MultiRank: Co-Ranking for Objects and Relations in Multi-Relational Data

Michael K. Ng
Department of Mathematics
Hong Kong Baptist University
Kowloon Tong
Hong Kong
mng@math.hkbu.edu.hk

Xutao Li and Yunming Ye
Department of Computer Science
Shenzhen Graduate School
Harbin Institute of Technology
xutaolee08@gmail.com,yeyunming@hit.edu.cn

ABSTRACT

The main aim of this paper is to design a co-ranking scheme for objects and relations in multi-relational data. It has many important applications in data mining and information retrieval. However, in the literature, there is a lack of a general framework to deal with multi-relational data for co-ranking. The main contribution of this paper is to (i) propose a framework (MultiRank) to determine the importance of both objects and relations simultaneously based on a probability distribution computed from multi-relational data; (ii) show the existence and uniqueness of such probability distribution so that it can be used for co-ranking for objects and relations very effectively; and (iii) develop an efficient iterative algorithm to solve a set of tensor (multivariate polynomial) equations to obtain such probability distribution. Extensive experiments on real-world data suggest that the proposed framework is able to provide a co-ranking scheme for objects and relations successfully. Experimental results have also shown that our algorithm is computationally efficient, and effective for identification of interesting and explainable co-ranking results.

Categories and Subject Descriptors

H.3.3 [**Information Systems**]: Information Search and Retrieval; H.2.8 [**Database Applications**]: Data mining

General Terms

Algorithms

Keywords

Ranking, multi-relational data, transition probability tensors, rectangular tensors, stationary probability distribution

1. INTRODUCTION

Evaluation of object importance or popularity is an important research problem in information retrieval that can assist in many data mining tasks, e.g., ranking query results of search engines, extracting communities in social networks and studying evolution of communities in dynamic networks. In the literature, there are many approaches to evaluating object importance [9, 3, 15, 12, 19, 8]. In these relation analysis methods, a single relation type is focused and studied. For example, both PageRank [19] and HITS [12] consider the structure of web, decompose and study the adjacency matrix that representing the hyperlink structure.

In this paper, we are interested in data with multiple relation types. There are many data mining and information retrieval applications in multi-relational data which objects have interactions with the others based on different relations. For example, researchers cite the other researchers in different conferences, and based on different concepts/topics, papers cite the other papers based on text analysis such as title, abstract, keyword and authorship [14], webpages link to the other webpages through different semantic meanings [13]. A social network [2, 17] where objects are connected via multiple relations, by their organizational structure, communication protocols, etc. This additional link structure can provide a way of incorporating multiple relations among objects into the calculation of object importance or popularity. In Figure 1, we show an example of a multi-relational data set. There are five objects and three relations (R1: green, R2: blue, R3: red) among them. We can also represent such multi-relational data set in a tensor format. A tensor is a multidimensional array. In the figure, a three-way array is used, where each two dimensional slice represents an adjacency matrix for a single relation. The data can be represented as a tensor of size $5 \times 5 \times 3$ where (i, j, k) entry is nonzero if the ith object is connected to the jth object by using the kth relation.

In the literature, there is a lack of a general framework to deal with multi-relational data or the corresponding tensor representation for co-ranking purpose. The main aim of this paper is to propose an algorithm, MultiRank, to determine the importance of both objects and relations simultaneously in multi-relational data. The MultiRank values indicate an importance of a particular object and an importance of a particular relation. In our proposal, the MultiRank of an object depends on the number and MultiRank metric of all objects that have multiple relations to this object, and also the MultiRank values of these multiple relations. An object that is linked via high MultiRank relations by objects with high MultiRanks, receives a high MultiRank itself. Similarly, the MultiRank of a relation depends on which objects

KDD'11, August 21–24, 2011, San Diego,California, USA.

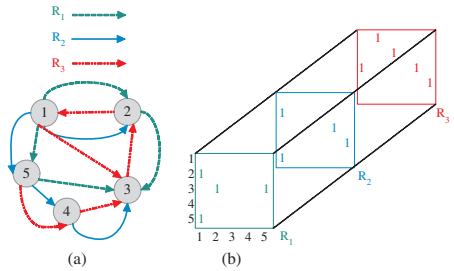


Figure 1: (a) An example of a multi-relational data in a graph representation (left) and (b) the corresponding tensor representation.

to be linked and their MultiRank values. A relation that is linked objects with high MultiRanks, receives a high MultiRank itself.

Similar to PageRank [19], our idea is to imagine infinite random surfers in a multi-relational data/tensor, and consider an equilibrium/stationary probability distribution of objects and relations as evaluation scores for objects and relations respectively. Thus we can consider that MultiRank is a stationary probability distribution that is used to represent the likelihood that we randomly visiting objects and using relations will arrive at any particular object and use any particular relation. However, instead of finding an eigenvector corresponding to the largest eigenvalue of the PageRank matrix [19] under a single relation type, our approach is to solve a set of tensor (multivariate polynomial) equations to determine a stationary probability distribution arising from a tensor that represents multiple relation types among objects.

The main contribution of this paper can be summarized as follows. (i) We propose a framework (MultiRank) to determine the importance of both objects and relations simultaneously based on a stationary probability distribution computed from multi-relational data. (ii) We show the existence and uniqueness of such stationary probability distribution so that it can be used in co-ranking for objects and relations very effectively. (iii) We develop an efficient iterative algorithm to solve a set of tensor equations to obtain such stationary probability distribution. Extensive experiments on real-world data suggest that the proposed framework is able to provide a co-ranking scheme for objects and relations based on stationary probability distribution successfully. Experimental results have also shown that our algorithm is computationally efficient, and effective for identification of interesting and explainable co-ranking results.

The rest of the paper is organized as follows. In Section 2, we describe notations in this paper and some preliminary knowledge. In Section 3, we present the proposed framework. In Section 4, we analyze the proposed methodology. In Section 5, we figure out the differences between the existing work and the proposed one. In Section 6, we show and discuss the experimental results for real-world data sets. In Section 7, we give some concluding remarks and mention some future work.

2. PRELIMINARY

In this section, we describe notations and present some preliminary knowledge on tensors. As we analyze objects under multiple relations and also consider interaction between relations based on objects, we make use of rectangular tensors to represent them.

Let R be the real field. We call $\mathcal{A} = (a_{i_1,i_2,j_1})$ where $a_{i_1,i_2,j_1} \in R$, for $i_k = 1, \cdots, m$, $k = 1, 2$ and $j_1 = 1, \cdots, n$, a real $(2,1)$th order $(m \times n)$-dimensional rectangular tensor. In this setting, we refer (i_1, i_2) to be the indices for objects and j_1 to be the index for relations. For instance, five objects $(m = 5)$ and three relations $(n = 3)$ are used in the example in Figure 1. Let $\mathbf{x}$ and $\mathbf{x}'$ be vectors of length m, and $\mathbf{y}$ be a vector of length n. Let $\mathcal{A}\mathbf{xy}$ be a vector in R^m such that

$$(\mathcal{A}\mathbf{xy})_{i_1} = \sum_{i_2=1}^{m} \sum_{j_1=1}^{n} a_{i_1,i_2,j_1} x_{i_2} y_{j_1},\ i_1 = 1, 2, \cdots, m.$$

Similarly, $\mathcal{A}\mathbf{xx}'$ is a vector in R^n such that

$$(\mathcal{A}\mathbf{xx}')_{j_1} = \sum_{i_1=1}^{m} \sum_{i_2=1}^{m} a_{i_1,i_2,j_1} x_{i_1} x'_{i_2},\ j_1 = 1, 2, \cdots, n.$$

In addition, $\mathcal{A}$ is called non-negative if $a_{i_1,i_2,j_1} \geq 0$.

As we consider infinite random surfers in a nonnegative rectangular tensor arising from multi-relational data, and study the likelihood that we will arrive at any particular object and use at any particular relation, we can construct two transition probability tensors $\mathcal{O} = (o_{i_1,i_2,j_1})$ and $\mathcal{R} = (r_{i_1,i_2,j_1})$ with respect to objects and relations by normalizing the entries of $\mathcal{A}$ as follows:

$$o_{i_1,i_2,j_1} = \frac{a_{i_1,i_2,j_1}}{\sum_{i=1}^{m} a_{i,i_2,j_1}}, \quad i_1 = 1, 2, \cdots, m,$$

$$r_{i_1,i_2,j_1} = \frac{a_{i_1,i_2,j_1}}{\sum_{j=1}^{n} a_{i_1,i_2,j}}, \quad j_1 = 1, 2, \cdots, n.$$

These numbers gives the estimates of the following conditional probabilities:

$$o_{i_1,i_2,j_1} = \text{Prob}[X_t = i_1 | X_{t-1} = i_2, Y_t = j_1]$$

$$r_{i_1,i_2,j_1} = \text{Prob}[Y_t = j_1 | X_t = i_1, X_{t-1} = i_2]$$

where X_t and Y_t are random variables referring to visit at any particular object and to use at any particular relation respectively at the time t. o_{i_1,i_2,j_1} can be interpreted as the probability of visiting the i_1th object by given that the i_2th object is currently visited and the j_1th relation is used, and r_{i_1,i_2,j_1} can be interpreted as the probability of using the j_1th relation given that the i_1th object is visited from the i_2th object.

We note that if a_{i_1,i_2,j_1} is equal to 0 for all $1 \leq i_1 \leq m$, this is called the dangling node [19], and the values of o_{i_1,i_2,j_1} can be set to $1/m$ (an equal chance to visit any object). Similarly, if a_{i_1,i_2,j_1} is equal to 0 for all $1 \leq j_1 \leq n$, then the values of r_{i_1,i_2,j_1} can be set to $1/n$ (an equal chance to use any relation). With the above construction, we have

$$0 \leq o_{i_1,i_2,j_1} \leq 1, \quad \sum_{i_1=1}^{m} o_{i_1,i_2,j_1} = 1,$$

$$0 \leq r_{i_1,i_2,j_1} \leq 1, \quad \sum_{j_1=1}^{n} r_{i_1,i_2,j_1} = 1.$$

Both $\mathcal{O}$ and $\mathcal{R}$ are nonnegative tensors. We call them transition probability tensors which are high-dimensional analog of transition probability matrices in Markov chains [20].

In addition, it is necessary for us to know the connectivity among the objects and the relations within a tensor. We remark that the concept of irreducibility has been used in the PageRank matrix in order to compute the PageRank vector [19].

DEFINITION 1. *A $(2,1)$th order nonnegative rectangular tensor $\mathcal{A}$ is called irreducible if (a_{i_1,i_2,j_1}) (m-by-m matrices) for fixed j_1 ($j_1 = 1, 2, \cdots, n$) are irreducible. If $\mathcal{A}$ is not irreducible, then we call $\mathcal{A}$ reducible.*

When $\mathcal{A}$ is irreducible, any two objects in multi-relational data can be connected via some relations. As we would like to determine the importance of both objects and relations simultaneously in multi-relational data, irreducibility is a reasonable assumption that we will use in the following analysis and discussion. It is clear that when $\mathcal{A}$ is irreducible, the two corresponding tensors $\mathcal{O}$ and $\mathcal{R}$ are also irreducible.

3. THE PROPOSED FRAMEWORK

The PageRank matrix can be regarded as a transition probability matrix of a Markov chain in a random walk. Given two transition probability tensors $\mathcal{O}$ and $\mathcal{R}$, we study the following probabilities:

$$\text{Prob}[X_t = i_1] = \sum_{i_2=1}^{m}\sum_{j_1=1}^{n} o_{i_1,i_2,j_1} \times \text{Prob}[X_{t-1} = i_2, Y_t = j_1] \tag{1}$$

$$\text{Prob}[Y_t = j_1] = \sum_{i_1=1}^{m}\sum_{i_2=1}^{m} r_{i_1,i_2,j_1} \times \text{Prob}[X_t = i_1, X_{t-1} = i_2], \tag{2}$$

where $\text{Prob}[X_{t-1} = i_2, Y_t = j_1]$ is the joint probability distribution of X_{t-1} and Y_t, and $\text{Prob}[X_t = i_1, X_{t-1} = i_2]$ is the joint probability distribution of X_t and X_{t-1}. In our approach, we consider an equilibrium/stationary distribution of objects and relations, i.e., we are interested in MultiRank values of objects and relations given by

$$\bar{\mathbf{x}} = [\bar{x}_1, \bar{x}_2, \cdots, \bar{x}_m]^T \quad \text{and} \quad \bar{\mathbf{y}} = [\bar{y}_1, \bar{y}_2, \cdots, \bar{y}_n]^T$$

respectively, with

$$\bar{x}_{i_1} = \lim_{t\to\infty} \text{Prob}[X_t = i_1] \quad \text{and} \quad \bar{y}_{j_1} = \lim_{t\to\infty} \text{Prob}[Y_t = j_1].$$

for $1 \le i_1 \le m$ and $1 \le j_1 \le n$.

In general, it may be difficult to obtain $\bar{x}_{i_1}$ and $\bar{y}_{j_1}$ as (1) and (2) are coupled together and they involves two joint probability distributions. Here we employ a product form of individual probability distributions for joint probability distributions in (1) and (2). More precisely, we assume that

$$\text{Prob}[X_{t-1} = i_2, Y_t = j_1] = \text{Prob}[X_{t-1} = i_2]\text{Prob}[Y_t = j_1] \tag{3}$$

$$\text{Prob}[X_t = i_1, X_{t-1} = i_2] = \text{Prob}[X_t = i_1]\text{Prob}[X_{t-1} = i_2]. \tag{4}$$

Therefore, by using the above assumptions and considering t goes to infinity, (1) and (2) becomes

$$\bar{x}_{i_1} = \sum_{i_2=1}^{m}\sum_{j_1=1}^{n} o_{i_1,i_2,j_1}\bar{x}_{i_2}\bar{y}_{j_1}, \quad i_1 = 1, 2, \cdots, m, \tag{5}$$

$$\bar{y}_{j_1} = \sum_{i_1=1}^{m}\sum_{i_2=1}^{m} r_{i_1,i_2,j_1}\bar{x}_{i_1}\bar{x}_{i_2}; \quad j_1 = 1, 2, \cdots, n. \tag{6}$$

We see from (5) and (6) that the MultiRank of an object depends on the number and MultiRank metric of all objects that have multiple relations to this object, and also the MultiRank values of these multiple relations. Similarly, the MultiRank of a relation depends on which the objects to be linked and their MultiRank values of these objects. It is clear that an object that is linked via high MultiRank relations by objects with high MultiRank, receives a high MultiRank itself. Also a relation that is linked objects with high MultiRank values, receives a high MultiRank itself. Under the tensor operation for (5) and (6), we solve the following tensor (multivariate polynomial) equations:

$$\mathcal{O}\bar{\mathbf{x}}\bar{\mathbf{y}} = \bar{\mathbf{x}} \quad \text{and} \quad \mathcal{R}\bar{\mathbf{x}}^2 = \bar{\mathbf{y}}, \tag{7}$$

with

$$\sum_{i_1=1}^{m} \bar{x}_{i_1} = 1 \quad \text{and} \quad \sum_{j_1=1}^{n} \bar{y}_{j_1} = 1 \tag{8}$$

to obtain the MultiRank values of objects and relations.

We remark that the normalized eigenvector (corresponding to the largest eigenvalue 1) in the PageRank computation can be interpreted as the stationary probability distribution vector of the associated Markov chain [19]. When we consider a single relation type, we can set $\bar{\mathbf{y}}$ to be a vector $\frac{1}{n}\mathbf{1}$ (equal chance of all relations) in (7), and thus we obtain a matrix equation $\mathcal{O}\bar{\mathbf{x}}\frac{1}{n}\mathbf{1} = \bar{\mathbf{x}}$. This is exactly the same as that we solve for the normalized eigenvector to get the PageRank vector. As a summary, the proposed framework MultiRank is a generalization of PageRank to deal with multi-relational data.

3.1 The Algorithm

In this subsection, we present an efficient iterative algorithm to solve the tensor equations in (7) to obtain $\bar{\mathbf{x}}$ and $\bar{\mathbf{y}}$ for the MultiRank values of objects and relations. The MultiRank algorithm is summarized in the following algorithm.

Algorithm 1 The MultiRank Algorithm

Input: Two tensors $\mathcal{O}$ and $\mathcal{R}$, two initial probability distributions $\mathbf{x}_0$ and $\mathbf{y}_0$ ($\sum_{i_1=1}^{m}[\mathbf{x}_0]_{i_1} = 1$ and $\sum_{j_1=1}^{n}[\mathbf{y}_0]_{j_1} = 1$) and the tolerance ϵ

Output: Two stationary probability distributions $\bar{\mathbf{x}}$ and $\bar{\mathbf{y}}$

Procedure:

1: Set $k = 1$;
2: Compute $\mathbf{x}_k = \mathcal{O}\mathbf{x}_{k-1}\mathbf{y}_{k-1}$;
3: Compute $\mathbf{y}_k = \mathcal{R}\mathbf{x}_k^2$;
4: If $||\mathbf{x}_k - \mathbf{x}_{k-1}|| + ||\mathbf{y}_k - \mathbf{y}_{k-1}|| < \epsilon$, then stop, otherwise set $k = k + 1$ and goto Step 2.

In Algorithm 1, the MultiRank computations require several iterations, through the collection to adjust approximate MultiRank values of objects and relations to more closely reflect their theoretical true values (underlying stationary probability distributions). The iterative method is similar

to the power method for computing the eigenvector corresponding to the largest eigenvalue of a matrix [19]. The main computational cost of the MultiRank algorithm depends on the cost of performing tensor operations in Steps 2 and 3. Assume that there are $O(N)$ nonzero entries (sparse data) in $\mathcal{O}$ and $\mathcal{R}$, the cost of these tensor calculations are of $O(N)$ arithmetic operations.

4. THEORETICAL ANALYSIS

In this section, we show the existence and uniqueness of stationary probability distributions $\bar{\mathbf{x}}$ and $\bar{\mathbf{y}}$ so that it can be used in co-ranking for objects and relations very effectively. Based on these results, the convergence of MultiRank algorithm can be shown.

We let $\Omega_m = \{\mathbf{x} = (x_1, x_2, \cdots, x_m) \in R^m | x_i \geq 0,\ 1 \leq i \leq m,\ \sum_{i=1}^m x_i = 1\}$ and $\Omega_n = \{\mathbf{y} = (y_1, y_2, \cdots, y_n) \in R^n | y_j \geq 0,\ 1 \leq j \leq n,\ \sum_{j=1}^n y_j = 1\}$. We also set $\Omega = \{[\mathbf{x}, \mathbf{y}] \in R^{m+n} | \mathbf{x} \in \Omega_m, \mathbf{y} \in \Omega_n\}$. We note that Ω_m, Ω_n and Ω are closed convex sets. We call $\mathbf{x}$ and $\mathbf{y}$ to be positive (denoted by $\mathbf{x} > 0$ and $\mathbf{y} > 0$) if all their entries are positive.

It is easy to check that if both $\mathbf{x}$ and $\mathbf{y}$ are probability distributions, then the output $\mathcal{O}\mathbf{x}\mathbf{y}$ and $\mathcal{R}\mathbf{x}^2$ are also probability distributions (the correctness of Steps 2 and 3 in the MultiRank Algorithm).

THEOREM 1. *Suppose $\mathcal{O}$ and $\mathcal{R}$ are constructed in Section 3. For any $\mathbf{x} \in \Omega_m$ and $\mathbf{y} \in \Omega_n$, then $\mathcal{O}\mathbf{x}\mathbf{y} \in \Omega_m$ and $\mathcal{R}\mathbf{x}^2 \in \Omega_n$.*

By using Theorem 1, we show the existence of positive solutions for the set of tensor equations in (7) and (8).

THEOREM 2. *Suppose $\mathcal{O}$ and $\mathcal{R}$ are constructed in Section 3. If $\mathcal{O}$ and $\mathcal{R}$ are irreducible, then there exist $\bar{\mathbf{x}} \in \Omega_m$ and $\bar{\mathbf{y}} \in \Omega_n$ such that $\mathcal{O}\bar{\mathbf{x}}\bar{\mathbf{y}} = \bar{\mathbf{x}}$ and $\mathcal{R}\bar{\mathbf{x}}^2 = \bar{\mathbf{y}}$, and both $\bar{\mathbf{x}}$ and $\bar{\mathbf{y}}$ are positive.*

PROOF. The problem can be reduced to a fixed point problem as follows. We define the following mapping $T : \Omega \to \Omega$ as follows

$$T([\mathbf{x}, \mathbf{y}]) = [\mathcal{O}\mathbf{x}\mathbf{y}, \mathcal{R}\mathbf{x}^2]. \tag{9}$$

It is clear that T is well-defined (i.e., when $[\mathbf{x}, \mathbf{y}] \in \Omega$, $T([\mathbf{x}, \mathbf{y}]) \in \Omega$) and continuous. According to the Brouwer Fixed Point Theorem, there exists $[\bar{\mathbf{x}}, \bar{\mathbf{y}}] \in \Omega$ such that $T([\bar{\mathbf{x}}, \bar{\mathbf{y}}]) = [\bar{\mathbf{x}}, \bar{\mathbf{y}}]$, i.e., $\mathcal{O}\bar{\mathbf{x}}\bar{\mathbf{y}} = \bar{\mathbf{x}}$ and $\mathcal{R}\bar{\mathbf{x}}^2 = \bar{\mathbf{y}}$.

Next we show that $\bar{\mathbf{x}}$ and $\bar{\mathbf{y}}$ are positive. Suppose $\bar{\mathbf{x}}$ and $\bar{\mathbf{y}}$ are not positive, i.e., there exist some entries of $\bar{\mathbf{x}}$ are zero and some entries of $\bar{\mathbf{y}}$ are zero. Let $I = \{i_1 | \bar{x}_{i_1} = 0\}$ and $J = \{j_1 | \bar{y}_{j_1} = 0\}$. It is obvious that I is a proper subset of $\{1, 2, \cdots, m\}$ and J is a proper subset of $\{1, 2, \cdots, n\}$. Let $\delta = \min\{\min\{\bar{x}_{i_1} | i_1 \notin I\}, \min\{\bar{y}_{j_1} | j_1 \notin J\}\}$. We must have $\delta > 0$. Since $\bar{\mathbf{x}}$ and $\bar{\mathbf{y}}$ satisfies $\mathcal{O}\bar{\mathbf{x}}\bar{\mathbf{y}} = \bar{\mathbf{x}}$, we have

$$\sum_{i_2=1}^{m} \sum_{j_1=1}^{n} o_{i_1,i_2,j_1} \bar{x}_{i_2} \bar{y}_{j_1} = \bar{x}_{i_1} = 0, \quad \forall i_1 \in I.$$

Let us consider the following quantity:

$$\begin{aligned} \delta^2 \sum_{i_2 \notin I} \sum_{j_1 \notin J} o_{i_1,i_2,j_1} &\leq \sum_{i_2 \notin I} \sum_{j_1 \notin J} o_{i_1,i_2,j_1} \bar{x}_{i_2} \bar{y}_{j_1} \\ &\leq \sum_{i_2=1}^{m} \sum_{j_1=1}^{n} o_{i_1,i_2,j_1} \bar{x}_{i_2} \bar{y}_{j_1} = 0, \end{aligned}$$

for all $i_1 \in I$. Hence we have $o_{i_1,i_2,j_1} = 0$ for all $i_1 \in I$ and for all $i_2 \notin I$ for any fixed $j_1 \notin J$. Thus the matrices (o_{i_1,i_2,j_1}) (for $j_1 \notin J$) are reducible. It implies that $\mathcal{O}$ is reducible. By using the similar argument and considering the equation $\mathcal{R}\mathbf{x}^2 = \mathbf{y}$, we can find that $\mathcal{R}$ is also reducible. According to these results, we obtain a contradiction. Hence both $\bar{\mathbf{x}}$ and $\bar{\mathbf{y}}$ must be positive. □

In [11], it has been given a general condition which guarantees the uniqueness of the fixed point in the Brouwer Fixed Point Theorem, namely, (i) 1 is not an eigenvalue of the Jacobian matrix of the mapping, and (ii) for each point in the boundary of the domain of the mapping, it is not a fixed point. In our case, we have shown in Theorem 2 that all the fixed points of T are positive when $\mathcal{O}$ and $\mathcal{R}$ are irreducible, i.e., they do not lie on the boundary $\partial\Omega$ of Ω. The Jacobian matrix of T is an $(m+n)$-by-$(m+n)$ matrix:

$$DT([\mathbf{x}, \mathbf{y}]) = \begin{pmatrix} DT_{11}([\mathbf{x}, \mathbf{y}]) & DT_{12}([\mathbf{x}, \mathbf{y}]) \\ DT_{21}([\mathbf{x}, \mathbf{y}]) & 0 \end{pmatrix}, \quad [\mathbf{x}, \mathbf{y}] \in \Omega,$$

where $DT_{11}([\mathbf{x}, \mathbf{y}]) = (\sum_{j_1=1}^n o_{i_1,i_2,j_1} y_{j_1})_{i_1,i_2}$ is an m-by-m matrix corresponding to the derivative of $\mathcal{O}\mathbf{x}\mathbf{y}$ with respect to $\mathbf{x}$, $DT_{12}([\mathbf{x}, \mathbf{y}]) = (\sum_{i_2=1}^m o_{i_1,i_2,j_1} x_{i_2})_{i_1,j_1}$ is an m-by-n matrix corresponding to the derivative of $\mathcal{O}\mathbf{x}\mathbf{y}$ with respect to $\mathbf{y}$, and $DT_{21}([\mathbf{x}, \mathbf{y}]) = (\sum_{i_1=1}^m r_{i_1,i_2,j_1} x_{i_1})_{j_1,i_2} + (\sum_{i_2=1}^m r_{i_1,i_2,j_1} x_{i_2})_{j_1,i_1}$ is an n-by-m matrix corresponding to the derivative of $\mathcal{R}\mathbf{x}^2$ with respect to $\mathbf{x}$. To conclude, we have the following theorem:

THEOREM 3. *Suppose $\mathcal{O}$ and $\mathcal{R}$ are constructed in Section 3, and they are irreducible. If 1 is not the eigenvalue of $DT([\mathbf{x}, \mathbf{y}])$ for all $[\mathbf{x}, \mathbf{y}] \in \Omega/\partial\Omega$, then then the solution vectors $\bar{\mathbf{x}}$ and $\bar{\mathbf{y}}$ in Theorem 2 are unique.*

According to Theorem 3, when $\mathbf{x}_k = \mathbf{x}_{k-1}$ and $\mathbf{y}_k = \mathbf{y}_{k-1}$ in the MultiRank algorithm, then we obtain the unique solution vectors $\bar{\mathbf{x}}$ and $\bar{\mathbf{y}}$ for $\mathcal{O}\mathbf{x}\mathbf{y} = \mathbf{x}$ and $\mathcal{R}\mathbf{x}^2 = \mathbf{y}$. When $\mathbf{x}_k \neq \mathbf{x}_{k-1}$ and $\mathbf{y}_k \neq \mathbf{y}_{k-1}$, there exist a subsequence $[\mathbf{x}_{k_s}, \mathbf{y}_{k_s}]$ converges to $[\bar{\mathbf{x}}, \bar{\mathbf{y}}]$ by using the fact that Ω is compact in R^{m+n}. As we have shown that the solution vectors are unique, it implies that $[\mathbf{x}_k, \mathbf{y}_k]$ converges (up to a subsequence) to $[\bar{\mathbf{x}}, \bar{\mathbf{y}}]$ which are the stationary probability vectors giving MultiRank values of objects and relations respectively for co-ranking purpose effectively.

5. RELATED WORK

Recently, co-ranking has been studied in [7, 26]. In [7], a co-HITS algorithm is proposed to incorporate the bipartite graph with the content information from both sides as well as the constraints of relevance. In [26], a method for co-ranking authors and their publications using several networks is proposed and based on coupling two random walks that separately rank authors and documents following the PageRank paradigm. In these two methods, the score propagation is the mutual reinforcement to boost co-linked entities on the graph only. We see from (5) and (6) that our method involves mutual reinforcement among all the objects and the relations. PopRank [18] uses the PageRank framework by adding popularity to each link pointing to an object. The main challenge of this approach is required to add popularity which is not known in general. The novelty of the proposed method is to provide a framework to determine the weights of relations automatically.

In the literature, tensor factorization is a generalized approach for analyzing multi-way interactions among entities. For instance, Lin et al. [17] proposed a novel relational hypergraph representation for modeling multi-relational and multi-dimensional social data and studied an efficient factorization method for community extraction on a given metagraph. Sun et al. proposed a general and efficient framework for analyzing high-order and high-dimensional stream data [21, 22]. Sun et al. [23] applied a 3-way Tucker decomposition [24] to the analysis of user, query-term and webpage data in order to personalize web search. Acar et al. [1] used various tensor decompositions of user, keyword and time data to separate different streams of conservations in chatroom data. Kolda et al. [14, 13] proposed TOPHITS by adding a third dimension to form an adjacency tensor that incorporates anchor text information, and then to increase the likelihood that the principal singular vectors relate to the query. They also employed a Three-way Parallel Factors (PARAFAC) decomposition [4, 10] to compute the singular vectors for query processing. In these methods, we need to select the number of decompositions (low-rank approximation) in the tensor factorization. The number of decompositions may not be known in advance. On the other hand, the computation of tensor factorization may not be unique as there are several numerical methods (e.g., the alternating least squares procedure) used to compute such factorization and the factorization results depend on the initial guess. There is no detailed algorithmic and mathematical analysis for the convergence of the method. Also the computational cost may be expensive for very large tensors [6]. Different from these methods, we compute stationary probability distributions of tensors for scores of objects and relations to handle the ranking problem. We have shown that such probability distributions can be unique and computed very efficiently.

6. EXPERIMENTAL RESULTS

In this section, we report three experiments to show the effectiveness of the proposed model for co-ranking objects and relations from real-world data sets. We crawled publication information of five conferences (SIGKDD, WWW, SIGIR, SIGMOD, CIKM) from DBLP[1]. Their publication periods are as follows: SIGKDD (1999-2010), WWW (2001-2010), SIGIR (2000-2010), SIGMOD (2000-2010) and CIKM (2000 and 2002-2009) [2]. Publication information includes title, authors, reference list, and classification categories associated with this publication[3]. There are in total 6848 publications, 10305 authors and 617 different categories in the data set. Based on this multi-relational data, we construct different types of tensors in the following three experiments.

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

[2]Missing information for CIKM 2001 is due to fact that DBLP does not provide links to ACM Digital Library.

[3]For each publication, there are several strings indicating the classification categories of this publication, where each string provides the information from the most general concept to the most specific concept. For example, a string may be "H. information systems—>H.3 information storage and retrieval—>H.3.3 information search and retrieval". For each string, we choose the most specific concept as the classification category it indicates for the publication.

6.1 Experiment 1

Tensor construction. In this experiment, we construct a tensor $\mathcal{A}$ based on citations of authors (objects) through different category concepts (relations). In this case there are 10305 objects and 617 relations. The tensor is constructed as follows: If a publication written by the i_2th author cites a publication written by the i_1th author, and the two publications have the same j_1th category concept, then we can add one to the entry a_{i_1,i_2,j_1} of $\mathcal{A}$. By considering all the publications, a_{i_1,i_2,j_1} refers to the number of citations of publications written by the i_2th author to publications written by the i_1th author where these publications have the same j_1th category concept. Here we do not consider any self-citation, i.e., $a_{i_1,i_1,j_1} = 0$ for all $1 \leq i_1 \leq 10305$ and $1 \leq j_1 \leq 617$. The purpose is to avoid an ambiguous high self-citations in $\mathcal{A}$. The size of $\mathcal{A}$ is $10305 \times 10305 \times 617$ and there are 39851 nonzeros entries in $\mathcal{A}$. The percentage of nonzero entries is $6.08 \times 10^{-5}\%$, and thus $\mathcal{A}$ is sparse. After we construct $\mathcal{A}$, we can generate both transition probability tensors $\mathcal{O}$ and $\mathcal{R}$. The only nonzero entries and their locations are stored in the computational process. It is not necessary to store values $1/m$ or $1/n$ for the dangling nodes in $\mathcal{O}$ and $\mathcal{R}$.

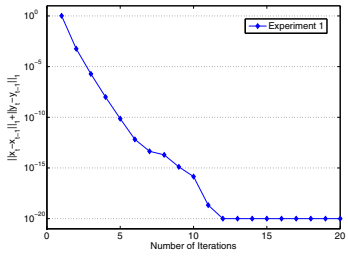
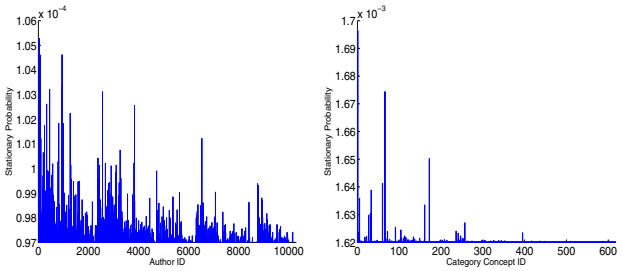


Figure 2: The difference between two successive calculated probability vectors against iterations.



(a) (b)

Figure 3: The stationary probability vectors (a) authors; (b) category concepts in Experiment 1.

Results and discussion. Figure 2 shows the convergence of the MultiRank algorithm. We see from the figure that the changes of stationary probabilities for objects and relations, $||\mathbf{x}_k - \mathbf{x}_{k-1}||_1 + ||\mathbf{y}_k - \mathbf{y}_{k-1}||_1$, decreases when iteration number increases, and the successive difference after 12 iterations, is less than 10^{-20} which is small enough for convergence criterion. The computational time is about 114 seconds under the MATLAB implementation in a standard desktop PC. In Figures 3(a)-(b), we show the two resulting stationary probability vectors corresponding to the authors

and the category concepts respectively. It is clear in the figures that some authors (relations) have higher stationary probabilities (MultiRank values). These results indicate that some authors (category concepts) are more important (popular) than the others.

ranking	author name	ranking	category concept
1	Bing Liu	1	data mining
2	Pedro Domingos	2	information search and retrieval
3	Wei-Ying Ma	3	retrieval models
4	Jiawei Han	4	search process
5	Philip S. Yu	5	query processing
6	ChengXiang Zhai	6	clustering
7	Thorsten Joachims	7	miscellaneous
8	W. Bruce Croft	8	query formulation
9	Matthew Richardson	9	information filtering
10	Susan T. Dumais	10	performance evaluation

Table 1: The top ten authors (left) and category concepts (right) in Experiment 1.

ranked author/concept	1	2	3	4	5	6	7	8	9	10
1	122	0	0	4	0	2	20	0	14	0
2	180	0	0	0	0	0	0	0	0	0
3	41	111	57	36	0	34	15	0	12	0
4	377	0	0	0	9	0	2	0	0	0
5	138	5	0	10	21	51	0	0	8	0
6	0	151	149	21	0	25	0	0	7	0
7	0	29	65	60	0	0	0	0	0	0
8	0	80	128	2	0	0	0	64	0	4
9	75	39	0	2	0	0	0	0	3	0
10	0	46	0	34	0	0	0	5	13	0

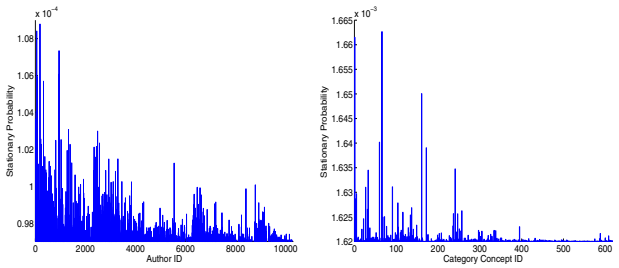
Table 2: The numbers that the top ten authors are cited by the others via the top ten category concepts.

In Table 1, we show the top ten authors and category concepts based on their MultiRank values. As there are two types of conferences in the data set, one is based on data mining like SIGKDD and SIGMOD, and the other is based on information retrieval like WWW, SIGIR and CIKM. Therefore, these two concepts are ranked numbers one and two respectively. In Table 2, we show the numbers that the top ten authors are cited by the others via the top ten category concepts. We find that there are many citations for the cited authors (ranked numbers 1 to 5) via the data mining concept (ranked number 1), and there are many citations for the cited authors (ranked numbers 6 to 10) via the information search and retrieval and retrieval model concepts (ranked numbers 2 and 3). This observation can explain why the first five authors have higher MultiRank values than the last five authors. Also there are more citations based on the first concept for the cited authors (ranked numbers 2, 4 and 5) than for the ranked number one author in Table 2, we find in (7) that his MultiRank value by summing the scores of his linked objects and used relations together, is higher than those of the other cited authors.

Next we compare the MultiRank and PageRank results. In the PageRank algorithm, we aggregate all the citations of different concepts together to construct relations among the authors. In the PageRank results, Philip Yu, Thorsten Joachims and Susan T. Dumais do not appear in the list of the top ten cited authors, and the others still appear. Now they are ranked numbers 14, 29, and 13 respectively. However, there are three new persons: D. R. Mani (rank #8), James Drew (ranked #9) and Andrew Betz (rank #10), who are co-authors of papers cited by Pedro Domingos (rank #1) in the PageRank results. Because Pedro Domingos has the highest PageRank score and cites less papers in the data set, these three persons also receive high PageRank scores even their citations are not many. Because the MultiRank algorithm gives the scores for concepts, it differentiates authors and concepts, and provides a more accurate and comparable ranking results than those by the PageRank algorithm.

6.2 Experiment 2

Tensor construction. In this experiment, we construct a $(2,1)$th order tensor $\mathcal{A}$ based on author-author collaborations through different category concepts. When the i_1th and i_2th authors publish a paper together under the j_1th category concept, we add one to the entries a_{i_1,i_2,j_1} and a_{i_2,i_1,j_1} of $\mathcal{A}$. In this case, $\mathcal{A}$ is symmetric with respect to the index j_1. By considering all the publications, a_{i_1,i_2,j_1} (or a_{i_2,i_1,j_1}) refers to the number of collaborations by the i_1th and the i_2th authors under the j_1th category concept. Again, we do not consider any self-collaboration, i.e., $a_{i_1,i_1,j_1} = 0$ for all $1 \le i_1 \le 10305$ and $1 \le j_1 \le 617$. The size of $\mathcal{A}$ is $10305 \times 10305 \times 617$ and there are 95722 nonzeros entries in $\mathcal{A}$. The percentage of nonzero entries is $1.46 \times 10^{-4}\%$.



(a) (b)

Figure 4: The stationary probability vectors (a) authors; (b) category concepts in Experiment 2.

ranking	author name	ranking	category concept
1	C. Lee Giles	1	information search and retrieval
2	Philip S. Yu	2	data mining
3	Wei-Ying Ma	3	miscellaneous
4	Zheng Chen	4	search process
5	Jiawei Han	5	retrieval models
6	Christos Faloutsos	6	general
7	Bing Liu	7	query processing
8	Johannes Gehrke	8	web-based services
9	Gerhard Weikum	9	information filtering
10	Elke A. Rundensteiner	10	clustering

Table 3: The top ten authors (left) and category concepts (right) in Experiment 2.

ranked author/concept	1	2	3	4	5	6	7	8	9	10
1	31	5	25	20	17	6	0	5	12	18
2	6	71	3	5	0	8	14	4	7	17
3	38	29	11	54	34	0	0	7	24	27
4	93	11	22	44	22	5	0	8	14	8
5	10	126	7	0	7	18	7	4	2	0
6	0	80	5	11	11	3	4	0	0	2
7	0	38	11	4	9	0	0	4	9	6
8	0	13	6	0	0	6	22	0	0	2
9	8	0	26	19	12	8	1	5	2	0
10	0	5	12	0	0	4	26	0	0	0

Table 4: The numbers that the top ten authors collaborate with the others via the top ten category concepts.

Results and discussion. We apply the MultiRank algorithm to the constructed tensors, and the algorithm converges in 10 iterations. In Figures 4(a) and 4(b), we show the resulting stationary probabilities for the authors and category concepts. Based on the MultiRank values, we show in Table 3 the top ten authors and category concepts. When we compare the results in Experiments 1 and 2, we find that the top ten category concepts changes slightly, i.e., the

concepts of query formulation and performance evaluation disappear in Experiment 2 while the concept of general and web-based services appear. This phenomena may suggest the former two concepts are more likely to be used when authors cite the others, while the latter concept are more likely to be used when authors collaborate together. The top ten authors also changes significantly, e.g., Philip S. Yu, Wei-Ying Ma, Jiawei Han and Bing Liu remain in the list of the top ten collaborated authors. This phenomena suggests these four authors not only have high citations via the mostly used concepts, but also have more collaborations via the mostly used concepts in the five conferences. It is clear that the other top cited authors in Table 1(a) have less collaborations in the five conferences. In Table 4, we show the numbers that the top ten authors collaborate with the others via the top ten category concepts. We see from the table that there are many collaborations for the authors (ranked numbers 1 to 5) via the top ten concepts, and there are less collaborations for the authors (ranked numbers 6 to 10) via the top ten concepts.

6.3 Experiment 3

Tensor construction. In this experiment, we co-rank both authors and their papers. We need to construct a $(2,2)$th order tensor for calculating MultiRanks of both authors and papers. Our idea is that if a paper or an author has a high MultiRank value, then the publications or the authors cited by this paper/author have a high chance to obtain high MultiRank values (see the tensor equations in (10)). There are four indices of the entry of $\mathcal{A}$: a_{i_1,i_2,j_1,j_2}. The first two indices (for objects) refer to the authors and the last two indices (for relations) refer to the papers. Here the role of objects and relations can be swapped. We set a_{i_1,i_2,j_1,j_2} to be 1 (or 0) when the j_2th paper written by the i_2th author cite (do not cite) the j_1th paper written by the i_1th author.

The five conferences may have different topics or tasks, e.g., WWW, SIGIR, and CIKM mainly focus on tasks related to information retrieval, while SIGKDD and SIGMOD mainly focus on tasks related data mining or databases. Here we construct three different tensors $\mathcal{A}$ based on the data from (a) all the five conferences; (b) SIGKDD and SIGMOD conferences; and (c) WWW, SIGIR and CIKM conferences. For each case, we construct the tensor as follows. Firstly, we select the top thirty cited authors in Experiment 1, who have at least ten papers in the conferences. For each selected author, we select his/her ten most cited papers by the other papers in the same set of conferences. Again we do not consider any self-citation, i.e., we set $a_{i_1,i_1,j_1,j_2} = 0$ for all i_1, j_1 and j_2. The sizes of $\mathcal{A}$ are $30 \times 30 \times 261 \times 261$, $30 \times 30 \times 252 \times 252$ and $30 \times 30 \times 258 \times 258$ for the cases (a), (b) and (c) respectively. They have different sizes as there are some overlapped papers in the construction. There are 504, 458, 493 nonzeros entries in $\mathcal{A}$ for the case (a), (b) and (c) respectively. The corresponding percentages of nonzero entries are $8.22 \times 10^{-4}\%$, $8.01 \times 10^{-4}\%$ and $8.23 \times 10^{-4}\%$.

In this experiment, we determine the MultiRanks of authors $\bar{\mathbf{x}}$ and papers $\bar{\mathbf{y}}$ by solving the following tensor equations:

$$\mathcal{O}\bar{\mathbf{x}}\bar{\mathbf{y}}^2 = \bar{\mathbf{x}} \quad \text{and} \quad \mathcal{R}\bar{\mathbf{x}}^2\bar{\mathbf{y}} = \bar{\mathbf{y}}, \tag{10}$$

with $\sum_{i_1=1}^{m} \bar{x}_{i_1} = 1$ and $\sum_{j_1=1}^{n} \bar{y}_{j_1} = 1$. Based on the idea in (7), we can interpret the first equation in (10) by viewing paper-paper citations $\bar{\mathbf{y}}^2$ as relations to give authors' scores $\bar{\mathbf{x}}$ as objects, and the second equation in by viewing author-author citations $\bar{\mathbf{x}}^2$ as relations to give papers' scores $\bar{\mathbf{y}}$ as objects. We can employ the same MultiRank algorithm in Section 4.1 to solve the equations, except we compute $\mathbf{x}_k = \mathcal{O}\mathbf{x}_{k-1}\mathbf{y}_{k-1}^2$ and $\mathbf{y}_k = \mathcal{R}\mathbf{x}_k^2\mathbf{y}_{k-1}$ in the iterative process. By using the similar arguments in Section 5, we can show existence and uniqueness of the solution in (10) and the algorithm can converge to the unique solution.

Results and discussion. We apply the MultiRank algorithm to three different cases, and obtain the stationary probabilities (MultiRanks) of the authors and their papers. The resulting top ten authors and top ten papers for the three cases are shown as in Table 5. We see from the table that there are five top cited authors appearing in both Experiments 3(a) and 3(c), and there are only two top cited authors appearing in both Experiments 3(a) and 3(b). Also there are five top cited papers appearing in both Experiments 3(a) and 3(c), and there is only one top cited paper appearing in both Experiments 3(a) and 3(b). These results suggest that the data set may be imbalanced in the areas of data mining and information retrieval, and it is reasonable to perform co-ranking for two separated cases (b) and (c).

Moreover, we see from the Table 5 that the top ten cited authors and the top ten cited papers are strongly related in the three experiments, i.e., most of the top ten cited authors have papers in the top ten cited papers, and most of the top ten cited papers are written by the top ten cited authors. This phenomena implies that our co-ranking results are reasonable as an author (or a paper) cited by other papers/authors with high MultiRanks, receives a high MultiRank value. To further verify the co-ranking results, we show in Table 6 the numbers that the top ten cited authors (or the top ten cited papers) are cited by the others of different rankings in Experiment 3(a). We see that the authors who are ranked higher usually have either more citations by highly-ranked authors and more citations in total. Similar findings are observed for the top ten cited papers. These observations reflect the design of the MultiRank paradigm by setting equations in (10). Similar observations can also be found in Experiment 3(b) and 3(c).

Experiment 3(a) author				Experiment 3(a) paper			
ranking	(1-10)	(11-20)	(21-30)	ranking	(1-100)	(101-200)	(others)
1	18	24	17	1	16	8	4
2	12	13	2	2	12	8	2
3	13	8	14	3	11	3	6
4	8	13	5	4	9	2	9
5	11	6	13	5	4	6	6
6	18	6	8	6	0	12	0
7	0	19	8	7	6	5	0
8	3	7	9	8	8	8	0
9	8	4	6	9	6	3	2
10	6	12	4	10	3	6	2

Table 6: The numbers that the top ten cited authors (papers) are cited by the other authors (papers) of different rankings in Experiments 3(a).

It is interesting to note that Thorsten Joachims and Jon M. Kleinberg who are ranked number one and number two in Experiment 3(a), disappear in both Experiments 3(b) and 3(c). They are the top cited authors for all five conferences, however, they are not the top cited authors either in data mining conferences or in information retrieval conferences. For Thorsten Joachims, this is because he does not have ten publications either in data mining conferences (Experiment 3(b)) or information retrieval conferences (Experiment

Experiment 3(a)		
ranking	author name	papers
1	Thorsten Joachims	1
2	Jon M. Kleinberg	7
3	Wei-Ying Ma	3, 6, 8
4	Susan T. Dumais	5, 9
5	Zheng Chen	3, 6
6	Ji-Rong Wen	2, 8
7	Jiawei Han	4
8	W. Bruce Croft	none
9	Pedro Domingos	none
10	Andrew Tomkins	10

Experiment 3(a)		
ranking	paper title	authors
1	Optimizing search engines using clickthrough data.	1
2	Clustering user queries of a search engine.	6
3	Optimizing web search using web click-through data.	3, 5
4	Mining frequent patterns without candidate generation.	7
5	Improving web search ranking by incorporating user behavior information.	4
6	A study of relevance propagation for web search.	3, 5
7	Bursty and hierarchical structure in streams.	2
8	Probabilistic query expansion using query logs.	3, 6
9	Learning user interaction models for predicting web search result preferences.	4
10	On the bursty evolution of blogspace.	10

Experiment 3(b)		
ranking	author name	papers
1	Pedro Domingos	5
2	Philip S. Yu	1, 2, 9
3	Jiawei Han	3, 9
4	David J. DeWitt	5, 8
5	Johannes Gehrke	4
6	Mohammed Javeed Zaki	none
7	Jian Pei	3
8	Joseph M. Hellerstein	7
9	Rajeev Rastogi	4
10	Jeffrey F. Naughton	8

Experiment 3(b)		
ranking	paper title	authors
1	Clustering by pattern similarity in large data sets.	2
2	Mining asynchronous periodic patterns in time series data.	2
3	Mining frequent patterns without candidate generation.	3, 7
4	Processing complex aggregate queries over data streams.	5, 9
5	Mining high-speed data streams.	1
6	NiagaraCQ: A scalable continuous query system for internet databases.	4
7	Continuously adaptive continuous queries over streams.	8
8	On supporting containment queries in relational database management systems.	4, 10
9	Graph indexing: A frequent structure-based approach.	2, 3
10	Mining association rules with multiple minimum supports.	none

Experiment 3(c)		
ranking	author name	papers
1	W. Bruce Croft	6
2	Susan T. Dumais	1, 8
3	Rosie Jones	5
4	Zheng Chen	2, 3
5	ChengXiang Zhai	none
6	Wei-Ying Ma	2, 3, 10
7	Andrew Tomkins	9
8	James P. Callan	none
9	Tie-Yan Liu	2, 7
10	Eugene Agichtein	1, 8

Experiment 3(c)		
ranking	paper title	authors
1	Improving web search ranking by incorporating user behavior information.	2, 10
2	A study of relevance propagation for web search.	4, 6, 9
3	Optimizing web search using web click-through data.	4, 6
4	Clustering user queries of a search engine.	none
5	Generating query substitutions.	3
6	Predicting query performance.	1
7	Adapting ranking SVM to document retrieval.	9
8	Learning user interaction models for predicting web search result preferences.	2, 10
9	Propagation of trust and distrust.	7
10	Learning block importance models for web pages.	6

Table 5: The top ten authors (left) and the top ten papers (right) in Experiment 3(a), 3(b) and 3(c). The column of papers in the left hand side table indicates associated top ten papers. The column of authors in the right hand side table indicates associated top ten authors. The authors who (or papers which) appear in the top ten of both Experiments 3(a) and 3(b) or Experiments 3(a) and 3(c), are indicated with green color (or blue color).

3(c)). When we intentionally add him in Experiment 3(c) and re-run the MultiRank algorithm, he appears in the list of the top ten cited authors, and he is ranked the number three. For Jon M. Kleinberg, even though he has enough publications in Experiment 3(b), he disappears in the list of the top ten cited authors. The main reason is that his high cited paper titled "Bursty and hierarchical structure in streams" published in SIGKDD conference, is more related to information retrieval conferences. Therefore, he does not have many citations from data mining conferences in Experiment 3(b).

7. CONCLUDING REMARKS

In this paper, we have proposed a framework (MultiRank) to determine the importance of both objects and relations simultaneously based on a probability distribution computed from multi-relational data. Both theoretical and experimental results have demonstrated that the proposed algorithm is efficient and effective. Here we give several future research work based on the proposed framework.

(i) In the framework, we assume probability distributions satisfying (3) and (4). It is interesting to employ the other possible forms to set up other tensor equations, compute and analyze stationary probability distributions.

(ii) We consider the (2,1)th and (2,2)th order rectangular transition probability tensors theoretically and numerically in this paper. We can further study and extend to the other types of rectangular transition probability tensors. For instance, (p, q)th order rectangular transition probability tensors can be studied. In this setting, we employ the pth order links among the objects and qth order links among the relations. Such higher-order links have been studied in higher-order Markov chains, see for instance [5]. Based on the proposed framework, we expect to solve the following set of tensor equations:

$$\mathcal{O}\mathbf{x}^{p-1}\mathbf{y}^{q} = \mathbf{x} \quad \text{and} \quad \mathcal{R}\mathbf{x}^{p}\mathbf{y}^{q-1} = \mathbf{y}.$$

In the MultiRank algorithm, we compute $\mathbf{x}_t = \mathcal{O}\mathbf{x}_{t-1}^{p-1}\mathbf{y}_{t-1}^{q}$ and $\mathbf{y}_t = \mathcal{R}\mathbf{x}_t^{p}\mathbf{y}_{t-1}^{q-1}$ in Steps 2 and 3 respectively. By using a similar argument, we expect to show existence and uniqueness of $\mathbf{x}$ and $\mathbf{y}$ in these higher-order links among the objects and relations.

(iii) On the other hand, we can study $(p_1, p_2, \cdots, p_s)$th order rectangular tensors where there are s different kinds of objects/relations to be analyzed. For instance, there are s networks and each network are related to each other based on their objects and relations. In the literature, there are many multiple networks applications and analysis, see for instance [17, 25, 26]. In this setting, we expect to set up a set of tensor equations similar to (7) and (8), and solve it for MultiRank values of objects/relations in these multiple networks.

(iv) We give a generalization of PageRank in this paper. We can study and extend HITS algorithm for computing the scores of hub and authority in multi-relational data. By using the graph structure of hub and authority, we solve a new set of tensor equations to obtain such scores:

$$\mathcal{H}\mathbf{y}\mathbf{z} = \mathbf{x}, \quad \mathcal{A}\mathbf{x}\mathbf{z} = \mathbf{y}, \quad \mathcal{R}\mathbf{x}\mathbf{y} = \mathbf{z},$$

Here three transition probability tensors $\mathcal{H}$, $\mathcal{A}$ and $\mathcal{R}$ corresponds to hubs, authorities and relations, $\mathbf{x}$, $\mathbf{y}$ are the authority and hub scores of objects, and $\mathbf{z}$ are relevance scores of relations. The detailed work can be found in [16].

(v) Theoretically, we need the assumption in Theorem 3 in order to show that $\bar{\mathbf{x}}$ and $\bar{\mathbf{y}}$ are unique. We have tested many different initial probability distributions in the examples in Section 6, the same resulting probability vector is obtained. We conjecture that the assumption in the theorem may be

true for irreducible transition probability tensors with some additional properties. We do not have a proof yet, so we leave it as an open problem for future analysis.

Acknowledgments: M. Ng's research supported in part by Centre for Mathematical Imaging and Vision, HKRGC grant and HKBU FRG grant. Y. Ye's research supported in part by NSFC under Grant no.61073195, and Shenzhen Science and Technology Program under Grant no. CXB201005250024A.

8. REFERENCES

[1] E. Acar, S. Camtepe, M. Krishnamoorthy, and B. Yener. Modeling and multiway analysis of chatroom tensors. *Intelligence and Security Informatics*, pages 256–268, 2005.

[2] L. Adamic and E. Adar. How to search a social network. *Social Networks*, 27(3):187–203, 2005.

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

[4] J. Carroll and J. Chang. Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart-Young" decomposition. *Psychometrika*, 35(3):283–319, 1970.

[5] W. Ching, M. Ng, and W. Ching. Markov chains: models, algorithms and applications. International Series on Operations Research and Management Science, 2006.

[6] P. Comon, X. Luciani and A. de Almeida. Tensor decompositions, Alternating least squares and other tales. *Journal of Chemometrics*, 23:393–405, 2009.

[7] H. Deng, M. Lyu, and I. King. A generalized Co-HITS algorithm and its application to bipartite graphs. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 239–248. ACM, 2009.

[8] C. Ding, X. He, P. Husbands, H. Zha, and H. Simon. PageRank, HITS and a unified framework for link analysis. In *Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 353–354. ACM, 2002.

[9] L. Freeman. A set of measures of centrality based on betweenness. *Sociometry*, 40(1):35–41, 1977.

[10] R. Harshman. Foundations of the PARAFAC procedure: Models and conditions for an" explanatory" multi-modal factor analysis. *UCLA working papers in phonetics*, 16(1):84, 1970.

[11] R. Kellogg, Uniqueness in the Schauder fixed point theorem. *Proceedings of the American Mathematical Society*, 60: 207–210, 1976.

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

[13] T. Kolda and B. Bader. The TOPHITS model for higher-order web link analysis. In *Workshop on Link Analysis, Counterterrorism and Security*, volume 7, pages 26–29, 2006.

[14] T. Kolda, B. Bader, and J. Kenny. Higher-order web link analysis using multilinear algebra. In *Data Mining, Fifth IEEE International Conference on*, page 8. IEEE, 2005.

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

[16] X. Li, M. Ng and Y. Ye. HAR: hub and authority scores in multi-relational data for query search. preprint, March 2011, http://www.math.hkbu.edu.hk/∼mng/tensor-research/har.pdf

[17] Y. Lin, J. Sun, P. Castro, R. Konuru, H. Sundaram, and A. Kelliher. MetaFac: community discovery via relational hypergraph factorization. In *Proceedings of the 15th ACM SIGKDD*, pages 527–536. ACM, 2009.

[18] Z. Nie, Y. Zhang, J. Wen and W. Ma. Object-level ranking: bringing order to web objects, WWW 2005.

[19] L. Page, S. Brin, R. Motwani, and T. Winograd. The pagerank citation ranking: Bringing order to the web. 1998.

[20] S. Ross. *Introduction to probability models*. Academic Pr, 2007.

[21] J. Sun, D. Tao, and C. Faloutsos. Beyond streams and graphs: dynamic tensor analysis. In *Proceedings of the 12th ACM SIGKDD*, pages 374–383. ACM, 2006.

[22] J. Sun, D. Tao, S. Papadimitriou, P. Yu, and C. Faloutsos. Incremental tensor analysis: Theory and applications. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 2(3):1–37, 2008.

[23] J. Sun, H. Zeng, H. Liu, Y. Lu, and Z. Chen. CubeSVD: a novel approach to personalized Web search. In *Proceedings of the 14th WWW*, pages 382–390. ACM, 2005.

[24] L. Tucker. Some mathematical notes on three-mode factor analysis. *Psychometrika*, 31(3):279–311, 1966.

[25] W. Xi, B. Zhang, Z. Chen, Y. Lu, S. Yan, W. Ma, and E. Fox. Link fusion: a unified link analysis framework for multi-type interrelated data objects. In *Proceedings of the 13th international conference on World Wide Web*, pages 319–327. ACM, 2004.

[26] D. Zhou, S. Orshanskiy, H. Zha, and C. Giles. Co-ranking authors and documents in a heterogeneous network. In *Data Mining, 2007. ICDM 2007. Seventh IEEE International Conference on*, pages 739–744. IEEE, 2008.