Can I have more of these please?

Assisting researchers in finding similar research papers from a seed basket of papers

Aravind Sesagiri Raamkumar, Schubert Foo and Natalie Pang
Wee Kim Wee School of Communication and Information, Nanyang Technological University, Singapore

Received 8 April 2017
Revised 1 September 2017
Accepted 30 October 2017

Abstract

Purpose – During the literature review phase, the task of finding similar research papers can be a difficult proposition for researchers due to the procedural complexity of the task. Current systems and approaches help in finding similar papers for a given paper, even though researchers tend to additionally search using a set of papers. This paper aims to focus on conceptualizing and developing recommendation techniques for key literature review and manuscript preparatory tasks that are interconnected. In this paper, the user evaluation results of the task where seed basket-based discovery of papers is performed are presented.

Design/methodology/approach – A user evaluation study was conducted on a corpus of papers extracted from the ACM Digital Library. Participants in the study included 121 researchers who had experience in authoring research papers. Participants, split into students and staff groups, had to select one of the provided 43 topics and run the tasks offered by the developed assistive system. A questionnaire was provided at the end of each task for evaluating the task performance.

Findings – The results show that the student group evaluated the task more favourably than the staff group, even though the difference was statistically significant for only 5 of the 16 measures. The measures topical relevance, interdisciplinarity, familiarity and usefulness were found to be significant predictors for user satisfaction in this task. A majority of the participants, who explicitly stated the need for assistance in finding similar papers, were satisfied with the recommended papers in the study.

Originality/value – The current research helps in bridging the gap between novices and experts in terms of literature review skills. The hybrid recommendation technique evaluated in this study highlights the effectiveness of combining the results of different approaches in finding similar papers.

Keywords Digital libraries, Information retrieval, Scholarly article recommender systems, Scholarly articles, Seed baskets

Paper type Research paper

Introduction

Literature review (LR) is an important phase of a research project, as it has direct impact on the subsequent phases. During LR, there are transitions in focus state, activity type and search style. The user moves through three stages starting from pre-focus to a problem formulation stage and then on to the final post-focus stage (Vakkari, 2000). These stages apply to both general-purpose and scientific/academic information-seeking domains. In a typical pre-focus stage, researchers use exploratory search tactics to get an initial set of papers for the given research area. These papers are either manually collated or acquired from experts (Ellis *et al.*, 1993). The initial set of papers can be referred to as the reading list, and this list ideally comprises a mix of seminal, recent, literature survey papers covering the

This research was supported by the National Research Foundation, Prime Minister's Office, Singapore under its International Research Centres in Singapore Funding Initiative and administered by the Interactive Digital Media Programme Office.

The Electronic Library

0264-0473
DOI 10.1108/EL-04-2017-0077

sub-topics of the research area. After obtaining a holistic understanding of the research area, researchers select a few papers from the list for finding more similar papers. The transition to directed search happens during this task. A seed set of papers (called *seed basket* in the context of the current study) from the reading list are used as inputs to this task. During this stage, the researcher executes multiple activities, such as chaining, metadata hyperlinking and extended topical searching, to name a few. Current academic search systems, databases and citation indices are tools used by researchers for performing the aforementioned activities, even though these systems are mainly designed for *ad hoc* searching. As this task involves variegated activities, information sources and relevance criteria, researchers need both skills and additional time. Moreover, novice researchers need assistance in performing this type of information-seeking task (Du and Evans, 2011). Two types of interventions provide the mitigatory measure for this scenario. They are *process-based* and *technology-oriented* interventions. In process-based interventions, the role of librarians and experts in helping other researchers has been underlined (Du and Evans, 2011; Spezi, 2016).

Under technology-oriented interventions, prior studies in the area of scientific paper information retrieval (IR) and recommender systems (RS) have looked at proposing techniques for finding similar papers. Most of the approaches have looked at either one or two of the aforementioned sub-tasks for finding similar papers. In these studies, evaluations have been conducted in an offline environment using simulations, without involving actual users. Most importantly, most of the studies have proposed techniques for finding similar papers for a single input paper. In a real-world scenario, researchers also tend to find similar papers for a set of seed papers (Raamkumar *et al.*, 2016).

With a view to address the abovementioned issues, the researchers developed a system for assisting researchers in three LR and manuscript preparatory tasks. The three tasks are:

(1) building a reading list of research papers;

(2) finding similar papers based on a set of papers; and

(3) shortlisting papers from the final reading list for inclusion in a manuscript based on article type.

These tasks are interconnected using two paper collection features – seed basket (SB) and reading list. The recommendation techniques for these tasks have been conceptualized using a set of pre-computed features (criteria) that capture the important characteristics of a research paper and its relations with bibliographic references and citations. Reproducibility in other environments has been taken as the key characteristic while designing the techniques for tasks. A data set of research papers extracted from the ACM Digital Library (ACM DL) comprising 103,739 articles is used as the corpus for the system.

In this paper, the focus is on the second task addressed in the assistive system: that is, the task of finding similar papers based on a seed set (seed basket) of papers. The conceptual design of the task is first described, followed by the findings of a user evaluation study conducted with 121 researchers. This research contributes to the existing literature in several ways:

- the importance of considering multiple seed papers while designing recommendation tasks for finding similar papers is highlighted;
- the proposed technique of finding similar papers using a SB of papers; and
- identification of measures that have predictive ability over user satisfaction in this task.

Related work

Studies based on implicit data, such as user log footprints and browsing histories, are excluded because papers identified through implicit data are not always considered as seed papers by researchers. Prior studies are categorized into two high level categories:

(1) inter-paper similarity measurement techniques; and

(2) seed set-based discovery of similar papers.

In the former category, studies cover approaches for ascertaining similarity between two research papers, using both citation-based relations and textual relations. In the latter category, studies deal with the case of finding similar papers based on multiple papers.

Inter-paper similarity measurement techniques

The introduction of CiteSeer digital library brought forth the common citation inverse document frequency (CCIDF) technique which was inspired by the popular TF-IDF technique (Jones, 1972). The CCIDF algorithm (Lawrence *et al.*, 1999) calculates similarity of a paper with all the other papers in the corpus. The algorithm factors together the citation count and co-citations of papers (White and Griffith, 1981) towards calculating the CCIDF value. As the algorithm requires the whole corpus for processing, it is considered to be a computationally expensive method. This algorithm has been improved by combining co-references (Kessler, 1963) data in a later study (Huynh *et al.*, 2012) where the results show that the modified algorithm provides better results. Co-citations and co-references have been used in many studies as a base model for incorporating further extensions. Such studies include:

- nested referencing in co-references used to enhance similarity calculation (Yoon *et al.*, 2010); and
- co-citation scores calculation based on the entire network in contrast to the immediate neighbours used in the traditional method (Jeh and Widom, 2002).

Dependency based on relation type of a citation between a citing and a cited paper along with graph distance in citation networks are combined to form two similarity metrics in another citation-oriented study (Liang *et al.*, 2011). These metrics are used for identifying relevant papers via the depth-first-search technique. During experimentation, the proposed approach outperformed the basic citation chaining methods, CCIDF, and graph distance methods. As the aforementioned approach is solely citation-based, it is limited to discovering papers only in citation networks. This apparent gap has been addressed in a recent study (Pan *et al.*, 2015) where citation data and textual data have been combined to form a heterogeneous graph. The graph is subsequently used in a semi-supervised learning algorithm to classify categories of similar papers. As this is a training-based approach, the model needs to be re-run whenever new papers are added. The combination of citations and textual content in formulating recommendations was adopted in another recent study (Chakraborty *et al.*, 2016). The recommendation technique is based on the random walk with the restart (RWR) algorithm and it classifies similar papers into different facets, such as alternate approaches, background, and methods, thereby facilitating easier understanding of the recommended papers for researchers.

The co-authorship network is another vital source for scientific paper recommendations as collaborators mostly tend to research on related topics. However, links between co-authors do not carry any topical information. This issue has been addressed in a study (Hwang *et al.*, 2017) where latent dirichlet allocation (LDA) topic models (Blei *et al.*, 2003) are

integrated with the co-authorship networks. In the recommendation model, scientific papers are conceptualized as author vectors in which elements represent both topic and co-authorship similarity between an author and a paper. This technique might be useful for recommending similar papers for experienced researchers.

Most of the proposed approaches are limited to a single data source or a data set. Data from federated sources form the base of a system where content-based similarity and collaborative filtering (CF) techniques are combined to produce recommendations (Zarrinkalam and Kahani, 2012). Purely text-based approaches in document similarity have also been applied in this domain. Language models and LDA topic models (Blei *et al.*, 2003) have been used to compare abstracts of research papers (Martin *et al.*, 2011). This study makes use of additional metadata fields, such as author-specified keywords, authors and journals, to compute similarity.

Seed set-based discovery of similar papers

One of the earliest studies by Mcnee (2006), the use of CF algorithms for finding similar papers was proposed. The user-item CF variant simulates the bibliographic coupling method (Kessler, 1963), whereas the item-item CF variant simulates the co-citation analysis (Small, 1973). Each paper in the seed set is passed one by one to the user-item matrix so that the recommendations could be generated. In the experiments, these CF variants performed better than content-based retrieval methods in finding more relevant papers. Based on a few papers-of-interest, a recent study (Küçüktunç *et al.*, 2015) used random walk algorithms for finding a diversified set of research papers. The Web service TheAdvisor (Küçüktunç *et al.*, 2013) uses the proposed technique for recommending papers. This technique is solely reliant on citation relations, so probability of finding new papers can be low. A sub-modular optimization approach has been used to propose a streaming algorithm in a study where the intention was to minimize computation when new papers are added to the citation network (Yu *et al.*, 2016). Similar to the previously discussed study, this study is also entirely reliant on the citation network, thereby disadvantaging researchers to a certain extent.

In this research, the aim is to address the problem of finding similar papers with multiple seed papers as input instead of a single research paper. In contrast to previous studies, the aim is to propose a technique that should consider all the seed papers together while formulating recommendations.

Finding similar papers based on a seed set of papers

Assumptions

The task of finding similar papers based on multiple papers is intrinsically different from a single paper. Certain assumptions are to be established so that the requirements of this task are clear. The first assumption is that the researcher may add papers from different sub-topics of a particular research area into the SB. The second assumption is that the researcher may add papers from different research areas into the SB. This scenario occurs for interdisciplinary and multidisciplinary research.

Basic approaches used by researchers

Before proposing the similar papers discovery technique, it is important to re-introduce the general approaches followed by researchers with the aid of academic search systems, databases and citation indices. The proposed technique should operationalize these basic approaches so that the recommendation of papers roughly simulates the manual process followed by researchers.

Chaining. Researchers generally use the *intellectual structure* method (White and Griffith, 1981) that involves the dual steps of backward and forward chaining in the underlying citation network of a research paper. In backward and forward chaining, bibliographic references and citations are mined, respectively, for finding similar papers. Complex chaining methods, such as bibliographic coupling (Kessler, 1963) and co-citation analysis (White, 1990), are used to find relevant papers based on co-references and co-citations, respectively.

Metadata hyperlinking. In most of the academic systems, metadata fields of research papers are displayed as hyperlinks. Researchers generally try to follow the publication trails of certain authors to find recent or even old papers written on the same topic. The same behaviour is repeated for journals, conferences and author-specified keywords. In principle, these hyperlinks facilitate the option of using the *follow-your-nose* method (Hausenblas, 2009) for finding relevant papers by following hyperlinks between papers.

Extended topical searching. In this type of topical searching, researchers make use of certain specific terms from the seed paper(s) for further searching. These terms could be extracted from the title, author-specified keywords, abstract and/or the full text of a research paper. Using this approach, researchers become cognizant of the different sub-topics in the particular research area.

Proposed integrated discovery of similar papers technique

The proposed integrated discovery of similar papers (IDSP) technique for the task of finding similar papers is described in this section. The steps in the technique have been selected with the aim of simulating the manual approaches used by researchers. The process flow of the technique is displayed in Figure 1. Three methods are used to find similar papers based on the input set of papers. The methods are classified under two modules.

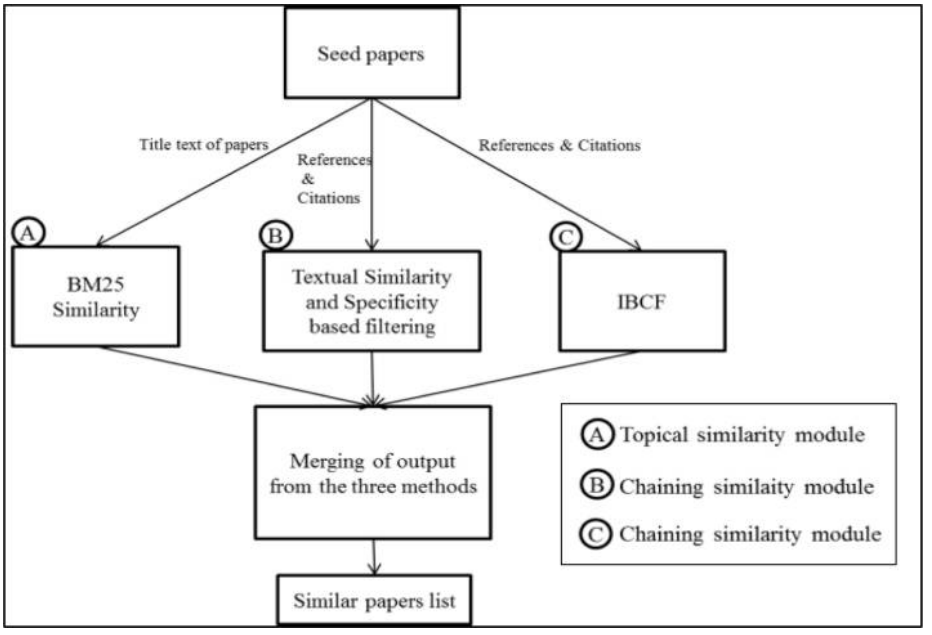


Figure 1.
Process flow in the IDSP technique

Topical similarity module

This module is meant to simulate the extended topical searching approach. For the current study, the title field is the singular field considered for computing similarity. Text from the title fields of all the seed set papers is concatenated to form a single string. This string becomes the input query. The Okapi BM25 similarity score (Jones *et al.*, 2000) is used for the similarity matching of the query with the documents in the corpus. The BM25 method is used, as it offers better performance than other retrieval models (Speriosu and Tashiro, 2006). The top 200 matching papers are retrieved to form set A.

Chaining similarity module

Collaborative filtering. In earlier studies (Ekstrand *et al.*, 2010; Mcnee, 2006), the CF algorithm has been found to perform better than content-based algorithms. The item-based CF variant (IBCF) is selected as it has provided better results than user-based CF variant (UBCF) (Mcnee, 2006). In the user-item matrix of the IBCF algorithm, the user rows are the research papers, whereas the item columns are occupied by the references and citations of the corresponding research papers. The rating is set to 1 between a user and an item (unary item space), as there are no ratings between papers and citations. The value 1 is set if a paper cites the reference. The pictorial representation of the matrix is presented in Figure 2. Five recommendations are retrieved for each SB paper to form set B.

Feature-based filtering. While traditional chaining methods are based on co-occurrences, the relation between a paper and its citations/references can be inferred through further analysis. In the feature-based filtering method, both textual and non-textual relations are measured. The textual and non-textual relations are measured using the features textual similarity and specificity. In the case of textual similarity, the bigram-based dice coefficient is used (Brew and McKelvie, 1996) for calculating the similarity between the paper title and the reference/citation title. Dice coefficient performs better than other methods, such as soundex (Holmes and McCabe, 2002) and edit distance (Kukich, 1992). Semantic textual similarity methods (Han *et al.*, 2012) have not been used in this study, as the incumbent knowledge base needs to be separately trained with the ACM DL corpus. The formula for textual similarity (TS) between two strings $S1$ and $S2$ is given as follows:

$$TS\ (S1, S2) = \frac{2\ \times |pairs(S1) \cap pairs(S2)|}{|pairs(S1)| + |pairs(S2)|} \tag{1}$$

where pairs (X) is a function that generates the pairs of adjacent letters (characters) from the string. This feature can be explained with a simple example. Let us take two strings $S1$ and $S2$ as data and datum, respectively. The bigram sets of these two strings are as {da, at, ta} and {da, at, tu, um}. TS is calculated in the following manner:

	Reference 1	Reference 2	Reference 3	.	.	.	Reference N
Article 1	1	1					1
Aritcle 2			1				
Article 3	1		1				1
.							
.							
.							
.							
Article N	1						1

Figure 2.
User-item matrix in IBCF

$$TS\ (S1,\ S2) = \frac{2\ \times\ |\{da,\ at\}|}{|\{da,\ at,\ ta\}| + |\{da,\ at,\ \ tu,\ um\}|} = \frac{2 \times 2}{3+4} = 0.57 \quad (2)$$

The non-textual relation feature specificity is meant to identify the similarity between a research paper and its references/citations based on commonality in the metadata fields: author-specified keywords, primary category and secondary category. The two category fields are specific to the ACM DL; therefore, corresponding fields are to be identified if this feature is implemented with a different corpus. The formula for specificity (*SP*) between a paper *P* and a reference/citation *F* is given as follows:

$$SP\ (P,F) = CK(P,F) + CP(P,F) + CS(P,F) \quad (3)$$

where *F* is a citation or reference of paper *P*. *CK* (common keywords) is a function that counts the number of shared author-specified keywords between *P* and *F*. *CP* (common primary categories) is a function that counts the number of shared primary categories between *P* and *F*. *CS* (Common Secondary categories) is a function that counts the number of shared second categories between *P* and *F*. The TS and specificity (SP) values of research papers are combined to form set *C*.

Merging of outputs from the two modules

The three sets *A*, *B* and *C* from the two modules are to be merged to form set *D*. The papers in *D* are sorted based on the descending order of citation count of the papers. First, the papers that are already present in the user's initial reading list are excluded. Second, the papers that are present in all the three sets *A, B* and *C* are retrieved from *D* to the final recommendation list *L*. The remaining papers for *L* are retrieved based on their respective positions in *D*. The count of recommended papers in *L* can be adjusted as per requirement.

Prototypical assistive system

Prior studies (Küçüktunç *et al.*, 2015; Mcnee, 2006) have evaluated individual LR and manuscript writing tasks as separate tasks. In the current study, a requirement to combine the tasks as a part of a single system is identified so that there is a semblance of progress in LR for researchers. In this section, the assistive system is introduced. The task screens and display features along with the underlying corpus/data set are described.

Overview

The Rec4LRW system is a prototypical assistive system developed to help researchers in two main search tasks of LR and one manuscript preparatory task (Sesagiri Raamkumar *et al.*, 2015). The three tasks are:

(1) building an initial reading list of research papers;

(2) finding similar papers based on a set of papers; and

(3) shortlisting papers from the final reading list for inclusion in a manuscript based on article-type choice.

A minimalist design principle was adapted for the task screens so that the user's focus is retained for evaluating the recommendations. A screenshot of the reading list task (Task 1) screen is displayed in Figure 3. Apart from the regular display features, such as article year, author name(s), abstract, publication year and citation count, the system displays some new features: author-specified keywords, references count and a short summary of the paper (if

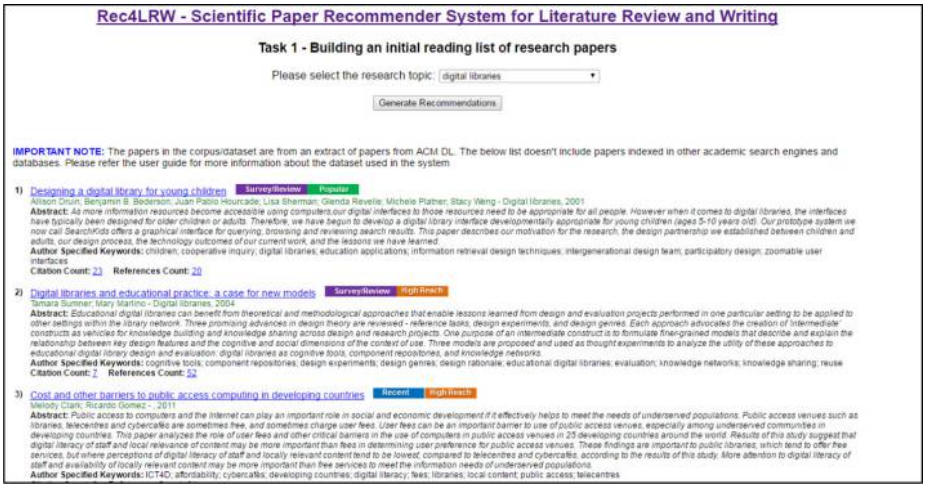


Figure 3.
Reading list task screen in the Rec4LRW system

the abstract of the paper is missing). Information cue labels are placed beside the title for each paper. There are four labels: popular, recent, high reach and survey/review.

A screenshot of Task 2 (the focus of this paper) is provided in Figure 4. Seed papers from Task 1 are input into this task. The first step for the user is to re-run Task 1 so that the seed papers can be selected. In Figure 4, there is a checkbox provided at the left of each paper title for selecting the paper. After selecting the required number of papers, the user can click on "Generate Recommendations". The output for Task 2 is provided in Figure 5. The system displays the SB papers at the top, followed by the recommended papers. For each recommended paper, there are two hyperlinks – shared co-references and shared co-citations – which can be clicked to view the shared relations with the SB papers.

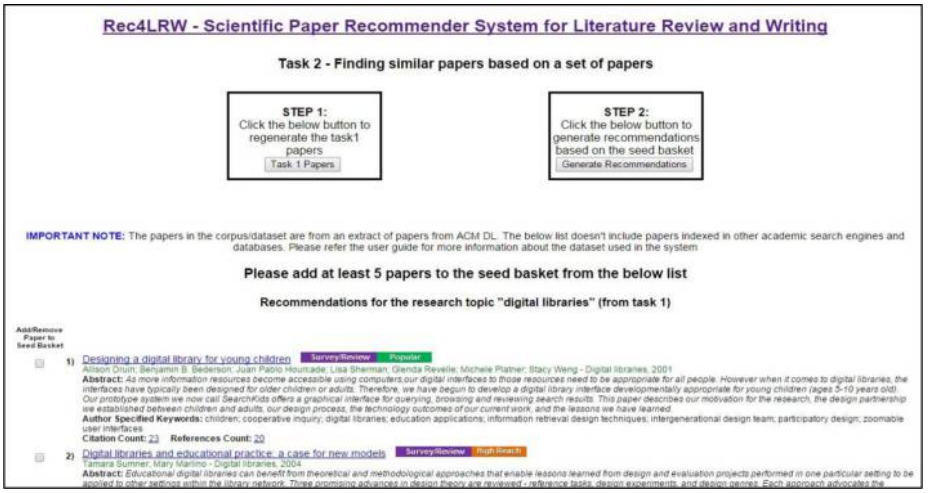
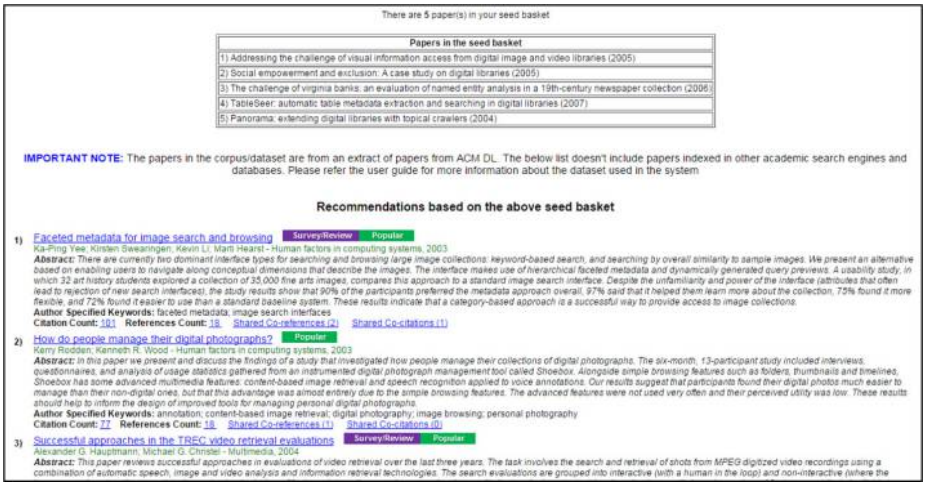


Figure 4.
Selecting seed papers before executing task 2


There are 5 paper(s) in your seed basket

Papers in the seed basket
1) Addressing the challenge of visual information access from digital image and video libraries (2005)
2) Social empowerment and exclusion: A case study on digital libraries (2005)
3) The challenge of virginia banks: an evaluation of named entity analysis in a 19th-century newspaper collection (2006)
4) TableSeer: automatic table metadata extraction and searching in digital libraries (2007)
5) Panorama: extending digital libraries with topical crawlers (2004)

IMPORTANT NOTE: The papers in the corpus/dataset are from an extract of papers from ACM DL. The below list doesn't include papers indexed in other academic search engines and databases. Please refer the user guide for more information about the dataset used in the system

Recommendations based on the above seed basket

1) Faceted metadata for image search and browsing — Survey/Review — Popular
Ka-Ping Yee, Kirsten Swearingen, Kevin Li, Marti Hearst - Human factors in computing systems, 2003
Abstract: *There are currently two dominant interface types for searching and browsing large image collections: keyword-based search, and searching by overall similarity to sample images. We present an alternative based on enabling users to navigate along conceptual dimensions that describe the images. The interface makes use of hierarchical faceted metadata and dynamically generated query previews. A usability study, in which 32 art history students explored a collection of 35,000 fine arts images, compares this approach to a standard image search interface. Despite the unfamiliarity and power of the interface (attributes that often lead to rejection of new search interfaces), the study results show that 90% of the participants preferred the metadata approach overall, 97% said that it helped them learn more about the collection, 75% found it more flexible, and 72% found it easier to use than a standard baseline system. These results indicate that a category-based approach is a successful way to provide access to image collections.*
Author Specified Keywords: faceted metadata; image search interfaces
Citation Count: 107 **References Count:** 18 Shared Co-references (2) Shared Co-citations (1)

2) How do people manage their digital photographs? — Popular
Kerry Rodden, Kenneth R. Wood - Human factors in computing systems, 2003
Abstract: *In this paper we present and discuss the findings of a study that investigated how people manage their collections of digital photographs. The six-month, 13-participant study included interviews, questionnaires, and analysis of usage statistics gathered from an instrumented digital photograph management tool called Shoebox. Alongside simple browsing features such as folders, thumbnails and timelines, Shoebox has some advanced multimedia features: content-based image retrieval and speech recognition applied to voice annotations. Our results suggest that participants found their digital photos much easier to manage than their non-digital ones, but that this advantage was almost entirely due to the simple browsing features. The advanced features were not used very often and their perceived utility was low. These results should help to inform the design of improved tools for managing personal digital photographs.*
Author Specified Keywords: annotation; content-based image retrieval; digital photography; image browsing; personal photography
Citation Count: 77 **References Count:** 18 Shared Co-references (1) Shared Co-citations (0)

3) Successful approaches in the TREC video retrieval evaluations — Survey/Review — Popular
Alexander G. Hauptmann, Michael G. Christel - Multimedia, 2004
Abstract: *This paper reviews successful approaches in evaluations of video retrieval over the last three years. The task involves the search and retrieval of shots from MPEG digitized video recordings using a combination of automatic speech, image and video analysis and information retrieval technologies. The search evaluations are grouped into interactive (with a human in the loop) and non-interactive (where the*

Figure 5.
Sample list of recommended papers in task 2

Data set

An extract from the ACM Digital Library (ACM DL) was used as the data set/corpus for the Rec4LRW system. Papers from proceedings and periodicals (journals) for the period 1951 to 2011 form the data set. The papers were shortlisted based on full text and metadata availability in the data set, to form the sample set/corpus for the system. The sample set contains a total of 103,739 articles and corresponding 2,320,345 references. The original data from ACM were received in the form of 4,500 XML files. Data were transferred to a MySQL database to facilitate easier storage, processing and retrieval. The references of papers were parsed using the AnyStyle parser (AnyStyle, 2015) for extracting article title and publication year. Apache Lucene and Mahout libraries were used for the IR and RS algorithm implementations.

User evaluation study

Purpose

The purpose of the user evaluation study is to determine whether researchers using the tasks provided by Rec4LRW system can be efficient and effective in conducting the corresponding LR tasks. In this context, researchers' perceptions of the system features, individual characteristics of the recommended papers and overall quality of the recommendation list were measured. In this paper, the findings for the second task (i.e. the task of finding similar papers) are reported. The specific evaluation goals are:

- to ascertain the agreement percentages of the evaluation measures;
- to test the hypothesis that students are more benefitted from the recommendation task in comparison to staff;
- to measure the correlation between the measures and build a regression model with user satisfaction as the dependent variable (DV); and
- to compare the pre-study and post-study variables for understanding whether the target participants are benefitted from the task.

Participant recruitment

The target population for the evaluation study was researchers who had the experience of working on research projects and writing research papers. Hence, the recruitment strategy was designed for a specific audience. Three communication channels were used for advertising the study. Invitation e-mails were sent to students and staff of the authors' university. Advertisement posters were put up in notice boards across the university. Invitation e-mails were also sent to mailing lists related to library and information science and information systems. A pre-screening survey was conducted to shortlist the potential participants. In this survey questionnaire, participants were requested to provide their demographic details and research experience. The main selection criteria were that each participant should have authored at least one conference or journal paper. Based on the responses, only researchers who had written research papers were invited. The pre-screening survey questionnaire can be accessed in this document[1]. The study was conducted from the second week of November 2015 to the end of January 2016. The Rec4LRW system was made available through the internet so that the user evaluation study could be conducted. Participants were permitted to perform the experiment from any location.

Study procedure

The participants had to select a research topic from a list of 43 research topics. In the context of this study, the term *research topic* refers to the author-specified keywords specified by authors in publications. As we were constrained by the size of the data set, we could not provide the free-text search feature to participants. Out of the provided topics, participants used 29 topics. The reading task was the first task run by the participant. Before running the similar papers task, the participant had to add at least five papers in the SB. Subsequently, the system provided 30 recommendations for the similar papers task (Task 2). The detailed study guide provided to the participants can be accessed in this document[2]. The evaluation screen for each task was embedded at the bottom of the screen (a screenshot is provided in Figure 6). The participants had to answer mandatory survey questions and two optional subjective feedback questions as a part of the evaluation. A five-point Likert scale was provided for measuring participant responses for each question. The participants were required to evaluate the three tasks and the overall system.

The survey questions and the corresponding measures are provided in Table I. The measures are classified into three categories:

(1) feature-related (FR) measures are about the features provided as part of the task;

(2) individual aspect (IA) measures are for evaluating the specific characteristics of recommended papers; and

(3) output quality (OQ) measures are evaluating the overall recommendation list.

The FR and IA measures are novel and specific to this study, whereas the OQ measures are standard measures used in RS studies (Knijnenburg *et al.*, 2012).

Analysis

The response values "Agree" and "Strongly Agree" were the two values considered for the calculation of agreement percentages for the measures. Descriptive statistics were used to measure central tendency. Independent samples *t*-test was used to check the presence of statistically significant difference in the mean values of the students and staff group, for

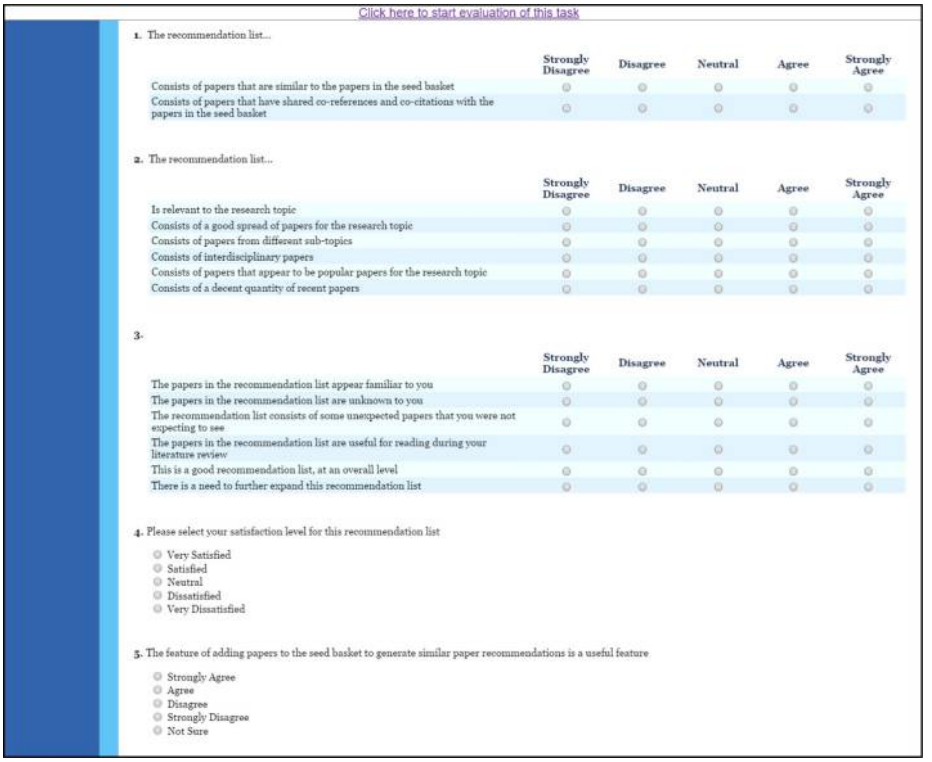


Figure 6.
Evaluation screen of the similar papers task

testing the hypothesis. Spearman correlation coefficient was used to measure the correlation between the measures. For the predictive model, multiple linear regression was used. Statistical significance was set at $p < 0.05$. Statistical analyses were done using SPSS 21.0 and R.

Participant demographics

Among the researchers who answered the pre-screening survey, 230 researchers were found to be eligible. After the study details were sent to them, 138 researchers signed the consent form, of which 119 participants completed the whole experiment inclusive of the three tasks in the system. The reading list task (first task) was completed by 132 participants, whereas 121 participants completed both the first and second task. Out of the 121 participants who completed the second task, which is presented in this paper, 60 participants were PhD/MSc students, whereas 61 were research staff, academic staff and librarians. The average research experience for PhD students was 2.84 years, while for staff it was 7.16 years. With respect to disciplines, 62 per cent of the participants were from the computer science, electrical and electronics; 26 per cent were from information and communication studies; and 12 per cent were from other disciplines. The sample size of 121 participants can be considered to be adequate as the participants comprised an equal mix of beginners and experts with varying degrees of experience within these two groups.

Question	Measure	Category
The recommendation list consists of papers that are similar to the papers in the seed basket	Seedbasket_Similarity	FR
The recommendation list consists of papers that have shared co-references and co-citations with the papers in the seed basket	Shared_Corelations	FR
The recommendation list is relevant to the research topic	Topical_Relevance	IA
The recommendation list consists of a good spread of papers for the research topic	Good_Spread	OQ
The recommendation list consists of papers from different sub-topics	Diversity	IA
The recommendation list consists of interdisciplinary papers	Interdisciplinarity	IA
The recommendation list consists of papers that appear to be popular papers for the research topic	Popularity	IA
The recommendation list consists of a decent quantity of recent papers	Recency	IA
The papers in the recommendation list appear familiar to you	Familiarity	IA
The papers in the recommendation list are unknown to you	Novelty	IA
The recommendation list consists of some unexpected papers that you were not expecting to see	Serendipity	IA
The papers in the recommendation list are useful for reading at the start of your literature review	Usefulness	OQ
This is a good recommendation list, at an overall level	Good_List	OQ
There is a need to further expand this recommendation list	Expansion_Required	IA
Please select your satisfaction level for this recommendation list	User_Satisfaction	OQ
The feature of adding papers to the seed basket to generate similar paper recommendations is a useful feature	Seedbasket_Usefulness	FR

Table I. Evaluation questions and corresponding measures

Results and discussion

Overall evaluation

In Figure 7, the agreement percentages for the 16 measures are displayed for the two groups. In the current study, an agreement percentage above 75 per cent is considered as an indication of higher agreement from the participants. The measures that met the above criteria are *Seedbasket_Similarity* (88.71 per cent for students, 75.41 per cent for staff), *Shared_Co-relations* (82.26 per cent for students), *Topical_Relevance* (90.32 per cent for students, 83.61 per cent for staff), *Good_Spread* (87.10 per cent for students), *Diversity* (79.03 per cent for students), *Usefulnes*s (82.26 per cent for students, 78.69 per cent for staff), *Good_List* (79.03 per cent for students) and *Seedbasket_Usefulness* (96.77 per cent for students, 95.08 per cent for staff). The high agreement on the FR measures indicate high topical similarity with the SB papers because the IDSP technique's design considers all the SB papers for formulating recommendations. The ability of the technique in covering a wide variety of sub-topics in the research area is vindicated with *Good_Spread* and *Diversity* measures. This finding is partially attributed to the ability of the IDSP technique's topical similarity module in finding papers that are not part of the citation networks of the SB papers. The agreement level of the OQ measures *Usefulness* and *Good_List* indicate higher levels of satisfaction. However, the *User_Satisfaction* percentage of both groups was around a decent range (approximately 70 per cent) which is below the set threshold for this study. This finding can be linked to the measure *Expansion_Required* (74.35 per cent for students,

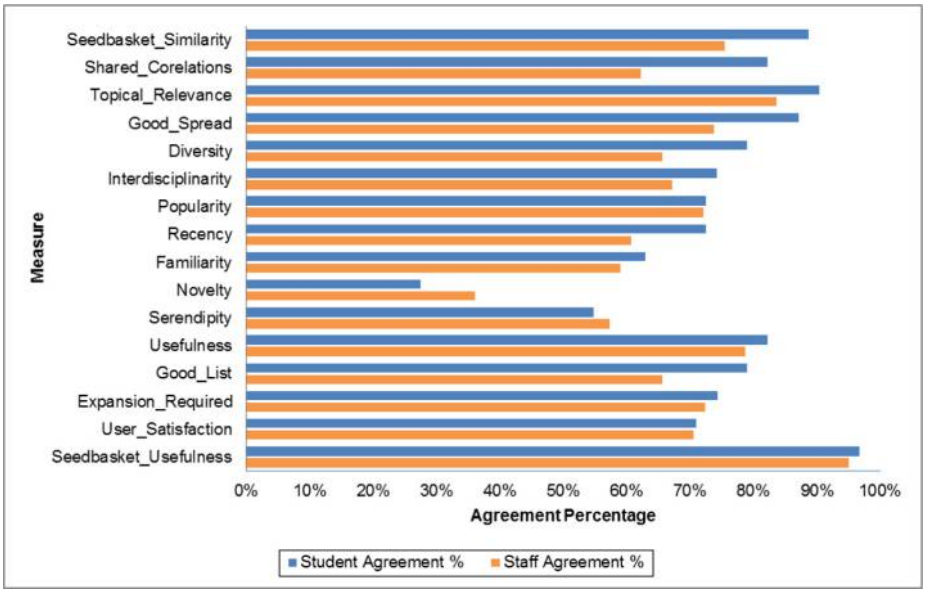


Figure 7. Agreement percentages for the two study groups

72.35 per cent for staff) which highlights the expectation of participants for acquiring more papers in the list, even though the responses for the other measures were favourable.

The measures *Novelty* (27.42 per cent for students, 36.07 per cent for staff) and *Serendipity* (54.84 per cent for students, 57.38 per cent for staff) had low agreement percentages. The ACM data set used in the study consists of papers published before 2011. Therefore, the recommended papers would have been largely familiar to participants. Serendipitous discovery of new papers is also affected by the data set. Second, the current design of the IDSP technique is not prioritized for finding serendipitous papers. The results of the other measures indicate that participants were largely in favour with the recommended papers.

Hypothesis testing

Table II lists the independent samples *t*-test results for the two groups. As for the hypothesis, the students' group rated higher than the staff group at a statistically significant level for five measures *Seedbasket_Similarity, Shared_Co-relations, Topical_Relevance, Good_Spread* and *Good_List*. The difference for the measures *Seedbasket_Similarity* ($M = 4.02$ for students, $M = 3.77$ for staff) and *Shared_Co-relations* ($M = 3.90$ for students, $M = 3.61$ for staff) is an interesting case as these are easily inferable measures. The participants have to merely identify whether the recommended papers are similar to the SB papers. It could be speculated that students were more confident of the similarity than staff. In addition, experienced researchers have higher awareness of their expertise areas; therefore, they would have expected more closely related papers in the list or noticed papers with weak relations to the SB papers. The differences for the OQ measures *Topical_Relevance* ($M = 4.13$ for students, $M = 3.84$ for staff), *Good_Spread* ($M = 3.98$ for students, $M = 3.69$ for staff) and *Good_List* ($M = 3.85$ for students, $M = 3.61$ for staff) can be explained by the

Measure	*t*	Students *M* (SD)	Staff *M* (SD)
Seedbasket_Similarity	1.699*	4.02 (0.558)	3.77 (0.99)
Shared_Co-relations	2.049*	3.90 (0.620)	3.61 (0.954)
Topical_Relevance	2.282*	4.13 (0.558)	3.84 (0.840)
Good_Spread	2.497*	3.98 (0.496)	3.69 (0.786)
Diversity	1.250	3.87 (0.640)	3.70 (0.823)
Interdisciplinarity	0.154	3.76 (0.645)	3.74 (0.814)
Popularity	0.135	3.85 (0.721)	3.84 (0.820)
Recency	0.554	3.69 (0.934)	3.61 (0.802)
Familiarity	0.840	3.60 (0.877)	3.46 (0.941)
Novelty	−1.356	2.81 (0.884)	3.03 (0.966)
Serendipity	0.055	3.47 (0.824)	3.46 (0.923)
Usefulness	0.820	3.97 (0.677)	3.85 (0.872)
Good_List	1.861*	3.85 (0.568)	3.61 (0.881)
Expansion_Required	0.949	3.71 (0.755)	3.56 (1.009)
User_Satisfaction	1.267	3.81 (0.649)	3.62 (0.934)
Seedbasket_Usefulness	0.390	4.32 (0.790)	4.26 (0.730)

Note: *Indicates $p < 0.05$

Table II. Independent samples *t*-test results

corresponding higher ratings for the IA measures by the students' group. The finding shows consistency in their evaluation across the IA measures to the OQ measures.

The lack of statistically significant differences for the other measures indicates two observable characteristics of the recommended papers list. The heterogeneity of the list in providing different types of paper (recent, paper, diverse and interdisciplinary) is acknowledged by the participants. Second, the findings vindicate the nature of the task in improving the discovery of relevant papers from the first task (reading list task). Therefore, there is a semblance of uniformity in participants' evaluation responses, with an inclination towards higher agreeability.

Correlation and regression analysis

In Table III, the measure combinations with correlation coefficient values above the threshold value of 0.5 are displayed. The moderate correlation between *Seedbasket_Similarity* and *Shared_Co-relations* ($R = 0.502$) is an expected observation as both the measures point to the same aspect of topical similarity between the recommended papers and SB papers. The correlation between *Shared_Co-relations* and *Good_Spread* ($R = 0.566$) is interesting because they are conceptually disparate features. The validity of this finding needs to be established

Measure 1	Measure 2	R (95% CI)
Seedbasket_Similarity	Shared_Co-relations	0.502
Shared_Co-relations	Good_Spread	0.566
Good_Spread	Good_List	0.503
Usefulness		0.569
Topical_Relevance	User_Satisfaction	0.633
Good_Spread		0.525
Familiarity		0.539

Table III. Measure combinations with moderate to high correlations

in future studies. The correlation of the measure *Good_List* with *Good_Spread* (R = 0.503) and *Usefulness* (R = 0.569) is another expected finding as they are OQ measures. These three OQ measures are in turn positively correlated with the fourth and important OQ measure *User_Satisfaction*. Hence, there is consistency among OQ measures. The IA measures *Topical_Relevance* (R = 0.633) and *Familiarity* (R = 0.539) are also correlated with *User_Satisfaction*. The inference is that if participants find known relevant papers for the given research topic, they tend to be more satisfied with the overall list. It is to be noted that *Familiarity* is a measure with little use in a real world setting as the recommended papers are supposed to be new to the user while searching for papers for an unknown research topic.

Results from multiple linear regression testing are displayed in Table IV. The model was built with *User_Satisfaction* as the DV and the 14 other measures as the independent variables (IV). The variable *SeedBasket_Usefulness* was not considered as one of the IVs for the model, as it is not related to the quality of the recommended papers. The multiple correlation coefficient R value of 0.83 and the adjusted R^2 value of 0.65 indicate a decent level of prediction at a statistically significant level. The model fit could potentially improve with more participants. Four independent variables *Topical_Relevance, Interdisciplinarity, Familiarity* and *Usefulness* were found to be statistically significant predictors in the model. Relevance is generally an important indicator of satisfaction (Saracevic, 2007). The case with interdisciplinary papers is interesting. Earlier approaches have not designed recommendation approaches for finding interdisciplinary papers. The predictive power of this measure is to be taken into consideration for future studies as researchers have indicated issues in finding interdisciplinary papers (George *et al.*, 2006). Even though the presence of familiar papers had an impact on user satisfaction, the scenario might turn out to be different when researchers are collecting papers for a new research topic. In such a situation, most of the papers would be novel. Therefore, the reliance on survey papers and popular papers would be high. The findings from this regression analysis provide potential to be tested in future studies for this task.

	Estimate	SE	t value	p
Intercept	−0.940	0.461	−2.037	0.044
Seedbasket_Similarity	−0.002	0.076	−0.032	0.974
Shared_Co-relations	0.033	0.075	0.435	0.665
Topical_Relevance	0.412*	0.094	4.365	0.000
Good_Spread	0.142	0.088	1.619	0.108
Diversity	−0.103	0.070	−1.467	0.145
Interdisciplinarity	0.171*	0.066	2.581	0.011
Popularity	0.011	0.071	0.162	0.872
Recency	0.078	0.063	1.240	0.218
Familiarity	0.129*	0.064	2.006	0.047
Novelty	0.044	0.056	0.785	0.434
Serendipity	−0.02	0.053	−0.368	0.714
Usefulness	0.193*	0.081	2.397	0.018
Good_List	0.141	0.08	1.769	0.080
Expansion_Required	−0.006	0.053	−0.118	0.906

Residual standard error: 0.472 on 108 *df*
Multiple R^2: 0.696, Adjusted R^2: 0.657
F-statistic: 17.672 on 14 and 108 *df*, p value: 1.391e−07

Note: *Indicates $p < 0.05$

Table IV.
Multiple linear regression results

Comparison of pre-study and post-study participants' opinions

In Figure 8, a clustered bar chart is illustrated for facilitating the comparison between the pre-study and post-study measures. The pre-study measure is *Issue_Frequency* where participants indicated the frequency of needing external assistance while finding topically similar papers during their LR sessions. The post-study measure is the OQ measure *User_Satisfaction*. From the figure, it is evident that participants, who frequently faced the issue in the past, were largely satisfied with the results of the task. For the frequency value "3" (corresponds to label "Sometimes"), 34 (31 satisfied and 3 very satisfied) out of 46 (94.44 per cent) of the participants were satisfied with the results. Similarly for the values "4" and "5" (labels "Very Often" and "Always"), the satisfaction percentages were 63.64 per cent (16 satisfied and 5 very satisfied out of 33) and 100 per cent (4 satisfied and 2 very satisfied out of 6), respectively. These findings show that participants were clearly supportive of the task's performance in lieu of their previous experiences of needing assistance while finding topically similar papers during LR.

Conclusion and future work

This paper outlined a hybrid technique called the IDSP technique for finding similar papers based on a SB of research papers. The findings from a user evaluation study conducted with 121 participants were presented. The IDSP technique was conceptualized based on two modules, namely, topical similarity and citation similarity module, so that similar papers could be identified from both citation networks of SB papers and the whole corpus. The technique takes multiple seed papers for formulating recommendations, thereby overcoming the gap in earlier studies where similar papers were found for an input paper.

The evaluation results indicated that the students' group found the recommended papers to be more useful than the staff group. In earlier studies (Du and Evans, 2011; Karlsson *et al.*, 2012), graduate research students were found to be in need for more assistance while conducting LR; therefore, the results from this study are encouraging for this group. Among

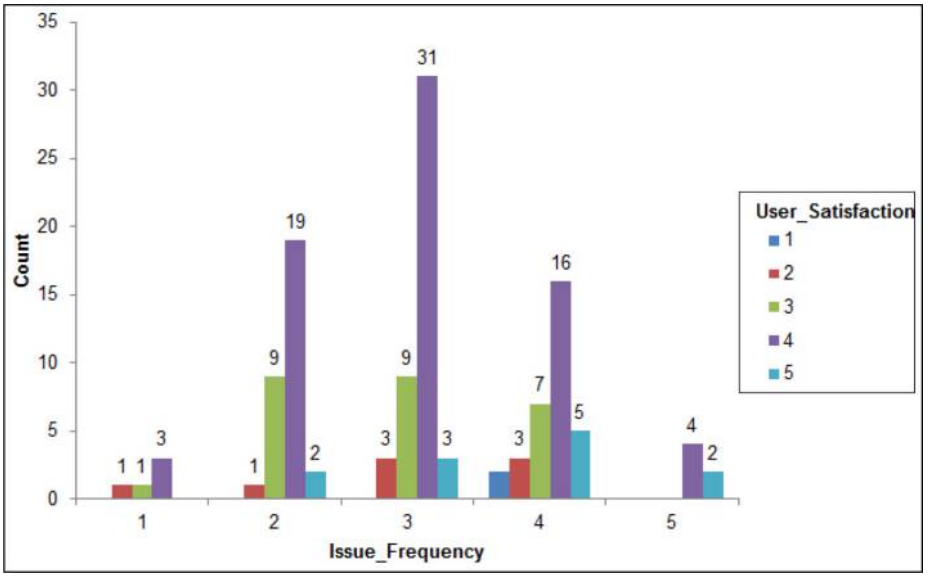


Figure 8.
Pre-study and post-study comparison

the IA measures, topical relevance and diversity found the highest agreement among participants. The same preference was observed for the OQ measures good list, usefulness and good spread. A majority of the participants, who indicated that they needed external assistance in the past for finding topically similar papers, were satisfied with the recommended papers in the study.

The primary contributions of this paper are the IDSP technique, the SB feature and the novel insights gleaned from the evaluation study. The IDSP recommendation technique can be directly implemented in academic digital libraries and task-based search systems. Similar to the tool TheAdvisor (Küçüktunç *et al.*, 2013), the users could also be given the flexible option of uploading BibTeX files of seed papers for generating recommendations from the IDSP technique. In terms of research implications, the evaluation measures topical relevance, interdisciplinarity, spread and usefulness were found to be important indicators for user satisfaction for this recommendation task. Based on the findings, researchers seem to apply a set of sequential criteria while evaluating recommended papers for this task. The starting criteria is topical relevancy where the recommended papers should be conceptually similar to the papers in the SB, followed by relatively complex criteria, such as the need for interdisciplinary and a diverse set of papers. This particular pattern corresponds to the *concertina* metaphor in searching (Levy and Ellis, 2006) where researchers narrow down to a set of papers followed by broadening out of their seeking to look out for more variety. A good spread of papers refers to both temporal and topical variety where papers are from different periods while covering a range of sub-topics of the particular research area. We believe these findings will be helpful for future studies.

There are certain limitations with the proposed technique and the user evaluation study. The technique has not been designed to give priority to any particular paper in the SB for formulating recommendations. Some participants felt that they needed the option of providing different weights to papers in the SB so that certain papers exerted more influence in the final recommendations list. Even though the technique takes all the SB papers together for finding similar papers, the recommended papers may not be related to all the papers in the SB as the final list is ranked based on citation count. Alternate ranking schemes are to be tested to overcome this issue. The latest papers in the dataset were published in 2011. Few participants indicated that they expected to see recently published papers. For executing the task, the minimum number of papers for the SB was five. Certain participants wanted to run the task with only one or two papers so that they could use such paper(s) as a starting point to explore deeper. This issue could have hampered their evaluation of the task.

As a part of future work, the authors plan to modify the IDSP technique for setting priority weights for papers in the seed papers so that the recommendations could be influenced by some papers. The next release of Rec4LRW system will have more UI features so that researchers could sieve through the results for a better understanding of the recommended papers. Additionally, they intend to set user roles in the system so that personalization and customization features are made available for users. Grey literature references from the corpus will be considered for the IDSP technique's re-design so that the recommendation list is a decent mix of different article-types including grey literature articles. The authors previously conducted an analysis of the ACM DL corpus to ascertain the extent of grey literature referencing in research papers and proposed a boosting scheme for pushing certain important articles (Raamkumar *et al.*, 2015). The proposed boosting functionality will be incorporated into the IDSP technique to ensure the presence of grey literature references in the recommendation list. The new system design will allow the user

to control the inclusion of functionalities so that the recommendation logic is more understandable.

Notes

1. Pre-screening survey questionnaire: https://goo.gl/ONGJSO
2. Rec4LRW user guide: http://goo.gl/dxUCuk

References

AnyStyle (2015), "AnyStyle.io", available at: http://anystyle.io (accessed 22 July 2015).

Blei, D.M., Ng, A.Y. and Jordan, M.I. (2003), "Latent dirichlet allocation", *Journal of Machine Learning Research*, Vol. 3, pp. 993-1022.

Brew, C. and McKelvie, D. (1996), "Word-pair extraction for lexicography", *Proceedings of the 2nd International Conference on New Methods in Language Processing*, pp. 45-55.

Chakraborty, T., Krishna, A., Singh, M., Ganguly, N., Goyal, P. and Mukherjee, A. (2016), "FeRoSA: a faceted recommendation system for scientific articles", Pacific Asia Knowledge Discovery and Data Mining Conference (PAKDD '16), Auckland, New Zealand, pp. 528-541.

Du, J.T. and Evans, N. (2011), "Academic users' information searching on research topics: characteristics of research tasks and search strategies", *The Journal of Academic Librarianship*, Vol. 37 No. 4, pp. 299-306.

Ekstrand, M.D., Kannan, P., Stemper, J.A., Butler, J.T., Konstan, J.A. and Riedl, J.T. (2010), "Automatically building research reading lists", *Proceedings of the Fourth ACM Conference on Recommender Systems, ACM Press, New York, NY*, pp. 159-166.

Ellis, D., Cox, D. and Hall, K. (1993), "A comparison of the information seeking patterns of researchers in the physical and social sciences", *Journal of Documentation*, Vol. 49 No. 4, pp. 356-369.

George, C.A., Bright, A., Hurlbert, T., Linke, E.C., Clair, G.S. and Stein, J. (2006), "Scholarly use of information: graduate students' information seeking behaviour", *Information Research*, Vol. 11 No. 4, available at: www.informationr.net/ir/11-4/paper272.html

Han, L., Kashyap, A., Finin, T., Mayfield, J. and Weese, J. (2012), "UMBC EBIQUITY-CORE: semantic textual similarity systems", *Proceedings of the Second Joint Conference on Lexical and Computational Semantics*, pp. 16-33.

Hausenblas, M. (2009), "Exploiting linked data to build web applications", *IEEE Internet Computing*, Vol. 13 No. 4, pp. 68-73.

Holmes, D. and McCabe, C.M. (2002), "Improving precision and recall for soundex retrieval", *Proceedings of International Conference on Information Technology: Coding and Computing*, pp. 22-26.

Huynh, T., Hoang, K., Do, L., Tran, H., Luong, H. and Gauch, S. (2012), "Scientific publication recommendations based on collaborative citation networks", 2012 International Conference on Collaboration Technologies and Systems (CTS), *IEEE, Piscataway, NJ*, pp. 316-321.

Hwang, S.-Y., Wei, C.-P., Lee, C.-H. and Chen, Y.-S. (2017), "Coauthorship network-based literature recommendation with topic model", *Online Information Review*, Vol. 41 No. 3, pp. 318-336.

Jeh, G. and Widom, J. (2002), "SimRank: a measure of structural-context similarity", *Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Edmonton, Alberta*, pp. 538-543.

Jones, K.S. (1972), "A statistical interpretation of term specificity and its application in retrieval", *Journal of Documentation*, Vol. 28 No. 1, pp. 11-21.

Jones, K.S., Walker, S. and Robertson, S.E. (2000), "A probabilistic model of information retrieval: development and comparative experiments: Part 2", *Information Processing & Management*, Vol. 36 No. 6, pp. 809-840.

Karlsson, L., Koivula, L., Ruokonen, I., Kajaani, P., Antikainen, L. and Ruismäki, H. (2012), "From novice to expert: information seeking processes of university students and researchers", *Procedia - Social and Behavioral Sciences*, Vol. 45, pp. 577-587.

Kessler, M.M. (1963), "Bibliographic coupling between scientific papers", *American Documentation*, Vol. 14 No. 1, pp. 10-25.

Knijnenburg, B.P., Willemsen, M.C., Gantner, Z., Soncu, H. and Newell, C. (2012), "Explaining the user experience of recommender systems", *User Modeling and User-Adapted Interaction*, Vol. 22 Nos 4-5, pp. 441-504.

Küçüktunç, O., Saule, E., Kaya, K. and Çatalyürek, Ü.V. (2013), "TheAdvisor: a webservice for academic recommendation", *Proceedings of the 13th ACM/IEEE-CS Joint Conference on Digital Libraries, New York, NY*, pp. 433-434.

Küçüktunç, O., Saule, E., Kaya, K. and Çatalyürek, Ü.V. (2015), "Diversifying citation recommendations", *ACM Transactions on Intelligent Systems and Technology*, Vol. 5 No. 4, pp. 55:1-55:21.

Kukich, K. (1992), "Technique for automatically correcting words in text", *ACM Computing Surveys*, Vol. 24 No. 4, pp. 377-439.

Lawrence, S., Lee Giles, C. and Bollacker, K. (1999), "Digital libraries and autonomous citation indexing", *Computer*, Vol. 32 No. 6, pp. 67-71.

Levy, Y. and Ellis, T.J. (2006), "A systems approach to conduct an effective literature review in support of information systems research", *Informing Science: The International Journal of an Emerging Transdiscipline*, Vol. 9 No. 1, pp. 181-212.

Liang, Y., Li, Q. and Qian, T. (2011), "Finding relevant papers based on citation relations", *Web-Age Information Management (WAIM '11): Proceedings of the 12th International Conference, Wuhan*, September 14-16, Vol. 6897, pp. 403-414.

Mcnee, S.M. (2006), "Meeting user information needs in recommender systems", Doctoral dissertation, University of Minnesota, Minneapolis, MN.

Martin, G.H., Schockaert, S., Cornelis, C. and Naessens, H. (2011), "q", *2nd Workshop on Semantic Personalized Information Management (SPIM '11): Retrieval and Recommendation*, pp. 106-113.

Pan, L., Dai, X., Huang, S. and Chen, J. (2015), "Academic paper recommendation based on heterogeneous graph", *Chinese Computational Linguistics and Natural Language Processing Based on Naturally Annotated Big Data*, Vol. 9427, pp. 381-392.

Raamkumar, A.S., Foo, S. and Pang, N. (2015), "More than just black and white: a case for grey literature references in scientific paper information retrieval systems", *Proceedings of the 17th International Conference on Asia-Pacific Digital Libraries (ICADL '15), Seoul, Korea*, December 9-12, *Springer International Publishing*, pp. 252-257.

Raamkumar, A.S., Foo, S. and Pang, N. (2016), "Survey on inadequate and omitted citations in manuscripts: a precursory study in identification of tasks for a literature review and manuscript writing assistive system", *Information Research*, Vol. 21 No. 4, available at: www.informationr.net/ir/21-4/paper733.html (accessed 20 December 2016).

Saracevic, T. (2007), "Relevance: a review of the literature and a framework for thinking on the notion in information science, Part III: behavior and effects of relevance", *Journal of the American Society for Information Science and Technology*, Vol. 58 No. 13, pp. 2126-2144.

Sesagiri Raamkumar, A., Foo, S. and Pang, N. (2015), "Rec4LRW: scientific paper recommender system for literature review and writing", *Proceedings of the 6th International Conference on Applications of Digital Information and Web Technologies, IOS Press, Hong Kong*, pp. 106-119.

Small, H. (1973), "Co-citation in the scientific literature: a new measure of the relationship between two documents", *Journal of the American Society for Information Science*, Vol. 24 No. 4, pp. 265-269.

Speriosu, M. and Tashiro, T. (2006), *Comparison of Okapi BM25 and Language Modeling Algorithms for NTCIR-6*, available at: https://webspace.utexas.edu/mas5622/www/speriosu06.pdf (accessed 17 March 2014).

Spezi, V. (2016), "Is information-seeking behavior of doctoral students changing? A review of the literature (2010-2015)", *New Review of Academic Librarianship*, Vol. 22 No. 1, pp. 78-106.

Vakkari, P. (2000), "Relevance and contributing information types of searched documents in task performance", *Proceedings of the 23rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, ACM Press, New York, NY*, pp. 2-9.

White, H.D. (1990), "Author co-citation analysis: overview and defense", *Scholarly Communication*, pp. 84-106. in Borgman, C. (Ed.), Sage Publications, Newbury Park, CA.

White, H.D. and Griffith, B.C. (1981), "Author cocitation: a literature measure of intellectual structure", *Journal of the American Society for Information Science*, Vol. 32 No. 3, pp. 163-171.

Yoon, S.-H., Kim, S.-W. and Park, S. (2010), "A link-based similarity measure for scientific literature", *Proceedings of the 19th International Conference on World Wide Web, Raleigh, North Carolina, USA*, pp. 1213-1214.

Yu, Q., Xu, E.L. and Cui, S. (2016), "Submodular maximization with multi-knapsack constraints and its applications in scientific literature recommendations", *IEEE Global Conference on Signal and Information Processing (GlobalSIP '16)*, IEEE, Piscataway, NJ, pp. 1295-1299.

Zarrinkalam, F. and Kahani, M. (2012), "A multi-criteria hybrid citation recommendation system based on linked data", *2nd International eConference on Computer and Knowledge Engineering (ICCKE '12)*, IEEE, Piscataway, NJ, pp. 283-288.

About the authors

Aravind Sesagiri Raamkumar is a Research Associate at the Wee Kim Wee School of Communication and Information, Nanyang Technological University, Singapore. He received his PhD in Information Studies and MSc in Knowledge Management from Nanyang Technological University. His research interests include recommender systems, information retrieval, scholarly metrics, scholarly communication, social media and linked data. Aravind Sesagiri Raamkumar is the corresponding author and can be contacted at: aravind002@ntu.edu.sg

Schubert Foo is a Deputy Associate Provost in the President's Office and Professor of Information Science at the Wee Kim Wee School of Communication and Information, Nanyang Technological University, Singapore. He received his BSc (Hons), MBA and PhD from the University of Strathclyde, UK. He is a Chartered Engineer, Chartered IT Professional, Fellow of the Institution of Mechanical Engineers and Fellow of the British Computer Society. He has authored more than 300 publications in his research areas of multimedia technology, internet technology, multilingual information retrieval, digital libraries, information literacy, knowledge management and social media innovations.

Natalie Pang is an Assistant Professor in the Wee Kim Wee School of Communication and Information, Nanyang Technological University. Prior to joining NTU, she has worked on public opinion research in The Gallup Organization, citizen science and participatory methods in Monash University and Museum Victoria. She specializes in the area of social and community informatics, examining the use and impacts of social media of communities in crises, social movements and everyday life.