

A Three-Layered Mutually Reinforced Model for Personalized Citation Recommendation

Xiaoyan Cai, Junwei Han¹, Wenjie Li, Renxian Zhang, Shirui Pan, and Libin Yang

Abstract—Fast-growing scientific papers pose the problem of rapidly and accurately finding a list of reference papers for a given manuscript. Citation recommendation is an indispensable technique to overcome this obstacle. In this paper, we propose a citation recommendation approach via mutual reinforcement on a three-layered graph, in which each paper, author or venue is represented as a vertex in the paper layer, author layer, and venue layer, respectively. For personalized recommendation, we initiate the random walk separately for each query researcher. However, this has a high computational complexity due to the large graph size. To solve this problem, we apply a three-layered interactive clustering approach to cluster related vertices in the graph. Personalized citation recommendations are then made on the subgraph, generated by the clusters associated with each researcher’s needs. When evaluated on the ACL anthology network, DBLP, and CiteSeer ML data sets, the performance of our proposed model-based citation recommendation approach is comparable with that of other state-of-the-art citation recommendation approaches. The results also demonstrate that the personalized recommendation approach is more effective than the nonpersonalized recommendation approach.

Index Terms—Mutually reinforced model, personalized citation recommendation, three-layered interactive clustering.

I. INTRODUCTION

WITH the rapid growth of scientific literature, it is impossible for researchers to go through and digest all the available literature. Traditional approaches perform keyword-based searches to retrieve a list of relevant papers, and require the researchers to manually review them and thus select appropriate papers as reference papers. This is time consuming and especially challenging for young and novice researchers. Thus, new techniques for the efficient processing of scientific papers are in great demand. Citation

recommendation, which recommends a list of reference papers that are relevant to the researchers’ information needs, is an essential technique to solve this problem.

There are a variety of forms for citation recommendation. For example, Elsevier, PubMed, and SpringLink offer paper recommendations that can meet the personalized interests of researchers through RSS subscription or by e-mail alerts. But these recommendations require researchers to state their interests explicitly, or to provide information about their categories of interest. Others assume researchers have provided citation contexts, a small part of a reference list, or a full-length manuscript as search queries. We argue that the above assumptions are impractical. In this paper, we focus on a more realistic personalized citation recommendation task. That is, given the identity of a researcher and the query text the researcher describes, a list of reference papers that are relevant to the query text, as well as the background knowledge of the researcher is recommended [1].

There exist three strategies for citation recommendation: graph-based approaches [1], content-based filtering (CBF) [2], and collaborative filtering (CF) [3]. CBF recommends a reference paper based on words or topic features of a paper and the identity of a researcher, CF makes citation recommendations by finding correlations among other researchers with similar research interests. Recently, most work has applied graph-based methods in citation recommendation. Graph-based citation recommendation approaches often consider citation recommendation as a link prediction problem and solve the problem using properties of random walks. A random walk considers relationships between researchers and papers globally, while CBF and CF focus on local pairwise similarities.

In this paper, we propose to model and formulate citation recommendation in a three-layered graph, where papers, venues, and researchers are incorporated in a mutually reinforcing manner. For personalized citation recommendation, we need to integrate the identity of the researcher and the query text information into a random walk process on the constructed three-layered graph in the training process, which has high computational complexity due to the large graph size. To solve the problem, we apply three-layered interactive clustering to reduce the graph size. Personalized citation recommendations are then made on the subgraph, generated by the clusters associated with a researcher’s information needs. Although recommending relevant reference papers is the primary goal of citation recommendation, eventually both venues and researchers could also be recommended, and

Manuscript received May 13, 2017; revised November 4, 2017, January 21, 2018, and February 12, 2018; accepted March 5, 2018. This work was supported in part by the National Natural Science Foundation of China under Grant 61772429 and Grant 61272291, in part by the Research Grants Council of Hong Kong under Grant PolyU 152094/14E, and in part by the China Postdoctoral Science Foundation under Grant 2017M613205 and Grant 2017M623241. (Corresponding author: Junwei Han.)

X. Cai, J. Han, and L. Yang are with the School of Automation, Northwestern Polytechnical University, Xi’an 710072, China (e-mail: xiaoyanc@nwpu.edu.cn; jhan@nwpu.edu.cn; libiny@nwpu.edu.cn).

W. Li is with the Department of Computing, The Hong Kong Polytechnic University, Hong Kong (e-mail: cswjli@comp.polyu.edu.hk).

R. Zhang is with Yappn Canada Corporation, Markham, ON LSR 0G4, Canada (e-mail: rx.zhang@hotmail.com).

S. Pan is with the Center for Artificial Intelligence, University of Technology Sydney, Sydney, NSW 2007, Australia (e-mail: shirui.pan@uts.edu.au).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TNNLS.2018.2817245

these by-products could serve other text processing purposes. The three main contributions of this paper are as follows.

- 1) A model is developed to integrate three different objects (i.e., papers, researchers, and venues) into a three-layered graph, on which citation recommendation is performed via a mutually reinforcing manner.
- 2) The personalized citation recommendation algorithm based on the model is proposed, by integrating the identity of the researcher and the query text information into the random walk process on the graph.
- 3) Three-layered interactive clustering is conducted on the graph to reduce the size of the graph and avoid computational complexity.

The rest of this paper is organized as follows. In Section II, we review some related work. We propose a three-layered mutually reinforced model-based personalized citation recommendation approach in Section III. In Section IV, we describe the three-layered interactive clustering approach, designed to improve the efficiency of the proposed personalized citation recommendation approach. We evaluate the experimental results in Section V. The conclusion is presented in Section VI.

II. RELATED WORK

A. Citation Recommendation

Searching relevant papers to cite is tedious work for researchers; however, citation recommendation can overcome this obstacle. Citation recommendation approaches can be classified into three categories: CF approaches, CBF approaches, and graph-based approaches.

McNee *et al.* [4] applied four CF approaches to recommend citations for research papers. Yang *et al.* [5] developed a ranking-oriented CF approach, based on users' access logs to recommend papers. Kang *et al.* [6] filled missing elements of the rating matrix using a low-rank assumption and then made recommendations based on the recovered matrix. Chandrasekaran *et al.* [7] created user profiles based on the user's authorship of previously published papers. Relevant papers are recommended by computing the similarities between the profiles of the papers in the collection and the user profile.

CF works by collecting user feedback on items and recommends an item based on the similarities displayed between the profiles of users. Therefore, it is "domain independent" [8]. Although it has been successfully applied in various domains, it suffers from both the sparsity and the first-rater problems [9]. CBF has many merits, such as the ability to generate recommendations over all items in the domain. Nascimento *et al.* [2] considered the content of a full paper as input to construct queries, and then applied the CBF-based recommendation algorithm to recommend candidate papers which were relevant to the input paper. CBF also has its shortcomings [10], such as content limitation in domains, narrow content analysis, and analysis of quality and taste. Therefore, several hybrid approaches have been proposed to combine the advantages of both CF and CBF methods.

Torres *et al.* [8] developed hybrid approaches combining CF and CBF to recommend research papers. CF used the KNN

algorithm to output an ordered list of input citations as a recommendation, while CBF recommended papers by computing cosine similarity between the current paper and the papers in the collection. Hybrid recommendation approaches were then generated by combining CF and CBF. Torres *et al.* [8] found that the hybrid approaches performed better than the individual ones. McNee [11] generated paper recommendations by a hybrid approach combining CF and CBF techniques.

Recent research has employed graph-based approaches to study the citation recommendation problem [1], [12]–[18]. Strohman *et al.* [12] treated citation recommendation as a link prediction problem. They represented each paper as a vertex, the citation relationship as the link between vertices and a new paper as a vertex without any in-link and out-link. Zhou *et al.* [13] measured paper similarities by combining the author-paper graph, the paper-citation graph, and the venue-paper graph. Then they recommended citations by treating some known citations as positive labels and applying semisupervised learning on the combined graphs. Gori and Pucci [14] developed a random walk-based approach to recommend research papers. Meng *et al.* [1] presented a personalized citation recommendation approach, which incorporated different kinds of information, such as content of papers, authorship, and citation into a unified graph model. Pan *et al.* [15] proposed an academic paper recommendation approach based on a heterogeneous graph containing various kinds of features. Jiang *et al.* [16] proposed a chronological citation recommendation approach that considers chronological nature. Wang *et al.* [17] developed a novel entity class-dependent discriminative mixture model for cumulative citation recommendation, avoiding the insufficient training data of less popular entities in a chronological stream corpus. Chakraborty *et al.* [18] presented a diversified citation recommendation framework that balanced the prestige, relevance, and diversity of reference papers, based on the given scientific query manuscript.

The above approaches are global citation recommendations, which aim at recommending a reference list based on a given manuscript. Local citation recommendations, on the other hand, aim at recommending citations for the specific context wherein a citation should be made. He *et al.* [19] built a CiteSeerX system, which not only recommends a reference list to a given manuscript, but also provides a citation list to a specific citation placeholder. Ebesu and Fang [20] developed an encoder–decoder framework for local citation recommendation. Huang *et al.* [21] proposed a novel neural probabilistic model for context-based citation recommendation. The proposed model can simultaneously learn the distributed representations of cited papers and citation contexts. We focus on the global citation recommendation in this paper.

B. Mutual Reinforcement

Mutual reinforcement [22]–[24] has been used in summarization research. Zha [25] developed a mutual reinforcement principle to simultaneously extract significant sentences and key phrases. He first constructed a weighted bipartite document graph by linking together the terms and the sentences

containing the terms. Then he applied mutual reinforcement principle to compute the largest singular vectors of the transition matrix of the constructed bipartite document graph. Wei *et al.* [26] further applied the mutual reinforcement principle to the term, sentence, and document mutual reinforcement chain, and proposed a reinforced ranking approach for query-oriented multidocument summarization. Cai and Li [27] proposed a mutually reinforced manifold-ranking-based relevance propagation model for query-focused multidocument summarization. In this paper, we integrate both interrelationships (author-to-paper, venue-to-paper, and author-to-venue relationships) and intrarelations (author-to-author, venue-to-venue, and paper-to-paper relationships) into a three-layered graph, and develop citation recommendation via a mutually reinforcing manner on the graph.

C. Multilayered Clustering

Recently, much attention has been drawn to multilayered clustering, in which more than two types of object are simultaneously clustered. Similar to pairwise co-clustering between two kinds of objects, clustering on multiple types of objects can be considered as high-order co-clustering [28]–[33]. Some existing methods iteratively applied a two-layered clustering algorithm to solve high-order clustering problem [34]. However, it is not a real solution. Bekkerman *et al.* [35] proposed a distributed clustering approach to simultaneously cluster different kinds of objects. Cheng *et al.* [36] simultaneously clustered terms, documents, and authors using nonnegative matrix factorization, but they only used interrelationships. Multilayered classification is different from multilayered clustering [37]. Cai and Li [38] proposed an integrated clustering framework and an interactive clustering framework to co-cluster different types of text objects. The integrated clustering framework utilized interrelationships among different types of text objects, while the interactive clustering framework utilized both interrelationships and intrarelations among different types of text objects. Here, we adopt the interactive clustering framework.

III. PERSONALIZED CITATION RECOMMENDATION

A. Three-Layered Graph Construction

Let us formulate the three-layered graph containing papers, authors and venues as $G = \langle P, A, V, E_{pp}, E_{aa}, E_{vv}, E_{ap}, E_{av}, E_{pv} \rangle$, where $P = \{p_i\}$ ($1 \leq i \leq n_p$, n_p is the total number of papers), $A = \{a_j\}$ ($1 \leq j \leq n_a$, n_a is the total number of authors), and $V = \{v_l\}$ ($1 \leq l \leq n_v$, n_v is the total number of venues). $E_{pp} = \{e_{ij}|p_i, p_j \in P\}$, $E_{aa} = \{e_{ij}|a_i, a_j \in A\}$, $E_{vv} = \{e_{ij}|v_i, v_j \in V\}$, $E_{ap} = \{e_{ij}|a_i \in A, p_j \in P\}$, $E_{av} = \{e_{ij}|a_i \in A, v_j \in V\}$ and $E_{pv} = \{e_{ij}|p_i \in P, v_j \in V\}$ correspond to the edges between papers, the edges between authors, the edges between venues, the edges between authors and papers, the edges between authors and venues, and the edges between papers and venues, respectively. Let $W_{pp} = [w_{p_i p_j}]_{n_p \times n_p}$, $W_{aa} = [w_{a_i a_j}]_{n_a \times n_a}$, $W_{vv} = [w_{v_i v_j}]_{n_v \times n_v}$, $W_{ap} = [w_{a_i p_j}]_{n_a \times n_p}$, $W_{av} = [w_{a_i v_j}]_{n_a \times n_v}$, and $W_{pv} = [w_{p_i v_j}]_{n_p \times n_v}$ be the paper-to-paper, author-to-author, venue-to-venue, author-to-paper, author-to-venue, and paper-to-venue affinity matrices.

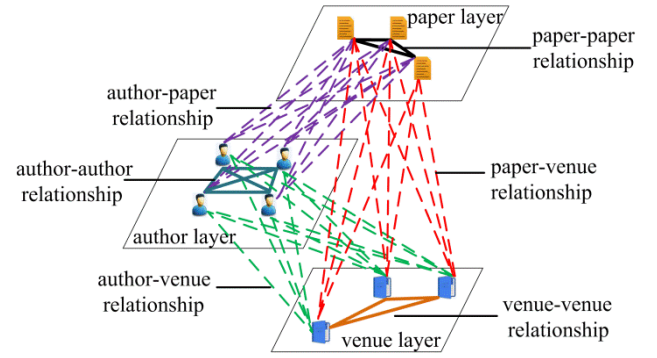


Fig. 1. Three-layered graph.

paper-to-venue affinity matrices. $w_{p_i p_j}$ is defined as the cosine similarity between paper p_i and paper p_j (we apply cosine similarity, Pearson correlation-based similarity, and Jaccard coefficient-based similarity measures [39] in the proposed algorithms, as we have found that the performance of the proposed algorithms, based on the three different similarity measures, is similar. For ease of illustration, we use cosine similarity in this paper); $w_{a_i a_j}$ is defined as 1 if author a_i and author a_j collaboratively write a paper, otherwise 0; $w_{v_i v_j}$ is defined as 1 if venue v_i and venue v_j have published papers written by the same author, otherwise 0; $w_{a_i p_j}$ is defined as 1 if author a_i is an author of paper p_j , otherwise 0; $w_{a_i v_j}$ is defined as 1 if a paper written by author a_i is published in venue v_j , otherwise 0; and $w_{p_i v_j}$ is defined as 1 if paper p_i is published in venue v_j , otherwise 0. E_{aa} , E_{pp} , and E_{vv} indicate the intrarelations in the graph G , while E_{ap} , E_{av} , and E_{pv} represent the interrelationships in graph G . Fig. 1 illustrates the constructed graph.

B. Three-Layered Mutually Reinforced Model for Personalized Citation Recommendation

In this paper, query is formulated by query author and query text, i.e., $Q = [q_a, q_t]$. q_a represents the identity of a researcher, whilst query text q_t can be the title of a paper, the abstract of a paper, or even the full manuscript of a paper. Inspired by Wei *et al.* [26], we propose a mutual reinforcement model, taking advantage of relations among papers, authors, and venues for citation recommendation. The mutual reinforcement model is built with three PageRank-like models for paper, author, and venue respectively, but in a unified and interrelated way. The ranking of each of them is derived not only from the relationship within itself, but is also affected by the other two objects. For personalized citation recommendation, we need to incorporate the information of query text and query author into the model. We then apply the following mutual reinforcement principle to personalized citation recommendation, i.e.,

- 1) A paper is recommended if a) it is relevant to other recommended papers and is relevant to the given query text, b) it is written by recommended authors, and c) it is published in a recommended venue.
- 2) An author is recommended if a) he/she writes recommended papers, b) he/she collaborates with other

recommended authors and collaborates with the given query author, and c) he/she writes papers published in recommended venues.

- 3) A venue is recommended if a) it publishes recommended papers, b) it publishes papers written by recommended authors, and c) it is relevant to other recommended venues.

Then, the ranking of papers, authors, and venues can be iteratively derived from the above principle. Let R_P , R_A , and R_V denote the ranking scores of P , A , V , respectively, the corresponding mathematical description of the above mutual reinforcement principle is presented in (1), where $\text{rel}(p_i|q_t)$ is the cosine similarity between paper p_i and query text q_t , $\text{rel}(a_i|q_a)$ is the collaborative relationship between the author a_i and the query author q_a (i.e., if a_i and q_a have written a paper collaboratively, $\text{rel}(a_i|q_a)$ is set to 1, otherwise 0). d is a damping factor. The definition of (1) is inspired by the definition of the relevance propagation model in [30]

$$\left[\begin{array}{l} R^{(t+1)}(p_i) = \alpha_1 \cdot \left[d \cdot \frac{\text{rel}(p_i|q_t)}{\sum_{p_i \in V_p} \text{rel}(p_i|q_t)} \right. \\ \quad \left. + (1-d) \cdot \sum_{p_j \in V_p} \frac{w_{p_i p_j}}{\sum_{p_l \in V_p} w_{p_l p_l}} R^{(t)}(p_j) \right] \\ \quad + \beta_1 \cdot \sum_{a_i \rightarrow p_i} R^{(t)}(a_i) + \gamma_1 \cdot R^{(t)}(v_i) \\ R^{(t+1)}(a_i) = \alpha_2 \cdot \sum_{a_i \rightarrow p_i} R^{(t)}(p_i) + \beta_2 \\ \quad \cdot \left[d \cdot \frac{\text{rel}(a_i|q_a)}{\sum_{a_i \in V_a} \text{rel}(a_i|q_a)} + (1-d) \right. \\ \quad \left. \cdot \sum_{a_j \in V_a} \frac{w_{a_i a_j}}{\sum_{a_l \in V_a} w_{a_l a_l}} R^{(t)}(a_j) \right] + \gamma_2 \cdot R^{(t)}(v_i) \\ R^{(t+1)}(v_i) = \alpha_3 \cdot \sum_{v_i \rightarrow p_i} R^{(t)}(p_i) + \beta_3 \\ \quad \cdot \sum_{a_i \rightarrow v_i} R^{(t)}(a_i) + \gamma_3 \cdot \sum_{v_j \in V_v} R^{(t)}(v_j). \end{array} \right] \quad (1)$$

Equation (1) can be reformulated in the following matrix-vector format:

$$\left[\begin{array}{l} R_P^{(t+1)} = \alpha_1 \cdot M_{pp} \cdot R_P^{(t)} + \beta_1 \cdot W_{ap}^T \cdot R_A^{(t)} + \gamma_1 \cdot W_{pv} \cdot R_V^{(t)} \\ R_A^{(t+1)} = \alpha_2 \cdot W_{ap} \cdot R_P^{(t)} + \beta_2 \cdot M_{aa} \cdot R_A^{(t)} + \gamma_2 \cdot W_{av} \cdot R_V^{(t)} \\ R_V^{(t+1)} = \alpha_3 \cdot W_{pv}^T \cdot R_P^{(t)} + \beta_3 \cdot W_{av}^T \cdot R_A^{(t)} + \gamma_3 \cdot W_{vv} \cdot R_V^{(t)} \end{array} \right] \quad (2)$$

where $M_{pp} = [dB + (1-d)C]^T$ and $M_{aa} = [dG + (1-d)H]^T$. B and G are the square matrices such that for a given index i , all the elements in the i th column are proportional to $\text{rel}(p_i|q_t)$ and $\text{rel}(a_i|q_a)$, respectively. C and H are also square matrices such that each entry $C(i, j)$ and $H(i, j)$ is proportional to $w_{p_i p_j}$ and $w_{a_i a_j}$, respectively

$$W = \begin{bmatrix} \alpha_1 & \beta_1 & \gamma_1 \\ \alpha_2 & \beta_2 & \gamma_2 \\ \alpha_3 & \beta_3 & \gamma_3 \end{bmatrix}$$

TABLE I
QR ALGORITHM

Input: The paper set P , the author set A , the venue set V , the query author q_a , the query text q_t , damping factor d , weight matrix M and $\tau = 0.0001$.

Output: Paper recommendation list.

- 1: Construct the matrices $W_{pp}, W_{ap}, W_{pv}, W_{aa}, W_{av}, W_{vv}$;
- 2: Make the matrix M stochastic and irreducible by transforming the nine block matrices, which is mentioned in [26];
- 3: Randomly initiated non-negative vectors $R_P^{(0)}, R_A^{(0)}$ and $R_V^{(0)}$, such that $\|R_P^{(0)}\| = 1, \|R_A^{(0)}\| = 1$ and $\|R_V^{(0)}\| = 1$;
- 4: $i \leftarrow 0, \nabla \leftarrow 0$;
- 5: Repeat
 - 6: $R_P^{(t+1)} = \alpha_1 \cdot M_{pp} \cdot R_P^{(t)} + \beta_1 \cdot W_{ap}^T \cdot R_A^{(t)} + \gamma_1 \cdot W_{pv} \cdot R_V^{(t)}$
 - 7: $R_A^{(t+1)} = \alpha_2 \cdot W_{ap} \cdot R_P^{(t)} + \beta_2 \cdot M_{aa} \cdot R_A^{(t)} + \gamma_2 \cdot W_{av} \cdot R_V^{(t)}$
 - 8: $R_V^{(t+1)} = \alpha_3 \cdot W_{pv}^T \cdot R_P^{(t)} + \beta_3 \cdot W_{av}^T \cdot R_A^{(t)} + \gamma_3 \cdot W_{vv} \cdot R_V^{(t)}$
- 9: $\nabla \leftarrow \max \left\{ \begin{array}{l} \|R_P^{(t+1)} - R_P^{(t)}\| \\ \|R_A^{(t+1)} - R_A^{(t)}\| \\ \|R_V^{(t+1)} - R_V^{(t)}\| \end{array} \right\}$;
- 10: $t \leftarrow t + 1$;
- 11: Until $\nabla < \tau$;
- 12: $R_P \leftarrow R_P^{(i)}, R_A \leftarrow R_A^{(i)}, R_V \leftarrow R_V^{(i)}$;
- 13: Generating paper recommendation list according to the top-ranked values of R_P ;
- 14: Return.

is a weight matrix which balances the relative weights of papers, authors, and venues. Equation (2) corresponds to a block matrix

$$M = \begin{bmatrix} \alpha_1 M_{pp} & \beta_1 W_{ap}^T & \gamma_1 W_{pv} \\ \alpha_2 W_{ap} & \beta_2 M_{aa} & \gamma_2 W_{av} \\ \alpha_3 W_{pv}^T & \beta_3 W_{av}^T & \gamma_3 W_{vv} \end{bmatrix}. \quad (3)$$

Let $R = [R_P, R_A, R_V]^T$, then R can be computed as the dominant eigenvector of M

$$M \cdot R = \lambda \cdot R. \quad (4)$$

Wei *et al.* [26] have proved that the three-layered mutual reinforcement model will converge on R . Finally, a ranked paper list is formed by ranking papers in P according to the scores assigned by the model. Top-ranked papers are recommended to the query author. We call the above approach query recommendation (QR) in this paper. Table I summarizes the whole process of the QR algorithm.

QR computes the ranking score of each vertex in the whole graph. But when the graph has massive vertices, the QR approach is computationally expensive. An alternative way to handle this problem is to perform offline computation, which will prevent researchers from renewing their identity and obtaining a real-time updated recommendation result. In Section IV, we will present a more efficient approach which applies three-layered interactive clustering to cluster-related vertices in the graph and then applies the above proposed

QR approach to a much smaller subgraph of the three-layered graph.

IV. IMPROVING EFFICIENCY VIA THREE-LAYERED INTERACTIVE CLUSTERING

When we apply the QR algorithm on the three-layered graph G , we find that the highly ranked vertices are always in the neighborhood of the vertices q_t and q_a . For example, the papers which are highly ranked are prone to cluster tightly around papers that are relevant to q_t , and the authors which are highly ranked are prone to cluster tightly around q_a . Thus, we develop a three-layered clustering approach to solve the curse of the dimensionality of the QR algorithm, in order to improve the efficiency of the algorithm.

A. Three-Layered Interactive Clustering Framework

As the specific author information and the venue information can enrich the paper information, we think that using the full author information, venue information and paper information can further boost clustering performance. So, we integrate papers, authors, and venues into a three-layered interactive framework to improve the clustering performance.

The clustering process is performed on the constructed three-layered graph G . We begin from the paper layer to obtain k_1 paper clusters $C_P^{(1)} = \{C_{P_1}^{(1)}, C_{P_2}^{(1)}, \dots, C_{P_{k_1}}^{(1)}\}$ based on W_{pp} , which is denoted as $W_{pp}^{(1)}$. After that, we calculate the association between each author and each paper cluster, and construct an author-to-paper-cluster affinity matrix $W_{aC_P}^{(1)}$. $W_{aC_P}^{(1)}(i, j)$ indicates the papers' number that author a_i writes in paper cluster $C_{P_j}^{(1)}$. Thus the refined similarity of a_i and a_j is defined as follows:

$$W_{aa}^{(1)}(i, j) = \varepsilon W_{aa}(i, j) + (1 - \varepsilon) W_{aC_P}^{(1)}(i, :) \times W_{aC_P}^{(1)T}(:, j) / (F^{(1)}(i) \cdot F^{(1)}(j)) \quad (5)$$

in which the row vector $W_{aC_P}^{(1)}(i, :)$ indicates the number of papers between author a_i and each paper cluster. $W_{aC_P}^{(1)T}$ is the transposed matrix of $W_{aC_P}^{(1)}$. The i th element of the column vector $F^{(1)}$ is defined as $F^{(1)}(i) = (\sum_{h=1}^{k_1} (W_{aC_P}^{(1)}(i, h))^2)^{1/2}$. Based on the author-to-author similarity which is defined in (5), author clusters $C_A^{(1)} = \{C_{A_1}^{(1)}, C_{A_2}^{(1)}, \dots, C_{A_{k_2}}^{(1)}\}$ can be obtained by applying one of the classical clustering algorithms. After that the venue-to-author-cluster matrix $W_{vC_A}^{(1)}$ is constructed, and $W_{vC_A}^{(1)}(i, j)$ indicates the number of authors in $C_{A_j}^{(1)}$ who write papers published in venue v_i . We can also get the venue cluster based on the venue-to-author-cluster similarities and the original venue-to-venue similarities. We set the weight parameter of the original similarities in the same way as in (5).

In the second round, as the author information and venue information can improve the performance of paper clustering, we define the newly updated similarity between two papers as the weighted combination of similarities between venue clusters and papers, similarities between author clusters and papers, and original similarities between two papers, in which

the venue clusters and author clusters are the newly updated clusters, i.e.,

$$W_{pp}^{(2)}(i, j) = \delta W_{pp}(i, j) + \mu W_{pC_V}^{(1)}(i, :) \times W_{pC_V}^{(1)T}(:, j) / (H^{(1)}(i) \cdot H^{(1)}(j)) + \eta W_{pC_A}^{(1)}(i, :) \times W_{pC_A}^{(1)T}(:, j) / (L^{(1)}(i) \cdot L^{(1)}(j)) \quad (6)$$

where $W_{pC_V}^{(1)}(i, :)$ is a row vector, the i th element of the column vector $H^{(1)}$ is defined as $H^{(1)}(i) = (\sum_{h=1}^{k_2} (W_{PC_V}(i, h))^2)^{1/2}$. $W_{pC_A}^{(1)}$ and $L^{(1)}$ are understood in the same way. Similarly, the similarity matrix $W_{aa}^{(2)}$ is obtained from W_{aa} , $W_{pp}^{(2)}$, and $W_{vv}^{(1)}$; and the similarity matrix $W_{vv}^{(2)}$ is calculated with W_{vv} , $W_{pp}^{(2)}$, and $W_{aa}^{(2)}$, which is similar to (6).

The second round is repeated until paper clusters, author clusters, and venue clusters are stable or above a threshold. ε, δ, μ and η are weighting parameters ranging from 0 to 1 and $\delta + \mu + \eta = 1$.

B. Personalized Query-Oriented Reference Paper Recommendation on Clustered Three-Layered Graph

1) *Clustering Algorithm*: We aim to simultaneously cluster papers, authors, and venues, so we need to apply a basic clustering algorithm based on the paper similarity, author similarity, and venue similarity, which have been described in Section IV-B. We apply the similarity and clustering with a single Kernel (SCSK) algorithm [40] in our work. A linear kernel $K(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T \mathbf{y}$ is used in the SCSK algorithm, where α is set to 0.001 and β is set to 1e-6, as suggested in [40]. To avoid exhaustively searching for an appropriate cluster number for each set, we apply the spectra approach introduced by Li *et al.* [41] to predict the number of the expected clusters. As for paper cluster number k_1 , we calculate its eigenvalues λ_i ($i = 1, 2, \dots, n_p$) based on the original paper similarity matrix W_{pp} , using the normalized 1-norm, and define the ratio $\sigma_i = \lambda_{i+1} / \lambda_2$ ($i \geq 1$). If $\sigma_i - \sigma_{i+1} > 0.05$ and σ_1 are still close to 1, we set paper cluster number k_1 as $k_1 = i + 1$. The author cluster number k_2 and venue cluster number k_3 can be set similarly.

2) *Reference Paper Recommendation*: The obtained clusters not can only provide a way of clustering related vertices, but can also help reduce the computational complexity of performing a personalized citation recommendation on a large three-layered graph. For example, given a query text and query author, we can identify a relevant paper cluster P^{sub} by calculating the similarity between each paper cluster and the query text. Similarly, we can obtain the corresponding author cluster A^{sub} by matching the query author. Finally, we can obtain the related venue cluster V^{sub} via the author-to-venue-cluster relation; that is, based on the final matrix W_{aC_V} , we find a row vector $W_{aC_V}(i, :)$ in which the row indicates the query author, then we select the largest value in the row and determine its corresponding column to be V^{sub} .

After the submatrices P^{sub} , A^{sub} , and V^{sub} are obtained, a subgraph of the three-layered graph is constructed. As the size of the subgraph is considerably smaller than that of the

TABLE II
CQR ALGORITHM

<p>Input: The paper set P, the author set A, the venue set V, the query author q_a, the query text q_t, damping factor d, weight matrix M, parameters $\varepsilon, \delta, \mu, \eta$ and $\tau = 0.0001$</p> <p>Output: Paper recommendation list.</p>
<ol style="list-style-type: none"> 1: Construct the matrices W_{pp}, W_{aa}, and W_{vv}, respectively; 2: Calculate the paper cluster number k_1, author cluster number k_2 and venue cluster number k_3 based on W_{pp}, W_{aa}, and W_{vv}, respectively; 3: Cluster authors using Eq. (5); 4: Cluster venues using equation similar to Eq. (5); 5: Repeat 6: Cluster papers using Eq. (6); 7: Cluster authors and venues separately using equations similar to Eq. (6); 8: Until the changes of the clustering results on the three sets are stable or above a threshold; 9: Identify relevant cluster P^{sub}, A^{sub} and V^{sub} related to q_a and q_t, as mentioned in Section IV-B-2); 10: Perform QR algorithm based on P^{sub}, A^{sub} and V^{sub}; 11: Generating paper recommendation list; 12: Return

original graph, the proposed QR approach can be performed efficiently to identify the relevant papers, which is called a CQR clustered QR (CQR). Table II summarizes the whole process of the CQR algorithm.

V. EXPERIMENT AND EVALUATION

A. Experiment Setup

1) *Experimental Data:* In order to evaluate the quality of the proposed model, we conduct experiments on three bibliographic data sets.

1) *The ACL anthology network (AAN) data set*,¹ which was established by Radev and Mutuhkrishnan [42] and consists of conference papers and journal papers in computational linguistics.

We remove the papers which do not have titles or abstracts in the data set, then we use the remaining 12555 papers published from 1965 to 2013 as the experimental data set. For evaluation purposes, we divide the entire data set into two disjoint sets, the papers published before 2013 are deemed as training set (11197 papers) and the remaining papers fall into the test set (1358 papers).

2) *The DBLP data set*,² which consists of bibliography data in computer science [43]. We select a list of conferences from five research areas: information retrieval (SIGIR, ACL, EAACL, ECIR, NAACL, CIKM, EMNLP, and COLING), machine learning (NIPS, ICML, SIGKDD, WSDM, ICDE, ICDM, and PAKDD), computer vision (CVPR, ECCV, ICCV, ACCV, ICIP, ICPR, and MM), networks and communications (INFOCOM, SIGCOMM, ICC, GLOBECOM, MOBICOM,

ICDCS, SECON, and ICNP), and computer security (SP, NDSS, FC, ACSAC, ARES, and ISI). The DBLP network consist of 64332 papers, among which the papers published before 2013 (included) make the training set (56304 papers), and papers published from 2014 to 2015 make the test set (8028 papers).

3) *The CiteSeer ML data set* [44], a subset of CiteSeerX data,³ which consists of scientific publications, was selected. The CiteSeer ML data set contains 139227 papers, which have been used in [45]. Papers published before 2012 (included) are selected as the training set (120291 papers), the remaining papers are selected as the test set (18936 papers). In this paper, we extract the title and abstract of the papers in the three data sets as document content of the papers, and we define a query as the title, abstract, and author of a paper.

2) *Evaluation Methods:* We define the cluster quality measure to evaluate the performance of generated clusters. We also use a paper's reference list as ground truth to evaluate the performance of recommendation. The performance of recommendation can be measured by a wide range of metrics, including user studies and click-through monitoring. In this paper, we focus on three evaluation methods: Recall@ N , mean average precision (MAP), and mean reciprocal rank (MRR).

a) *Intrinsic evaluation:* We treat cluster quality as an intrinsic evaluation in our citation recommendation task. First, we construct a paper graph model $G_P = (P, E_{pp})$, where P and E_{pp} are defined the same as in graph G . Modularity measure Q is defined in social network as [36]

$$Q = \sum_i (e_{ii} - a_i^2) = \text{Tre} - \|\mathbf{e}\|^2 \quad (7)$$

where the matrix \mathbf{e} is a $K \times K$ symmetric matrix whose element e_{ij} is the fraction of edges that join vertices in community i to vertices in community j . (K is the number of communities in the network.) $a_i = \sum_j e_{ij}$ represents the proportion of ends of edges that are attached to vertices in community i . The trace of the matrix $\text{Tre} = \sum_i e_{ii}$ gives the proportion of edges in the network that connect vertices in the same community. $\|\mathbf{x}\|$ is the sum of the elements of the matrix X . As the sentence graph which we construct is a connected graph, the traditional modularity measure is defined in a disconnected graph, but it cannot be applied in the constructed sentence graph directly. In order to solve this problem, we define the element e_{ij} of the matrix \mathbf{e} as the proportion of all edges' weight in G_P that connect vertices in cluster C_i to vertices in cluster C_j . Then we evaluate the generated paper clusters through the revised modularity measure Q_P . Author cluster quality measure Q_A and venue cluster quality measure Q_V can be defined in a similar way.

b) *Extrinsic evaluation:* Our ultimate aim is to recommend more relevant reference papers. We use three common metrics to evaluate the extrinsic recommendation performance as follows.

¹We download 2013 Release from <http://clair.eecs.umich.edu/aan/index.php>

²<http://arnetminer.org/citation> (V4 version is used)

³<http://citeseerx.ist.psu.edu>

- 1) *Recall@N*: is defined as the percentage of original reference papers that appear in the top- N recommended reference papers. Here we use Recall@N ($N = \{20, 40, 60, 80, 100\}$) for evaluation, where N is the number of top- N papers recommended by our proposed approaches.
- 2) *Mean Average Precision (MAP)* [46]: As Recall@N only considers the top- N ranking results and not the exact ranking position, the MAP measure can overcome this disadvantage. MAP is derived from average precision. The average precision for a query is the mean of the precisions obtained after each relevant reference (here “relevant” means the ground-truth reference papers of the submitted query) paper is retrieved. The corresponding quantity averaged over queries is called “MAP.”
- 3) *Mean Reciprocal Rank (MRR)* [47]: Evaluates how far the first relevant reference papers are from the top. It is defined as

$$\text{MRR} = \frac{1}{|T_p|} \sum_{i=1}^{|T_p|} \frac{1}{r_i} \quad (8)$$

where r_i is the rank of the highest ranking relevant reference papers for the i th query, and T_p is the test paper set.

3) *Baselines*: We use CBF algorithms (cosine similarity [46], Okapi [48], and K - L Divergence [49]) and CF algorithm as baseline approaches. As for CF, we first construct a matrix $\mathbf{M}_{n_a \times n_p}$ based on the whole data set. $\mathbf{M}_{n_a \times n_p}(i, j) = 1$ indicates that the paper p_j is one of the reference papers in a paper written by the author a_i , meanwhile p_j is in the training set, otherwise $\mathbf{M}_{n_a \times n_p}(i, j) = 0$. Then we apply the standard CF approach based on the matrix in order to compare against our proposed approaches.

B. Experimental Results

In order to obtain semantic similarity between texts, we first apply the doc2vec approach [50] to obtain a fixed-length vector representation of each text information, such as paper content and query content. (The fixed-length of the vector is set to 100 in this paper.) Then we use the cosine similarity measure to compute similarities between the two text objects. We set the weight matrix \mathbf{M} in the QR as

$$\begin{bmatrix} 1 & 0.5 & 0.5 \\ 0.5 & 1 & 0.5 \\ 0.5 & 0.5 & 1 \end{bmatrix}$$

based on the assumption that papers play important roles, while authors and venues play less important, but equal roles. \mathbf{M} is normalized to be column stochastic. The damping factor d is set at 0.85 as default.

1) *Parameter Setting*: In the first set of experiments, we examine and fix the values of the parameters ε , δ , η , and μ in the three-layered interactive clustering framework. We conduct these experiments on the AAN data set. In the first round of the framework, we first obtain paper clusters based on \mathbf{W}_{pp} , then we tune the values of ε from 0 to 1

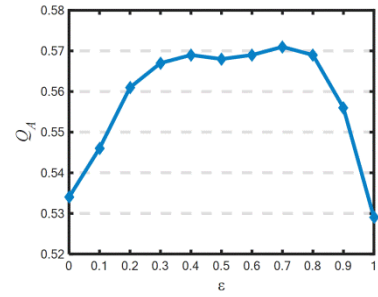


Fig. 2. Author cluster quality in the first round of the three-layered interactive clustering framework with different ε values on the AAN data set.

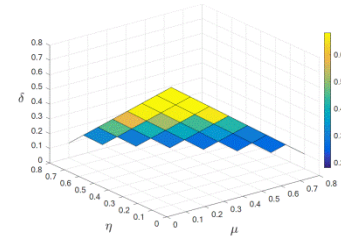


Fig. 3. Paper cluster quality in the second iteration of the three-layered interactive clustering framework with different μ , η and δ values on the AAN data set.

with step size 0.1 to obtain author clusters. The author cluster quality values are presented in Fig. 2.

We can see from Fig. 2 that when ε ranges from 0.3 to 0.8, author cluster quality in the first round of the three-layered interactive clustering framework is very stable, and the best result is obtained at $\varepsilon = 0.7$. Thus, we use $\varepsilon = 0.7$ in the following experiments.

From the second iteration of the three-layered interactive clustering framework, we first need to tune parameters δ , η , and μ to get paper clusters. We tune the three parameters with step size 0.1. Fig. 3 presents paper cluster quality values with three different δ , η , and μ values.

Fig. 3 shows that paper cluster quality reaches the best result when $\mu = 0.1$, $\eta = 0.2$, and $\delta = 0.7$. Thus, we evaluate the performance of the following experiments with the same parameter values. Moreover, we find that when fixing the δ value, the performance of a paper cluster quality value with a higher μ value and a lower η value is inferior to that with a lower μ value and a higher η value. We believe that the author information is more important for paper clustering than venue information. Parameters are set the same in DBLP and CiteSeer ML Data sets.

The number of clusters is predefined, as mentioned in Section IV-B1), so the number of paper clusters is 99, the number of author clusters is 92, and the number of venue clusters is 16 in the AAN data set. When the paper clusters, the author clusters, and the venue clusters are stable, we find that the largest paper cluster contains 136 papers, the smallest paper cluster contains 92 papers, and the average number of each paper cluster is 113. Likewise, the largest author cluster contains 128 authors, the smallest author cluster contains 85 authors, and the average number of each author

TABLE III
COMPARISON OF DIFFERENT APPROACHES ON AAN DATA SET

	MAP	MRR	Recall @20	Recall @40	Recall @60	Recall @80	Recall @100
QR	0.126	0.137	0.242	0.331	0.401	0.425	0.479
CQR	0.124	0.135	0.239	0.328	0.396	0.421	0.475
QR-PV	0.103	0.116	0.219	0.285	0.348	0.367	0.411
QR-PA	0.101	0.112	0.215	0.281	0.343	0.362	0.402
QR-P	0.098	0.107	0.206	0.273	0.339	0.351	0.396
CQR-P	0.096	0.106	0.206	0.269	0.332	0.349	0.395
cosine	0.082	0.096	0.196	0.253	0.319	0.337	0.382
CF	0.081	0.095	0.196	0.250	0.318	0.335	0.379
okapi	0.079	0.092	0.193	0.250	0.315	0.334	0.377
K-L	0.078	0.090	0.191	0.249	0.312	0.331	0.376

TABLE IV
COMPARISON OF DIFFERENT APPROACHES ON DBLP DATA SET

	MAP	MRR	Recall @20	Recall @40	Recall @60	Recall @80	Recall @100
QR	0.115	0.125	0.228	0.302	0.366	0.383	0.428
CQR	0.113	0.122	0.226	0.300	0.365	0.381	0.426
QR-PV	0.101	0.113	0.216	0.283	0.346	0.365	0.408
QR-PA	0.099	0.111	0.214	0.279	0.341	0.359	0.399
QR-P	0.097	0.106	0.205	0.271	0.337	0.349	0.395
CQR-P	0.095	0.105	0.204	0.267	0.330	0.347	0.393
cosine	0.080	0.094	0.195	0.251	0.317	0.336	0.380
CF	0.079	0.093	0.194	0.248	0.316	0.334	0.378
okapi	0.076	0.090	0.191	0.247	0.314	0.332	0.376
K-L	0.075	0.088	0.189	0.246	0.311	0.330	0.375

cluster is 107. The largest venue cluster contains 39 venues, the smallest venue cluster contains 21 venues, and the average number of each venue cluster is 28.

2) *Performance of Citation Recommendation*: We ignore the query author information in this set of experiments. So the query information only focuses on the query text, i.e., $Q = [q_t]$. QR is our full mutually reinforced approach on the three-layered graph, and CQR is its improved version, which applies the mutual reinforcement principle on a subgraph of the three-layered graph. To illustrate the benefits of applying the mutual reinforcement principle on the three-layered graph, we also compare our algorithm against the following:

- 1) QR-P, which applies query-sensitive random walk on the paper layer alone;
- 2) QR-PA, which applies the mutual reinforcement principle on the paper-author graph constructed from the paper-to-paper, author-to-author, and paper-to-author relationships;
- 3) QR-PV, which applies the mutual reinforcement principle on the paper-venue graph constructed from the paper-to-paper, venue-to-venue, and paper-to-venue relationships;
- 4) CQR-P, which first runs the clustering algorithm on the paper layer alone, then selects a relevant paper cluster related to the query text, and finally applies query-sensitive random walk on the relevant paper cluster.

Tables III–V show the comparison of the results of these approaches on the three data sets.

The results show that our proposed approaches outperform the three CBF approaches and the CF approach. It is not

TABLE V
COMPARISON OF DIFFERENT APPROACHES ON CITESEER ML DATA SET

	MAP	MRR	Recall @20	Recall @40	Recall @60	Recall @80	Recall @100
QR	0.117	0.128	0.231	0.305	0.368	0.387	0.430
CQR	0.115	0.127	0.230	0.303	0.367	0.385	0.428
QR-PV	0.102	0.115	0.218	0.284	0.347	0.366	0.410
QR-PA	0.100	0.113	0.216	0.281	0.343	0.361	0.401
QR-P	0.098	0.107	0.206	0.272	0.338	0.350	0.396
CQR-P	0.095	0.106	0.205	0.268	0.331	0.348	0.394
cosine	0.081	0.095	0.196	0.252	0.318	0.337	0.381
CF	0.080	0.094	0.195	0.249	0.317	0.335	0.379
okapi	0.078	0.091	0.192	0.248	0.316	0.334	0.378
K-L	0.076	0.089	0.190	0.247	0.312	0.332	0.376

TABLE VI
COMPARISON OF PERFORMANCE BETWEEN PERSONALIZED AND NONPERSONALIZED CITATION RECOMMENDATION ON AAN DATA SET

	MAP	MRR	Recall @20	Recall @40	Recall @60	Recall @80	Recall @100
$CQR_{,q_2}$	0.143	0.156	0.256	0.347	0.419	0.443	0.493
$CQR_{,q_1}$	0.124	0.135	0.239	0.328	0.396	0.421	0.475
$CQR-PA_{,q_2}$	0.120	0.133	0.237	0.325	0.389	0.403	0.452
$CQR-PA_{,q_1}$	0.100	0.111	0.213	0.280	0.342	0.361	0.401
cosine	0.082	0.096	0.196	0.253	0.319	0.337	0.382
CF	0.081	0.095	0.196	0.250	0.318	0.335	0.379
okapi	0.079	0.092	0.193	0.250	0.315	0.334	0.377
K-L	0.078	0.090	0.191	0.249	0.312	0.331	0.376

surprising to find that QR-P/CQR-P shows the poorest performance in the original graph/subgraph, because it only utilizes the paper layer information, whereas the QR-PV, QR-PA, and CQR approaches integrate the additional author layer and venue layer information. Besides this, the results also demonstrate that the performance of QR-PA and QR-PV is worse than QR. We attribute this to the ability of QR, which considers both the interrelationships and intrarelations among papers, authors, and venues; however, some of these relationships are ignored in QR-PA and QR-PV. We also notice that the performance of QR-PV is better than that of QR-PA, which means that adding the author information into the QR model has a little effect on the performance of nonpersonalized recommendation. It is reassuring to see that the performance of CQR is quite comparable to QR.

3) *Comparison With Personalized and Nonpersonalized Citation Recommendation*: We conduct this set of experiments to investigate whether personalized recommendation can provide more appropriate and customized results to the individuals than nonpersonalized recommendation. Let $q_1 = [q_t]$ denote a nonpersonalized query and $q_2 = [q_t, q_a]$ denote a personalized query. When a user who submits a query to the recommendation has not published papers, the query information is only based on the query text information. So the personalized recommendation will be reduced to a nonpersonalized recommendation for the user. We use the CQR approach in the following experiments to avoid high computational complexity.

From Tables VI–VIII, we can see that the performance of nonpersonalized recommendation is inferior to that of

TABLE VII
COMPARISON OF PERFORMANCE BETWEEN PERSONALIZED AND
NONPERSONALIZED CITATION RECOMMENDATION
ON DBLP DATA SET

	MAP	MRR	Recall @20	Recall @40	Recall @60	Recall @80	Recall @100
CQR- q_2	0.135	0.143	0.248	0.335	0.408	0.435	0.486
CQR- q_1	0.113	0.122	0.226	0.300	0.365	0.381	0.426
CQR-PA- q_2	0.110	0.115	0.219	0.286	0.351	0.370	0.413
CQR-PA- q_1	0.098	0.109	0.212	0.278	0.340	0.358	0.398
cosine	0.080	0.094	0.195	0.251	0.317	0.336	0.380
CF	0.079	0.093	0.194	0.248	0.316	0.334	0.378
okapi	0.076	0.090	0.191	0.247	0.314	0.332	0.376
K-L	0.075	0.088	0.189	0.246	0.311	0.330	0.375

TABLE VIII
COMPARISON OF PERFORMANCE BETWEEN PERSONALIZED AND
NONPERSONALIZED CITATION RECOMMENDATION
ON CITESEERML DATA SET

	MAP	MRR	Recall @20	Recall @40	Recall @60	Recall @80	Recall @100
CQR- q_2	0.139	0.146	0.251	0.340	0.411	0.439	0.490
CQR- q_1	0.115	0.127	0.230	0.303	0.367	0.385	0.428
CQR-PA- q_2	0.114	0.123	0.223	0.298	0.365	0.381	0.420
CQR-PA- q_1	0.098	0.112	0.214	0.280	0.342	0.360	0.400
cosine	0.081	0.095	0.196	0.252	0.318	0.337	0.381
CF	0.080	0.094	0.195	0.249	0.317	0.335	0.379
okapi	0.078	0.091	0.192	0.248	0.316	0.334	0.378
K-L	0.076	0.089	0.190	0.247	0.312	0.332	0.376

personalized recommendation. The personalized recommendation achieves an average gain of about 12.87% in the three data sets. When we compare the correct recommendation papers, with regard to nonpersonalized and personalized recommendation approaches, we observe that the personalized recommendation approach can find more papers published by co-authors or published in the venues with related research fields. We compare the difference between the top-60 recommendations made by CQR with q_1 and CQR with q_2 . The overlap between them is about 71.43% in the AAN data set, 66.28% in the DBLP data set, and 68.56% in the CiteSeer ML dataset, respectively. CQR with q_2 performs much better in the top-5 recommendations. For the top-3 recommendations, the accuracy of CQR with q_2 is about 78.61%, 73.62%, 76.59% more than that of CQR with q_1 in the AAN data set, the DBLP data set and the CiteSeer ML data set, respectively. Clearly, the proposed personalized CQR approaches outperform the CF approach.

4) *Comparison With Other Citation Recommendation Approaches*: In order to evaluate the performance of the proposed approaches, we compare them with the other citation recommendation approaches.

- 1) *The neural probabilistic model-based approach* [21], which simultaneously learns the distributed representations of cited papers and citation contexts. As this approach is used for context-based citation recommendation, we replace citation context with the content of manuscript in this paper.
- 2) *The Chronological Citation Recommendation Approach* [17], which assumes initial user information needs could shift while users are searching for papers in different time slices. As this approach focuses on chronological features, we split the AAN training

TABLE IX
COMPARISON OF DIFFERENT RECOMMENDATION
APPROACHES ON AAN DATA SET

	MAP	MRR	Recall @20	Recall @40	Recall @60	Recall @80	Recall @100
QR- q_1	0.126	0.137	0.242	0.331	0.401	0.425	0.479
CQR- q_1	0.124	0.135	0.239	0.328	0.396	0.421	0.475
Neural probabilistic Model	0.119	0.126	0.234	0.299	0.362	0.393	0.441
Chronologica l approach	0.116	0.123	0.229	0.295	0.358	0.390	0.437
LocDiSCern	0.113	0.120	0.226	0.292	0.355	0.387	0.421
GloDiSCern	0.112	0.119	0.225	0.290	0.353	0.385	0.418
Unified graph model	0.108	0.117	0.221	0.286	0.349	0.381	0.410
Multi-layer graph model	0.091	0.108	0.209	0.275	0.331	0.368	0.396
HITS	0.073	0.089	0.186	0.258	0.314	0.343	0.378

data set into nine time slices: 1965–1989, 1990–1999, 2000–2002, 2003–2005, 2006–2007, 2008–2009, 2009–2010, 2011, and 2012. Similarly, the DBLP training data set is split into ten time slices: pre-1990, 1990–1999, 2000–2002, 2003–2005, 2006–2007, 2008–2009, 2010, 2011, 2012, and 2013. We also split the CiteSeer ML training data set into eight time slices: pre-1990, 1990–1999, 2000–2002, 2003–2005, 2006–2007, 2008–2009, 2010, and 2011. Each time slice of each training data set contains a similar number of papers. For citation recommendation evaluation, we select test papers in each training data set according to the predefined criterion in [17]. Meanwhile, MetaPath2 achieves the best averaged evaluation scores in [17], so we use MetaPath2 in this paper.

- 3) *DiSCern* [18], which retrieves relevant and diversified citations in response to a given query. We need to extract keywords in the keyword section of each paper, if a paper does not contain a keyword section, we apply a graph degeneracy-based approach [51] to extract keywords. We use both LocDiSCern and GloDiSCern in this paper.
- 4) *HITS* [52], which first generates a bipartite graph consisting of two kinds of vertices, representing the papers and words, respectively, then applies the HITS algorithm to assign each candidate reference paper a hub score, and finally the top- N papers are recommended according to the hub scores.
- 5) *Multilayer Graph* [15], which first constructs a heterogeneous graph based on both citation and content information within papers, and then applies a graph-based similarity learning algorithm to perform the citation recommendation task and
- 6) *Unified Graph Model* [1], which incorporates various types of useful information into a multilayer graph to perform citation recommendation. We also ignore the query author information in this set of experiments. The results of our proposed approaches and the other citation recommendation approaches are shown in Tables IX–XI.
- 5) *Efficiency Comparison*: In order to examine the efficiencies of the proposed approaches, we use papers published

TABLE X
COMPARISON OF DIFFERENT RECOMMENDATION
APPROACHES ON DBLP DATA SET

	MAP	MRR	Recall @20	Recall @40	Recall @60	Recall @80	Recall @100
QR, q_i	0.115	0.125	0.228	0.302	0.366	0.383	0.428
CQR, q_i	0.113	0.122	0.226	0.300	0.365	0.381	0.426
Neural probabilistic Model	0.099	0.113	0.221	0.282	0.344	0.375	0.402
Chronologica l approach	0.096	0.110	0.218	0.279	0.341	0.372	0.399
LocDiSCern	0.092	0.107	0.214	0.275	0.337	0.368	0.395
GloDiSCern	0.091	0.105	0.212	0.273	0.334	0.365	0.392
Unified graph model	0.085	0.098	0.201	0.266	0.325	0.358	0.387
Multi-layer graph model	0.068	0.082	0.181	0.253	0.318	0.342	0.374
HITS	0.051	0.063	0.159	0.232	0.289	0.323	0.350

TABLE XI
COMPARISON OF DIFFERENT RECOMMENDATION
APPROACHES ON CITESEER ML DATA SET

	MAP	MRR	Recall @20	Recall @40	Recall @60	Recall @80	Recall @100
QR, q_i	0.117	0.128	0.231	0.305	0.368	0.387	0.430
CQR, q_i	0.115	0.127	0.230	0.303	0.367	0.385	0.428
Neural probabilistic Model	0.110	0.121	0.229	0.290	0.356	0.382	0.422
Chronologica l approach	0.109	0.121	0.232	0.289	0.349	0.385	0.410
LocDiSCern	0.105	0.116	0.227	0.286	0.345	0.380	0.405
GloDiSCern	0.103	0.114	0.226	0.285	0.343	0.378	0.401
Unified graph model	0.096	0.105	0.213	0.262	0.336	0.369	0.398
Multi-layer graph model	0.079	0.093	0.192	0.265	0.323	0.357	0.381
HITS	0.060	0.072	0.168	0.246	0.298	0.332	0.361

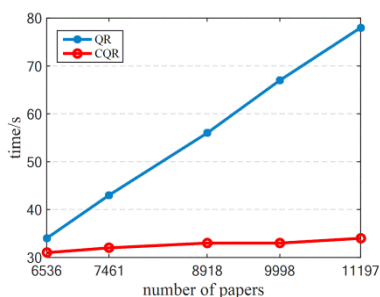


Fig. 4. Runtime comparison between QR and CQR.

before 2009, 2010, 2011, 2012, and 2013 from the AAN data set to construct the three-layered graphs, and evaluate the recommendations made from those papers for the remaining 1358 AAN data set papers published in 2013. The horizontal axis of Fig. 4 shows the number of published papers which were published before the above five years, and the vertical axis of Fig. 4 shows the runtime of each data set using the QR and CQR approaches. It clearly shows that when adding more vertices to the graph, the running time of both approaches increases, because the approaches have to process the increasing size of the similarity relationships. However, as CQR applies the mutual reinforcement principle on the induced subgraph, it is more efficient than the QR approach.

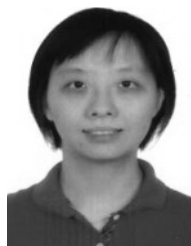
VI. CONCLUSION

In this paper, we propose a three-layered mutual reinforcement model to recommend personalized query-oriented reference papers. Experiments performed on the three different data sets demonstrate the effectiveness of the proposed approaches in comparison with the other citation recommendation approaches. We propose to apply three-layered interactive clustering on the constructed three-layered graph to solve the huge size problem of the graph. The mutual reinforcement principle can then be applied on the subgraph generated by the paper clusters. Our experiments confirm the efficiency of the improved approach.

REFERENCES

- [1] F. Meng, D. Gao, W. Li, X. Sun, and Y. Hou, "A unified graph model for personalized query-oriented reference paper recommendation," in *Proc. 22nd ACM Conf. Inf. Knowl. Manag.*, San Francisco, CA, USA, 2013, pp. 1509–1512.
- [2] C. Nascimento, A. Laender, A. Silva, and M. G. Alves, "A source independent framework for research paper recommendation," in *Proc. 11th Joint Conf. Dig. Libraries*, 2011, pp. 297–306.
- [3] A. Hernando, J. Bobadilla, and F. Ortega, "A non negative matrix factorization for collaborative filtering recommender systems based on a Bayesian probabilistic model," *Knowl. Based Syst.*, vol. 97, pp. 188–202, Apr. 2016.
- [4] S. M. McNee *et al.*, "On the recommending of citations for research papers," in *Proc. ACM Conf. Comput.*, 2002, pp. 116–125.
- [5] C. Yang, B. Wei, J. Wu, Y. Zhang, and L. Zhang, "CARES: A ranking-oriented CADAL recommender system," in *Proc. 9th Joint Conf. Dig. Libraries*, Austin, TX, USA, 2009, pp. 203–212.
- [6] Z. Kang, C. Peng, and Q. Cheng, "Top-N recommender system via matrix completion," in *Proc. 30th AAAI Conf.*, Phoenix, AZ, USA, 2016, pp. 179–185.
- [7] K. Chandrasekaran, S. Gauch, P. Lakkaraju, and H. Luong, "Concept-based document recommendations for CiteSeer authors," in *Proc. Int. Conf. Adap. Hypermedia Adap. Web Syst.*, Hannover, Germany, 2008, pp. 83–92.
- [8] R. Torres, S. M. McNee, M. Abel, J. A. Konstan, and J. Riedl, "Enhancing digital libraries with TechLens+," in *Proc. 4th Joint Conf. Dig. Libraries*, Tucson, AZ, USA, 2004, pp. 228–236.
- [9] M. Balabanović and Y. Shoham, "Fab: Content-based, collaborative recommendation," *Commun. ACM*, vol. 40, no. 3, pp. 66–72, 1997.
- [10] P. Lops, M. Gemmis, and G. Semeraro, "Content-based recommender systems: State of the art and trends," in *Recommender Systems Handbook*. Berlin, Germany: Springer, 2011.
- [11] S. McNee, "Meeting user information needs in recommender systems," Ph.D. dissertation, Dept. Comput. Sci. Eng., Univ. Minnesota, Minneapolis, MN, USA, 2006.
- [12] T. Strohmaier, W. Croft, and D. Jensen, "Recommending citations for academic papers," in *Proc. 30th SIGIR Conf.*, Amsterdam, The Netherlands, 2007, pp. 705–706.
- [13] D. Zhou *et al.*, "Learning multiple graphs for document recommendations," in *Proc. 17th Int. Conf. World Wide Web*, Beijing, China, 2008, pp. 141–150.
- [14] M. Gori and A. Pucci, "Research paper recommender systems: A random-walk based approach," in *Proc. IEEE/WIC/ACM Int. Conf. Web Intell.*, Hong Kong, Dec. 2006, pp. 778–781.
- [15] L. Pan, X. Dai, S. Huang, and J. Chen, "Academic paper recommendation based on heterogeneous graph," in *Chinese Computational Linguistics and Natural Language Processing Based on Naturally Annotated Big Data*. Berlin, Germany: Springer, 2015, pp. 381–392.
- [16] Z. Jiang, X. Liu, and L. Gao, "Chronological citation recommendation with information-need shifting," in *Proc. 25th ACM Conf. Inf. Knowl. Manag.*, 2016, pp. 1291–1300.
- [17] J. Wang, D. Song, Q. Wang, Z. Zhang, L. Si, and L. Liao, "An entity class-dependent discriminative mixture model for cumulative citation recommendation," in *Proc. 38th SIGIR Conf.*, 2015, pp. 635–644.
- [18] T. Chakraborty, N. Modani, R. Narayanan, and S. Nagar, "DiSCern: A diversified citation recommendation system for scientific queries," in *Proc. 31st ICDE Conf.*, 2015, pp. 555–566.
- [19] Q. He, J. Pei, D. Kifer, P. Mitra, and C. Giles, "Context-aware citation recommendation," in *Proc. 19th WWW Conf.*, Raleigh, NC, USA, 2010, pp. 421–430.

- [20] T. Ebesu and Y. Fang, "Neural citation network for context-aware citation recommendation," in *Proc. 40th SIGIR Conf.*, Tokyo, Japan, 2017, pp. 1093–1096.
- [21] W. Huang, Z. Wu, C. Mitra, and C. Giles, "A neural probabilistic model for context based citation recommendation," in *Proc. 29th AAAI Conf.*, Austin, TX, USA, 2015, pp. 2404–2410.
- [22] X. Xu, Z. Huang, L. Zuo, and H. He, "Manifold-based reinforcement learning via locally linear reconstruction," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 4, pp. 934–947, Apr. 2017.
- [23] C. Yu, M. Zhang, F. Ren, and G. Tan, "Emotional multiagent reinforcement learning in spatial social dilemmas," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 12, pp. 3083–3096, Dec. 2015.
- [24] Z. Lin, G. Ding, J. Han, and J. Wang, "Cross-view retrieval via probability-based semantics-preserving hashing," *IEEE Trans. Cybern.*, vol. 47, no. 12, pp. 4342–4355, Dec. 2017.
- [25] H. Zha, "Generic summarization and key phrase extraction using mutual reinforcement principle and sentence clustering," in *Proc. 25th SIGIR Conf.*, Tampere, Finland, 2002, pp. 113–120.
- [26] F. Wei, W. Li, Q. Lu, and Y. He, "Query-sensitive mutual reinforcement chain and its application in query-oriented multi-document summarization," in *Proc. 31st SIGIR Conf.*, Boston, MA, USA, 2009, pp. 283–290.
- [27] X. Cai and W. Li, "Mutually reinforced manifold-ranking based relevance propagation model for query-focused multi-document summarization," *IEEE Trans. Audio, Speech, Language Process.*, vol. 20, no. 5, pp. 1597–1607, Jul. 2012.
- [28] L. Labiod and M. Nadif, "A unified framework for data visualization and cocustering," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 26, no. 9, pp. 2194–2199, Sep. 2015.
- [29] B. Gao, T.-Y. Liu, and W.-Y. Ma, "Star-structured high-order heterogeneous data co-clustering based on consistent information theory," in *Proc. 6th IEEE Int. Conf. Data Min.*, Dec. 2006, pp. 880–884.
- [30] M. Rege and Q. Yu, "Efficient mining of heterogeneous star-structured data," *Int. J. Softw. Inf.*, vol. 2, no. 2, pp. 141–161, 2008.
- [31] Y. Chen, L. Wang, and M. Dong, "Non-negative matrix factorization for semisupervised heterogeneous data cocustering," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1459–1474, Oct. 2010.
- [32] Y. Wang, W. Zhang, L. Wu, X. Lin, M. Fang, and S. Pan, "Iterative views agreement: An iterative low-rank based structured optimization method to multi-view spectral clustering," in *Proc. 25th Joint Conf. Neural Netw.*, 2016, pp. 2153–2159.
- [33] D. Zhang, D. Meng, and J. Han, "Co-saliency detection via a self-paced multiple-instance learning framework," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 5, pp. 865–878, May 2017.
- [34] J. Bouvrie, "Multi-source contingency clustering," M.S. thesis, Dept. Elect. Eng. Comput. Sci., Massachusetts Inst. Technol., Cambridge, MA, USA, 2004.
- [35] R. Bekkerman, R. Yaniv, and A. McCallum, "Multi-way distributional clustering via pairwise interactions," in *Proc. 22nd Int. Conf. Mach. Learn.*, Bonn, Germany, 2005, pp. 41–48.
- [36] H. Cheng, P. Tan, J. Sticklen, and W. Punch, "Recommendation via query centered random walk on k-partite graph," in *Proc. 7th IEEE Int. Conf. Data Min.*, Omaha, NE, USA, Oct. 2007, pp. 457–462.
- [37] J. Han, R. Quan, D. Zhang, and F. Nie, "Robust object co-segmentation using background prior," *IEEE Trans. Image Process.*, vol. 27, no. 4, pp. 1639–1651, Apr. 2018.
- [38] X. Cai and W. Li, "Enhancing sentence-level clustering with integrated and interactive frameworks for theme-based summarization," *J. Assoc. Inf. Sci. Tech.*, vol. 62, no. 10, pp. 2067–2082, 2011.
- [39] Z. Kang, C. Peng, and Q. Cheng, "Kernel-driven similarity learning," *Neurocomputing*, vol. 267, pp. 210–219, Dec. 2017.
- [40] Z. Kang, C. Peng, and Q. Cheng, "Twin learning for similarity and clustering: A unified kernel approach," in *Proc. 31st AAAI Conf.*, San Francisco, CA, USA, 2017, pp. 2080–2086.
- [41] W. Li, W. K. Ng, Y. Liu, and K. L. Ong, "Enhancing the effectiveness of clustering with spectra analysis," *IEEE Trans. Knowl. Data Eng.*, vol. 19, no. 7, pp. 887–902, Jul. 2007.
- [42] D. Radev and V. Muthukrishnan, "The ACL anthology network corpus," in *Proc. Workshop Text Citation Anal. Scholarly Digit. Libraries*, 2009, pp. 54–61.
- [43] J. Tang, J. Zhang, L. Yao, J. Li, L. Zhang, and Z. Su, "Arnetminer: Extraction and mining of academic social networks," in *Proc. 14th ACM SIGKDD Conf.*, Las Vegas, NV, USA, 2008, pp. 990–998.
- [44] C. Giles, K. D. Bollacker, and S. Lawrence, "Citeseer: An automatic citation indexing system," in *Proc. 3rd DL Conf.*, Pittsburgh, PA, USA, 1998, pp. 89–98.
- [45] K. Lim and W. Buntine, "Bibliographic analysis with the citation network topic model," in *Proc. 6th ACML Conf.*, 2014, pp. 142–158.
- [46] R. Yates and B. Neta, *Modern Information Retrieval*. New York, NY, USA: ACM Press, 1999.
- [47] J. Han, D. Zhang, G. Cheng, N. Liu, and D. Xu, "Advanced deep-learning techniques for salient and category-specific object detection: A survey," *IEEE Signal Process. Mag.*, vol. 35, no. 1, pp. 84–100, Jan. 2018.
- [48] S. Robertson, S. Walker, M. Beaulieu, and M. Gatford, "Okapi at TREC-4," in *Proc. 4th Text Retr. Conf.*, 1993, pp. 73–86.
- [49] C. Zhai and J. Lafferty, "Model-based feedback in the KL-divergence retrieval model," in *Proc. 10th ACM Conf. Inf. Knowl. Manag.*, Atlanta, GA, USA, 2001, pp. 403–410.
- [50] Q. Le and T. Mikolov, "Distributed representations of sentences and documents," in *Proc. 31st ICML Conf.*, Beijing, China, 2014, pp. 1188–1196.
- [51] A. Tixier, F. Malliaros, and M. Vazirgiannis, "A graph degeneracy-based approach to keyword extraction," in *Proc. 16th EMNLP Conf.*, Austin, TX, USA, 2016, pp. 1860–1870.
- [52] M. Ohta, T. Hachiki, and A. Takasu, "Related paper recommendation to support online-browsing of research papers," in *Proc. 4th Int. Conf. Digit. Inf. Web Technol.*, 2011, pp. 130–136.



Xiaoyan Cai is currently an Associate Professor with the School of Automation, Northwestern Polytechnical University, Xi'an, China. Her current research interests include document summarization, information retrieval, and machine learning.



Junwei Han is currently a Professor with the School of Automation, Northwestern Polytechnical University, Xi'an, China. His current research interests include data mining, computer vision, multimedia information processing, and brain imaging analysis.



Wenjie Li is currently an Associate Professor with the Department of computing, The Hong Kong Polytechnic University, Hong Kong. Her current research interests include natural language processing and information extraction.



Renxian Zhang is currently a Senior Researcher and Linguistics Specialist with Yappn Canada Corporation, Markham, ON, Canada. His current research interests include natural language processing and data minings.



Shirui Pan is currently a Research Associate with the Centre for Artificial Intelligence, University of Technology Sydney, Ultimo, NSW, Australia. His current research interests include text mining and network representation.



Libin Yang is currently an Assistant Professor with the School of Automation, Northwestern Polytechnical University, Xi'an, China. His current research interests include information retrieval, computer network, and game theory.