact that the greatest common factor of the length of any closed path in the directed graph associated with $P^{(j,j)}$ is equal to 1. Clearly, the directed graph associated with $P^{(j,j)}$ is a subgraph of the directed graph associated with P, and therefore the greatest common factor of the length of any closed path in the directed graph associated with P must also be equal to 1. Hence P is also primitive. According to the previous lemma, we know that P is convergent.

□

Theorem 2.1. *Let $\lambda_{k,k}P^{(k,k)}$ be irreducible, $k = 1, 2, \cdots, m$, and $\Lambda = [\lambda_{k_1,k_2}]_{k_1,k_2=1}^{k_1,k_2=m}$ be irreducible. If there is an index j such that $P^{(j,j)}$ is primitive, then*

$$\lim_{t\to\infty} (P)^t = \mathbf{u}\mathbf{v}^T,$$

where $\mathbf{u}$ and $\mathbf{v}$ are positive n-by-1 vectors where $n = \sum_{i=1}^m n_i$.

For any given initial vector $\mathbf{x}_0$ with nonnegative entries, we have

$$\lim_{t\to\infty} \mathbf{x}_t = \lim_{t\to\infty} P^t \mathbf{x}_0 = \mathbf{u}\mathbf{v}^T \mathbf{x}_0.$$

This implies that $\mathbf{x}_t$ tends a stationary nonnegative vector $\mathbf{x}$ as t goes to infinity. We also note that if $\mathbf{x}_0$ has the following property

$$\sum_{i=1}^{n_k} [\mathbf{x}_0^{(k)}]_i = 1, \quad 1 \le k \le m,$$

then $\mathbf{x}_t$ and $\mathbf{x}$ have the same property.

Corollary 2.1. *Under the assumptions of Theorem 1, there is a unique vector* $\mathbf{x} = [\mathbf{x}^{(1)}, \cdots, \mathbf{x}^{(m)}]^T$ *such that* $\mathbf{x} = P\mathbf{x}$ *and*

$$\sum_{i=1}^{n_k} [\mathbf{x}^{(k)}]_i = 1, \quad 1 \le k \le m.$$

Algorithm 2.1 Power Method
Require: an $n \times n$ matrix P
Ensure: the principal eigenvector x of P

1. $\mathbf{x}_0 = [\frac{1}{n_1}\mathbf{e}, \frac{1}{n_2}\mathbf{e}, \cdots, \frac{1}{n_m}\mathbf{e}]^T$;
2. $t = 1$;
3. **while** until convergence **do**
4. $\mathbf{x}_t = P\mathbf{x}_{t-1}$;
5. $t = t + 1$;
6. **end while**

Remark: In the model, we need not to assume that all $P^{(i,j)}$ are irreducible and all $\lambda_{i,j}$ are positive as required in [5].

2.2. The Algorithm

According to the results in the previous subsection, we can employ the power method algorithm for computing principal eigenvector of P. Under the assumptions, we can make sure the algorithm is convergent for any given initial nonnegative vector.

As an example, we construct a synthetic two-network to demonstrate the calculation. We would like to show that the network-network interaction via common objects can bring interesting results for both common objects and non-common objects interaction. Two networks are generated and there are 13 objects all together, and they are shown in Fig. 1. We see that objects 1, 2, 3, 4, 5, 6, 12 and 13 appear in network 1 and objects 1, 2, 3, 6, 7, 8, 9, 10 and 11 appear in network 2. The common objects in both networks are 1, 2, 3 and 6. The others are non-common objects.

In this example, the adjacency matrices are given as follows:

$$A^{(1,1)} = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad A^{(2,2)} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$
A^{(1,2)} = \begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}, \quad
A^{(2,1)} = \begin{pmatrix}
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}.
$$

Their transition probability matrices are given as follows:

$$
P^{(1,1)} = \begin{pmatrix}
0 & 0 & 0 & 1/2 & 1/4 & 0 & 1 & 0 \\
1/4 & 0 & 1/4 & 0 & 1/4 & 1 & 0 & 0 \\
0 & 1/3 & 0 & 0 & 1/4 & 0 & 0 & 1 \\
1/4 & 0 & 1/4 & 0 & 1/4 & 0 & 0 & 0 \\
1/4 & 1/3 & 1/4 & 1/2 & 0 & 0 & 0 & 0 \\
0 & 1/3 & 0 & 0 & 0 & 0 & 0 & 0 \\
1/4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1/4 & 0 & 0 & 0 & 0 & 0
\end{pmatrix},
$$

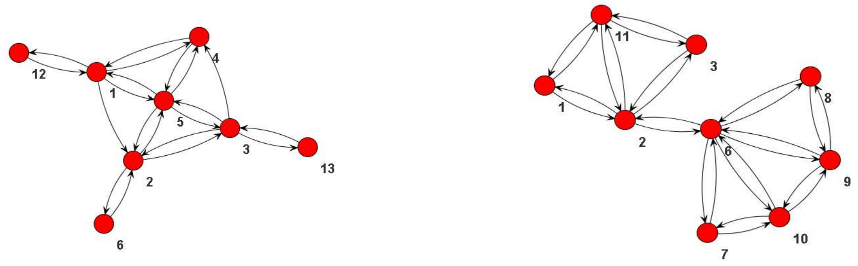
$$
P^{(2,2)} = \begin{pmatrix}
0 & 1/4 & 0 & 0 & 0 & 0 & 0 & 0 & 1/3 \\
1/2 & 0 & 1/2 & 1/5 & 0 & 0 & 0 & 0 & 1/3 \\
0 & 1/4 & 0 & 0 & 0 & 0 & 0 & 0 & 1/3 \\
0 & 1/4 & 0 & 0 & 1/2 & 1/2 & 1/3 & 1/3 & 0 \\
0 & 0 & 0 & 1/5 & 0 & 0 & 0 & 1/3 & 0 \\
0 & 0 & 0 & 1/5 & 0 & 0 & 1/3 & 0 & 0 \\
0 & 0 & 0 & 1/5 & 0 & 1/2 & 0 & 1/3 & 0 \\
0 & 0 & 0 & 1/5 & 1/2 & 0 & 1/3 & 0 & 0 \\
1/2 & 1/4 & 1/2 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix},
$$

$$
P^{(1,2)} = \begin{pmatrix}
0 & 0 & 0 & 0 & 1/8 & 1/8 & 1/8 & 1/8 & 1/8 \\
1/4 & 0 & 1/4 & 1 & 1/8 & 1/8 & 1/8 & 1/8 & 1/8 \\
0 & 1/3 & 0 & 0 & 1/8 & 1/8 & 1/8 & 1/8 & 1/8 \\
1/4 & 0 & 1/4 & 0 & 1/8 & 1/8 & 1/8 & 1/8 & 1/8 \\
1/4 & 1/3 & 1/4 & 0 & 1/8 & 1/8 & 1/8 & 1/8 & 1/8 \\
0 & 1/3 & 0 & 0 & 1/8 & 1/8 & 1/8 & 1/8 & 1/8 \\
1/4 & 0 & 0 & 0 & 1/8 & 1/8 & 1/8 & 1/8 & 1/8 \\
0 & 0 & 1/4 & 0 & 1/8 & 1/8 & 1/8 & 1/8 & 1/8
\end{pmatrix},
$$

$$
P^{(2,1)} = \begin{pmatrix}
0 & 1/4 & 0 & 1/9 & 1/9 & 0 & 1/9 & 1/9 \\
1/2 & 0 & 1/2 & 1/9 & 1/9 & 1/5 & 1/9 & 1/9 \\
0 & 1/4 & 0 & 1/9 & 1/9 & 0 & 1/9 & 1/9 \\
0 & 1/4 & 0 & 1/9 & 1/9 & 0 & 1/9 & 1/9 \\
0 & 0 & 0 & 1/9 & 1/9 & 1/5 & 1/9 & 1/9 \\
0 & 0 & 0 & 1/9 & 1/9 & 1/5 & 1/9 & 1/9 \\
0 & 0 & 0 & 1/9 & 1/9 & 1/5 & 1/9 & 1/9 \\
0 & 0 & 0 & 1/9 & 1/9 & 1/5 & 1/9 & 1/9 \\
1/2 & 1/4 & 1/2 & 1/9 & 1/9 & 0 & 1/9 & 1/9
\end{pmatrix}.
$$

We note that both $P^{(1,1)}$ and $P^{(2,2)}$ are irreducible and primitive. In this example, we set $\lambda_{1,1} = \lambda_{1,2} = \lambda_{2,1} = \lambda_{2,2} = 1/2$. According to our results, the power method converges and gives stationary probability distributions of two networks.

Table 1 shows the stationary probabilities obtained by the proposed algorithm and the PageRank algorithm. When we use the PageRank algorithm for individual network, the objects 5 and 6 have the highest stationary probabilities in the first and second networks respectively. We note that the object 5 is not a common object, and the object 6 is a common object which is not important in the first network. According to the PageRank



(a) (b)

Figure 1: An example of a synthetic two-network. (a) network 1; (b) network 2.

Table 1: The stationary probabilities calculated by multivariate Markov model and calculated by PageRank for individual networks.

Method	PageRank		Multivariate Markov model	
	network 1	network 2	network 1	network 2
Object 1	0.1579	0.0769	0.1075	0.1085
Object 2	0.1974	0.1539	**0.2256**	**0.1978**
Object 3	0.1579	0.0769	0.1530	0.1085
Object 4	0.1319	-	0.1115	-
Object 5	**0.2105**	-	0.1866	-
Object 6	0.0658	**0.1923**	0.0991	0.1291
Object 7	-	0.0769	-	0.0579
Object 8	-	0.0769	-	0.0579
Object 9	-	0.1154	-	0.0724
Object 10	-	0.1154	-	0.0724
Object 11	-	0.1154	-	0.1954
Object 12	0.0395	-	0.0555	-
Object 13	0.0395	-	0.0612	-

results, we cannot determine important common objects in both networks. However, when the multivariate Markov model is used, the common object 2 has the highest stationary probabilities in both networks. This result shows that the object 2 plays a very important role in the connection between two networks.

On the other hand, the non-common object 11 is the second important (the probability is 0.1954) in the network 2, while it is the third important object (the probability is 0.1154) when individual network is considered. This observation may be explained by the fact that the non-common object 11 is interacted with important common objects 1, 2 and 3 in network 2.

In the next section, we present two real data examples to demonstrate the usefulness of the proposed model in identification of important common objects.

3. Experimental Results

The first experiment is based on multiple collaboration networks in conferences, and the second experiment is based on multiple paper citation networks. We collected the data for several years of different data mining conferences from DBLP[‡]. Given the data, we construct multiple collaboration networks or multiple paper citation networks as follows:

- For collaboration networks, choose m conferences and construct a network for each conference by treating researchers who have publications in the conference as nodes and adding a bi-directed edge between two nodes if the corresponding researchers collaborate in the conference. In this case, we construct m networks in terms of the co-authorship of researchers in m conferences for experiment 1. The researchers appearing commonly in the m conferences are common objects.
- For paper citation networks, choose m categories and construct a network for each category by treating papers that belong to the category as nodes and adding a directed edge between two nodes if the corresponding paper cites another. In this case, we construct m networks in terms of the citation of papers belonging to m categories for experiment 2. The papers appearing commonly in the m categories are common objects.

According to lemmas in Section 2.1, we need to make sure $\lambda_{k,k}P^{(k,k)}$ $(k=1,2,\cdots,m)$ and $\Lambda = \left[\lambda_{k_1,k_2}\right]_{k_1,k_2=1}^{k_1,k_2=m}$ be irreducible. Therefore we preprocess $P^{(k,k)}$ $(k=1,2,\cdots,m)$ by applying the method used in PageRank:

$$P^{(k,k)} = (1-d)P^{(k,k)} + \frac{d}{n_k}S_k,$$

where S_k represents the n_k-by-n_k matrix of all ones, and we set d to be 0.1 in both experiments. For the parameters $\Lambda = \left[\lambda_{k_1,k_2}\right]_{k_1,k_2=1}^{k_1,k_2=m}$ in our model, we address how to estimate them in the next subsection.

3.1. Estimations of Parameters

In our model, we assume that multiple networks interact with each other via the common objects, therefore we can estimate parameters $\Lambda = \left[\lambda_{k_1,k_2}\right]_{k_1,k_2=1}^{k_1,k_2=m}$ by considering the influence from common objects among the networks. For collaboration networks, let $C=\left[c_{k_1,k_2}\right]_{k_1,k_2=1}^{k_1,k_2=m}$ be a matrix that represents the number of common objects between the two networks, i.e., c_{k_1,k_2} denotes the number of common objects between the k_1-th network and the k_2-th network. For paper citation networks, let $C=\left[c_{k_1,k_2}\right]_{k_1,k_2=1}^{k_1,k_2=m}$ be a matrix that represents the number of citations from one network to anther via common

[‡]http://www.informatik.uni-trier.de/~ley/db/

objects, i.e., c_{k_1,k_2} denotes the number of citations from the k_2-th network to the k_1-th network via common objects between them. Given the matrix C, we compute the matrix Λ as the following formula:

$$\lambda_{k_1,k_2} = \frac{c_{k_1,k_2}}{\sum_{l=1}^{m} c_{k_1,l}}, \quad 0 \le k_1, k_2 \le m.$$

Since we have the data for several years collected from DBLP, we can estimate how frequently an object is associated to a network. Based on this estimation, we can check whether the estimated parameters Λ is acceptable and reasonable. The idea is given as follows. For each network k, we estimate the stationary vector probability $\mathbf{q}_k$:

$$[\mathbf{q}_k]_j = \frac{f_{k,j}}{\sum_{l=1}^{n_k} f_{k,l}},$$

where $f_{k,j}$ represents the number of papers that j-th researcher has published in k-th conference for collaboration networks. Similarly, $f_{k,j}$ represents the times that j-th paper has been cited by other papers belonging to k-th category for paper citation networks. Given these m stationary probability vectors, we can check the following m approximation errors

$$\left\| \sum_{k_2=1}^{m} \lambda_{k_1,k_2} P^{(k_1,k_2)} \mathbf{q}_{k_2} - \mathbf{q}_{k_1} \right\|. \tag{3.1}$$

where $k_1 = 1, 2, \cdots, m$, to be small enough. In our experiments, we consider both l_∞-norm and l_2-norm in the evaluation because the former one shows the worst case of the approximation error and the latter one shows the averaged case of the approximation error.

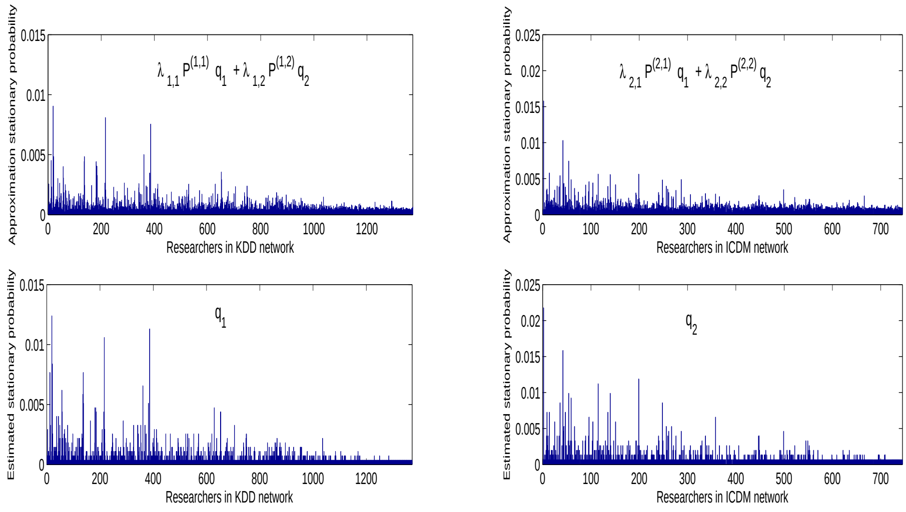
3.2. Experiment 1

In this experiment, we construct a collaboration network by considering the KDD conference and the ICDM conference from 1994 to 2009 and from 2001 to 2009 respectively. There are 1372 researchers in the KDD network and 745 researchers in the ICDM network. There are 240 common researchers that appear in both networks. The matrix C and the estimated parameters Λ are as follows:

$$C = \begin{bmatrix} 1372 & 240 \\ 240 & 745 \end{bmatrix}, \quad \Lambda_{est} = \begin{bmatrix} 0.8511 & 0.1489 \\ 0.2437 & 0.7563 \end{bmatrix}.$$

By using Λ_{est}, we compute the approximation errors as in (3.1). For the KDD network, we have

$$\left\| \sum_{k_2=1}^{2} \lambda_{1,k_2} P^{(1,k_2)} \mathbf{q}_{k_2} - \mathbf{q}_1 \right\|_\infty = 0.0038, \quad \left\| \sum_{k_2=1}^{2} \lambda_{1,k_2} P^{(1,k_2)} \mathbf{q}_{k_2} - \mathbf{q}_1 \right\|_2 = 0.0139.$$



(a) (b)

Figure 2: The approximation stationary probability vector and the estimated stationary probability vector for each network. (a) KDD network in experiment 1; (b) ICDM network in experiment 1.

Fig. 2(a) shows the approximation stationary probability vector and the estimated stationary probability vector for this network. For the ICDM network, we have

$$\left\| \sum_{k_2=1}^{2} \lambda_{2,k_2} P^{(2,k_2)} \mathbf{q}_{k_2} - \mathbf{q}_2 \right\|_\infty = 0.0063, \quad \left\| \sum_{k_2=1}^{2} \lambda_{2,k_2} P^{(2,k_2)} \mathbf{q}_{k_2} - \mathbf{q}_2 \right\|_2 = 0.0207.$$

Fig. 2(b) shows the approximation stationary probability vector and the estimated stationary probability vector for this network. Based on the figures and the computing results, we find that the approximation errors are small for both networks, which indicates that the estimated Λ is acceptable and reasonable.

Tables 2 and 3 show the top common and non-common researchers in KDD and ICDM networks when PageRank and the proposed Multivariate Markov chain model are used. According to Tables 2 and 3, we find the following interesting results.

(i) The top ten lists of common researchers by the two methods are about the same in the two networks. Christos Faloutsos cannot be identified in the ICDM network by the PageRank, but can be identified (rank number one) in the KDD network. It is clear that he is an important common researcher in both the KDD and ICDM networks. Indeed, the number of his collaborations with all the researchers and the top ten common researchers are 17 and 13 respectively in the ICDM network which are better than the number of collaborations (15 and 8) of Ruoming Jin ranked the number eight by the PageRank. These results show that the use of PageRank may not be effective, but the new model can consider the importance of a researcher in both collaborations with all the researchers and the common researchers together to

Table 2: The top ten lists of common and non-common researchers **in KDD network** by PageRank and multivariate Markov chain.

	Common Researchers		Non-Common Researchers	
Ranking	PageRank	Multivariate Markov	PageRank	Multivariate Markov
1	Christos Faloutsos	Jiawei Han	Padhraic Smyth	Martin Ester
2	Jiawei Han	Philip S. Yu	Martin Ester	Padhraic Smyth
3	Philip S. Yu	Christos Faloutsos	Rakesh Agrawal	Rakesh Agrawal
4	Heikki Mannila	Jian Pei	Ravi Kumar	Ravi Kumar
5	Jian Pei	Eamonn J. Keogh	Johannes Gehrke	Rong Jin
6	Ke Wang	Ke Wang	David Jensen	Foster J. Provost
7	Bing Liu	Heikki Mannila	Usama M. Fayyad	Wynne Hsu
8	Eamonn J. Keogh	Bing Liu	Ron Kohavi	David Jensen
9	Vipin Kumar	Srinivasan Parthasarathy	Foster J. Provost	Deepak Agarwal
10	Mohammed Javeed Zaki	Vipin Kumar	Gregory Piatetsky-Shapiro	Bishan Yang

Table 3: The top ten lists of common and non-common researchers **in ICDM network** by PageRank and multivariate Markov chain.

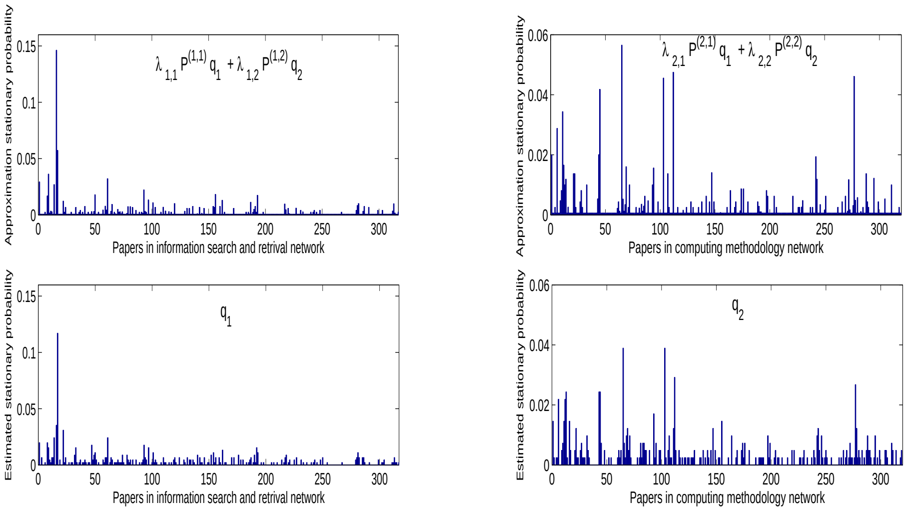
	Common Researchers		Non-Common Researchers	
Ranking	PageRank	Multivariate Markov	PageRank	Multivariate Markov
1	Philip S. Yu	Philip S. Yu	Xinsong Wu	Xindong Wu
2	Jiawei Han	Jiawei Han	Peng Zhang	Fei Wang
3	Haixun Wang	Haixun Wang	Fei Wang	Ben Kao
4	Eamonn J. Keogh	Eamonn J. Keogh	Jun Yan	Peng Zhang
5	Zheng Chen	Christos Faloutsos	Benyu Zhang	Sau Dan Lee
6	Wei Fan	Ruoming Jin	Yiyu Yao	Jing Peng
7	Hans-Peter Kriegel	Hans-Peter Kriegel	Ben Kao	Benyu Zhang
8	Ruoming Jin	Wei Fan	Ning Zhong	Jun Yan
9	Qiang Yang	Wei Wang	Jing Peng	Dongqing Yang
10	Wei Wang	Srinivasan Parthasarathy	Sau Dan Lee	Yiyu Yao

generate ranking results. Christos Faloutsos is ranked number three and five in the KDD and ICDM networks.

(ii) The lists of top non-common researchers by the two methods are quite different. In the KDD networks, there are four different researchers. In the ICDM networks, there are also one different researcher. As we study two collaboration networks, we are interested in the researchers collaborating with each other via common objects. The multivariate Markov chain model provides most frequent non-common researchers collaborating with common researchers in the two networks. The number of collaborations are 32 and 23 in the KDD and ICDM networks respectively. Their relations like supervisor-student and colleagues can be further identified via these results by the proposed model. However, the collaborations between common researchers and non-common researchers are weak in the results by the PageRank. The number of collaborations are only 17 and 19 in the KDD and ICDM networks respectively.

3.3. Experiment 2

In this experiment, we construct two citation networks by considering papers that belong to the category of “Information Search and Retrieval" and papers that belong to the category of “Computing Methodologies" respectively. The papers and their categories infor-



(a) (b)

Figure 3: The approximation stationary probability vector and the estimated stationary probability vector for each network. (a) information search and retrieval network; (b) computing methodology network.

mation are all obtained from the KDD and CIKM in DBLP[§]. There are 317 papers belonging to the category of "Information Search and Retrieval" and 320 papers belonging to the category of "Computing Methodologies" for our consideration. There are 56 common papers that appear in both networks. The matrix C and the estimated parameters Λ are given as follows:

$$C = \begin{bmatrix} 453 & 69 \\ 69 & 411 \end{bmatrix}, \quad \Lambda_{est} = \begin{bmatrix} 0.8678 & 0.1322 \\ 0.1437 & 0.8562 \end{bmatrix}.$$

By using Λ_{est}, we compute the approximation errors as in Equation 3.1. For information search and retrieval network, we have

$$\left\| \sum_{k_2=1}^{2} \lambda_{1,k_2} P^{(1,k_2)} \mathbf{q}_{k_2} - \mathbf{q}_1 \right\|_\infty = 0.1109, \quad \left\| \sum_{k_2=1}^{2} \lambda_{1,k_2} P^{(1,k_2)} \mathbf{q}_{k_2} - \mathbf{q}_1 \right\|_2 = 0.1376.$$

Fig. 3(a) shows the approximation stationary probability vector and the estimated stationary probability vector for this network. For computing methodology network, we have

$$\left\| \sum_{k_2=1}^{2} \lambda_{2,k_2} P^{(2,k_2)} \mathbf{q}_{k_2} - \mathbf{q}_2 \right\|_\infty = 0.0345, \quad \left\| \sum_{k_2=1}^{2} \lambda_{2,k_2} P^{(2,k_2)} \mathbf{q}_{k_2} - \mathbf{q}_2 \right\|_2 = 0.0748.$$

[§]We collect the papers from both conferences in which reference lists are provided in DBLP. More precisely, we collect papers from 1999 to 2010 for KDD and we collect papers in 2000 and also from 2002 to 2009 for CIKM.

Table 4: The top ten lists of common and non-common papers **in information search and retrieval network** by PageRank and Multivariate Markov chain.

Ranking	PageRank (Common Papers)
1	Agglomerative clustering of a search engine query log
2	Fast and effective text mining using linear-time document clustering
3	On the merits of building categorization systems by supervised clustering
4	Optimizing search engines using clickthrough data
5	Combining link-based and content-based methods for web document classification
6	Taxonomy-driven computation of product recommendations
7	Kernel k-means: spectral clustering and normalized cuts
8	Using appraisal groups for sentiment analysis
9	Determining the semantic orientation of terms through gloss classification
10	Enhanced word clustering for hierarchical text classification
Ranking	Multivariate Markov (Common Papers)
1	Agglomerative clustering of a search engine query log
2	Fast and effective text mining using linear-time document clustering
3	Optimizing search engines using clickthrough data
4	On the merits of building categorization systems by supervised clustering
5	Combining link-based and content-based methods for web document classification
6	Taxonomy-driven computation of product recommendations
7	Kernel k-means: spectral clustering and normalized cuts
8	Information-theoretic co-clustering
9	Determining the semantic orientation of terms through gloss classification
10	Using appraisal groups for sentiment analysis
Ranking	PageRank (Non-Common Papers)
1	Co-clustering documents and words using bipartite spectral graph partitioning
2	Simple BM25 extension to multiple weighted fields
3	A cross-collection mixture model for comparative text mining
4	Query association for effective retrieval
5	SimRank: a measure of structural-context similarity
6	Query expansion using random walk models
7	First story detection in TDT is hard
8	Efficient identification of Web communities
9	A local search mechanism for peer-to-peer networks
10	Document quality models for web ad hoc retrieval
Ranking	Multivariate Markov (Non-Common Papers)
1	Co-clustering documents and words using bipartite spectral graph partitioning
2	Simple BM25 extension to multiple weighted fields
3	A cross-collection mixture model for comparative text mining
4	SimRank: a measure of structural-context similarity
5	Query expansion using random walk models
6	Query association for effective retrieval
7	First story detection in TDT is hard
8	Efficient identification of Web communities
9	A local search mechanism for peer-to-peer networks
10	Document quality models for web ad hoc retrieval

Fig. 3(b) shows the approximation stationary probability vector and the estimated stationary probability vector for this network. Even though the approximation error in information search and retrieval network is large (compared with the errors in the last two experiments), we suggest the estimated Λ can be acceptable because the approximation stationary probability vector and the estimated stationary probability vector have quite similar patterns, see Fig. 3(a).

Table 5: The top ten lists of common and non-common papers **in computing methodology network** by PageRank and multivariate Markov chain.

Ranking	PageRank (Common Papers)
1	Information-theoretic co-clustering
2	Agglomerative clustering of a search engine query log
3	Fast and effective text mining using linear-time document clustering
4	Generative model-based clustering of directional data
5	Visualization of navigation patterns on a web site using model-based clustering
6	On the merits of building categorization systems by supervised clustering
7	Combining link-based and content-based methods for web document classification
8	Enhanced word clustering for hierarchical text classification
9	Using appraisal groups for sentiment analysis
10	Kernel k-means: spectral clustering and normalized cuts
Ranking	Multivariate Markov (Common Papers)
1	Agglomerative clustering of a search engine query log
2	Information-theoretic co-clustering
3	Fast and effective text mining using linear-time document clustering
4	Visualization of navigation patterns on a web site using model-based clustering
5	Generative model-based clustering of directional data
6	Combining link-based and content-based methods for web document classification
7	On the merits of building categorization systems by supervised clustering
8	Optimizing search engines using clickthrough data
9	Using appraisal groups for sentiment analysis
10	Enhanced word clustering for hierarchical text classification
Ranking	PageRank (Non-Common Papers)
1	Efficient progressive sampling
2	MetaCost: A general method for making classifiers cost-sensitive
3	Mining the most interesting rules
4	Efficient clustering of high-dimensional data sets with application to reference matching
5	Mining and summarizing customer reviews
6	Statistics and data mining techniques for lifetime value modeling
7	Data selection for support vector machine classifiers
8	Modeling and predicting personal information dissemination behavior
9	Using association rules for product assortment decisions: a case study
10	A classifier for semi-structured documents
Ranking	Multivariate Markov (Non-Common Papers)
1	MetaCost: a general method for making classifiers cost-sensitive
2	Mining the most interesting rules
3	Efficient progressive sampling
4	Mining and summarizing customer reviews
5	Efficient clustering of high-dimensional data sets with application to reference matching
6	Statistics and data mining techniques for lifetime value modeling
7	Data selection for support vector machine classifiers
8	Modeling and predicting personal information dissemination behavior
9	Using association rules for product assortment decisions: a case study
10	Incorporating prior knowledge with weighted margin support vector machines

Tables 4 and 5 show the top common and non-common papers in information search and retrieval network and computing methodology network when PageRank and the proposed Multivariate Markov chain model are used. According to Table 4, we find the following interesting results.

(i) In both networks, there are more citations by the common papers to the top ten common papers identified by the multivariate Markov chain than those by the PageRank, i.e., 33 against 27 in information search and retrieval network and 30 against 24 in computing methodology network. In the computing methodology network, there are also more citations by all the papers to the top ten common papers identified by the multivariate Markov chain than those by the PageRank (56 against 48).

(ii) One important common paper "Information-theoretic co-clustering" cannot be identified by the PageRank in information search and retrieval network, but can be identified by the PageRank in computing methodology network (ranked number one). These results again show that the PageRank cannot capture the connection of common objects between two networks. As this common paper has high citations, the multivariate Markov chain can rank it number eight in the top ten list of information search and retrieval network and number two in the top ten list of computing methodology network.

4. Concluding Remarks

In this paper, we have proposed to study common and non-common objects appearing in multiple networks. Usually, common objects are important in network-network interactions via object-object interactions, and they play a key connection point for other objects in multiple networks. We have developed a multivariate Markov chain model for analyzing multiple networks. New theoretical results of the multivariate Markov chain model are derived. Different from [5], our model is to analyze the importance of common objects and non-common objects in multiple networks. The other contribution of this paper is to give a less restricted model than that in [5] where each block in the whole transition matrix is required to be irreducible.

We have performed some experiments based on examples of co-authorship of researchers in different conferences and paper citations in different categories. The experimental results 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.

Acknowledgments W. Li's research is supported in part by National Natural Science Foundations of China (Nos. 10671077, 10971075), Research Fund for the Doctoral Program of Higher Education of China (No. 20104407110002) and Guangdong Provincial Natural Science Foundations, P. R. China (No. 9151063101000021). Y. Ye's research was supported in part by NSFC under Grant no.61073195, Shenzhen Science and Technology Program under Grant no. CXB201005250024A and ZD201006100018A, and Natural Scientific Research Innovation Foundation in HIT under Grant no. HIT.NSFIR.2010128.

References

[1] D. Bini, G. Del Corso, AND F. Romani, *Evaluating scientific products by means of citation-based models: a first analysis and validation*. Electronic Transactions on Numerical Analysis, 33:1–16, 2008.

[2] D. Bini, G. Del Corso, AND F. Romani, *A combined approach for evaluating papers, authors and scientific journals*. Journal of computational and applied mathematics, 234(11):3104–3121, 2010.

[3] S. Borgatti, *Centrality AND network flow*. Social Networks, 27(1):55–71, 2005.

[4] U. Brandes, *On variants of shortest-path betweenness centrality and their generic computation*. Social Networks, 30(2):136–145, 2008.

[5] W. Ching, E. Fung, AND M. Ng, *A multivariate Markov chain model for categorical data sequences and its applications in demand predictions*. IMA Journal of Management Mathematics, 13(3):187, 2002.

[6] D. Cohn AND H. Chang, *Learning to probabilistically identify authoritative documents*. In Proc. of International Conference on Machine Learning, pp. 167–174, 2000.

[7] G. Del Corso AND F. Romani, *Versatile weighting strategies for a citation-based research evaluation model*. Bulletin of the Belgian Mathematical Society-Simon Stevin, 16(4):723–743, 2009.

[8] H. Deng, M. Lyu, AND I. King, *A generalized Co-HITS algorithm and its application to bipartite graphs*. In Proc. of ACM SIGKDD, pp. 239–248, 2009.

[9] M. Ercsey-Ravasz AND Z. Toroczkai, *Centrality scaling in large networks*. Phys. Rev. Lett., 105(3):038701, Jul 2010.

[10] L. Freeman, *A set of measures of centrality based on betweenness*. Sociometry, pp. 35–41, 1977.

[11] L. Freeman, *Centrality in social networks conceptual clarification*. Social networks, 1(3):215–239, 1979.

[12] T. H. Haveliwala, *Topic-sensitive PageRank*. In Proc. of ACM World Wide Web, 2002.

[13] J. Kleinberg, *Authoritative sources in a hyperlinked environment*. Journal of the ACM (JACM), 46(5):604–632, 1999.

[14] T. Kolda, B. Bader, AND J. Kenny, *Higher-order web link analysis using multilinear algebra*. In Proc. of IEEE International Conference on Data Mining, pp. 242–249, 2005.

[15] V. Latora AND M. Marchiori, *A measure of centrality based on network efficiency*. New Journal of Physics, 9:188, 2007.

[16] C. Meyer, *Stochastic complementation, uncoupling Markov chains and the theory of nearly reducible systems*. SIAM Rev., 31(2):240–272, 1989.

[17] M. Newman, *A measure of betweenness centrality based on random walks*. Social networks, 27(1):39–54, 2005.

[18] Z. Nie, Y. Zhang, J. Wen, AND W. Ma, *Object-level ranking: Bringing order to web objects*. In Proc. of ACM World Wide Web, pp. 567–574, 2005.

[19] L. Page, S. Brin, R. Motwani, AND T. Winograd, *The pagerank citation ranking: Bringing order to the web*. Technical report, Stanford Digital Library Technologies Project, 1998.

[20] R. Puzis, Y. Elovici, AND S. Dolev, *Fast algorithm for successive computation of group betweenness centrality*. Phys. Rev. E, 76(5):056709, Nov 2007.

[21] D. Zhou, S. Orshanskiy, H. Zha, AND C. Giles, *Co-ranking authors and documents in a heterogeneous network*. In Proc. of IEEE International Conference on Data Mining, 2007.