# Comparison of Techniques for Measuring Research Coverage of Scientific Papers: A Case Study

Aravind Sesagiri Raamkumar, Schubert Foo, Natalie Pang
Wee Kim Wee School of Communication and Information
Nanyang Technological University, Singapore
*{aravind002, sfoo, nlspang}@ntu.edu.sg*

***Abstract*— With a plethora of research papers published all throughout the year, the task of measuring the research coverage of a paper has become ever-important. Different methods have been proposed for calculating coverage scores using the references and citations network of papers, based on co-ccurences techniques and graph ranking algorithms. In this paper, we propose two techniques for measuring coverage based on author-specified keywords in research papers. We evaluate these techniques with HITS-based coverage technique in a simple experiment using four paper-type inclusivity criteria. Results show that the proposed techniques perform better at providing higher scores for diverse, popular and recent papers while the HITS technique was better at identifying survey papers. The proposed techniques will be used in subsequent studies involving a task-based scientific paper recommender system. The theoretical and practical implications of this study are also been discussed in this paper.**

***Keywords—Research Coverage, Citation Analysis, Author-Specified Keywords, Keywords-based Coverage, Case Study***

## I. INTRODUCTION

Academic databases and search engines are used to find relevant research papers to meet researchers' information needs. These systems rank papers largely based on citation count. Previous studies have commented on the insufficiency of citation count in gauging a paper's contribution to a particular field [1]. Complementing the Information Retrieval (IR) based search engines, Recommender Systems (RS) have been employed to recommend resources based on contextual factors [2]. RS can be tailored to provide task-based recommendations. Such tasks include building a reading list of research papers for initiating literature review, finding similar papers based on a set of papers and other ad-hoc search tasks.

The task of building a reading list of research papers is essential and beneficial to both graduate research students and researchers who are venturing into new research areas. In this task, the relative importance of a research paper is the crucial aspect in formulating recommendations. This relative importance has been perceived as paper seminality and popularity in earlier studies [3]–[5]. The relative importance refers to the research coverage of a research paper in multiple aspects such as diversity, recency, seminality and popularity. Multiple aspects are essential as the expectations of a reading list can be different for researchers based on varying expertise levels, primary discipline and nature of research (academic, applied or translational). Therefore, the techniques for measuring the research coverage of research papers are to be perceived with multiple aspects.

In this paper, we propose two novel research coverage measuring techniques 'Topical Coverage (TC)' and 'Topical and Peripheral Coverage (TPC)' based on author-specified keywords in research papers. Similar to previously proposed approaches, these techniques also rely on references and citations networks of papers albeit the generation of these networks are guided by the author-specified keywords of research papers. At a conceptual level, the two techniques are aimed at uncovering diverse papers for research topics and also identify peripheral papers in the case of inter-disciplinary research topics. We have evaluated these techniques with a popular graph ranking technique 'HITS-based Coverage (HC)' for comparison purposes using four inclusivity evaluation criteria for diverse, popular, survey and recent papers. The evaluation was carried out as a case study with a subset of Information Retrieval (IR) related papers from the ACM Digital Library (ACM DL) collection.

The organization of this paper is as follows: Related works are discussed in section 2. The research coverage techniques are introduced in section 3. In sections 4 and 5, the experiment details along with the results are presented and discussed. The concluding remarks, study limitations and future works are presented in section 6.

## II. RELATED WORK

The traditional IR based mechanism used for finding research papers, improved with the advent of citations indexing systems [6] . Instead of just ranking the papers based on search keywords similarity, the citations data provided the leverage to improve the ranking so that important papers get better visibility. The metric *citation count* has largely remained the popular choice for ranking research papers. A study done to ascertain the ranking algorithm of Google Scholar indicated that highly cited papers dominate the search results on a consistent basis [7], thereby confirming to the *'Matthew Effect'* [8]. There is no proven correlation to show that highly cited papers have better coverage, however following the citations trail of these papers could lead to other important papers.

Alternatively, graph ranking algorithms such as PageRank [9] and HITS [10] have been employed in ranking research papers, since the in-links and out-links in citations network carry weightage in ascertaining the popularity of papers [11]. Some recent works have proposed enhanced versions of these

978-1-4673-9152-8/15/$31.00 ©2015 IEEE

algorithms incorporating textual [12] and temporal data [13] for providing improved results.

Studies conducted with the aim of generating reading lists for researchers, have also proposed methods for measuring research coverage of papers. The Collaborative Filtering (CF) technique of Recommender Systems (RS), has been found to produce a better starting list of citations for novice researchers, than traditional content-based filtering methods [14]. In a later study, CF techniques weighted with paper importance scores from graph ranking algorithms such as PageRank, HITS and SALSA provided even better results [4]. Bae et al. [3] proposed an algorithm based on random walk with restart (RWR) to measure the seminality score of papers. The method incorporates inter-paper similarity scores as well.

The earlier approaches in research coverage measurement have been largely based on the references and citations network while graph ranking methods have been predominantly used to ascertain the paper importance. However, the utility of these methods have not been tested yet in the interdisciplinary research sphere. Interestingly, none of the methods use author-specified keywords from research papers to build the references and citations network. We believe this data could help in identifying diverse set of papers. Secondly, the coverage score is best calculated separately so that the pre-computed values could be fed into IR or RS implementations for ranking or filtering purposes.

## III. TECHNIQUES FOR MEASURING RESEARCH COVERAGE

### A. State-of-the-Art Techniques

#### 1) HITS-based Coverage (HC)

Graph ranking algorithms provide the required mechanism for measuring the importance of nodes in graphs by using in-links and out-links. They have been used in earlier studies for measuring the importance of research papers in citations network. Among these algorithms, the HITS algorithm [10] has provided better results in both generic digital libraries context [11] and scientific paper recommender systems context [4]. When applied to the domain of scientific papers, the hub node is the paper which cites important papers whereas the authority node is the paper which is cited by other papers. In this paper, we will be using this technique for benchmarking purposes.

### B. Proposed Techniques

We propose two novel coverage measuring techniques based on author-specified keywords in research papers. The rationale for utilizing this data is that these keywords represent the central topics addressed in the paper. Most publications allow authors to specify a maximum of five keywords for each paper. These keywords not only provide scope for classifying the paper as detailed and broader topics but also an opportunity to identify inter-disciplinary topics. The two keywords-based coverage techniques are described as follows:-

#### 1) Topical and Peripheral Coverage (TPC)

The Topical and Peripheral Coverage (TPC) technique is meant at utilizing all the keywords provided by the authors for a research paper, in research coverage measurement. The measurement technique starts with identifying the keywords $K$ provided for a paper $P_i$, followed by extracting all the papers in the corpus which have the keywords in K. This extracted set of papers becomes the base set $P^k$. In the next step, extraction the references list $reflist_i$ and citations list $citelist_i$ of $P_i$ is performed. The coverage score is measured by counting the number of papers from $reflist_i$ and $citelist_i$ that are present in $P^k$. Alternatively, the coverage score can be also calculated as ratio score by dividing the combined count of $reflist_i$ and $citelist_i$ by count of papers in $P^k$. The conceptual model for this technique is provided in sub-section 3.

#### 2) Topical Coverage (TC)

Since the TPC technique takes into account all the author-specified keywords of a research paper, it is expected to identify research papers for a diverse set of sub-topics. This approach is also suited for inter-disciplinary topics with the precondition that authors from different disciplines have used common keywords in their research papers. The second coverage technique Topical Coverage (TC) is conceptualized at a single research topic level. The measurement technique starts with identifying a topic T. Papers having the topic T as author-specified keyword are retrieved from the corpus to form the set $P^T$. In the next step, we extract the references list $reflist_i$ and citations list $citelist_i$ of a particular paper $P_i$ which has topic T as author-specified keyword. The coverage score is measured by calculating the count of papers from $reflist_i$ and $citelist_i$ that are present in $P^T$. Alternatively, the coverage score can be also calculated as ratio score by dividing the combined count of $reflist_i$ and $citelist_i$ by count of papers in $P^T$. The conceptual model for this technique is also provided in the next sub-section. A pictorial representation of the network space in these techniques is illustrated in Figure 1.

#### 3) Conceptual Model for TPC and TC Techniques

$P_n$ : list of papers in the sample set of the corpus about the topic $T$

For each $P_i\,(1 \le i \le n)$, $reflist_i$ : $\{P_j \mid P_i \text{ references } P_j\}$

For each $P_i\,(1 \le i \le n)$, $citlist_i$ : $\{P_j \mid P_i \text{ is cited by } P_j\}$

$K$ : list of all author specified keywords in $P_n$

$P^k$ : list of all papers in the corpus having any $k$ in $K$ as author specified keyword

$P^T$: list of all papers in the corpus having $T$ as author specified keyword

$TPC$ for $P_i$ : $\left|reflist_i \in P^k\right| + \left|\ citlist_i \in P^k\right|$

$TC$ for $P_i$ : $\left|reflist_i \in P^T\right| + \left|\ citlist_i \in P^T\right|$

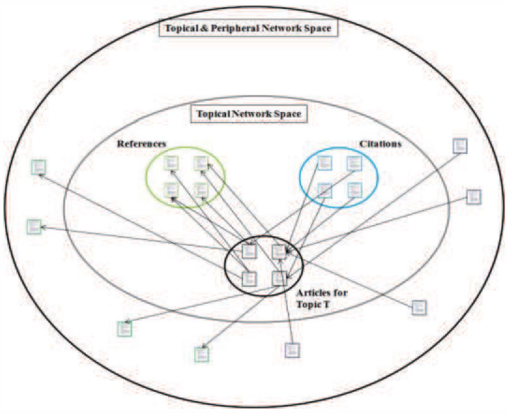


Fig. 1. Paper Network Space for TPC and TC Techniques

### C. Evaluation Criteria for Research Coverage

The evaluation criteria for a measure such as research coverage need to be selected based on its application context. The measure finds its primary use in academic digital libraries, databases and search engines in both IR and RS contexts. Particularly, it is most suited for the task of generating a reading list of research papers since a reading list is supposed to include papers that provide an overview of the research done in a particular research area. Even for this particular task, the expectations could vary as per the experience level of the researcher. Therefore, multiple evaluation criteria are required. We propose the use of four key evaluation criteria that measure the inclusivity extent of diverse papers, survey papers, popular papers and recent papers in the final list of papers with top coverage scores.

#### 1) Diverse papers inclusivity

A reading list is supposed to contain papers that cover sub-topics of the main research topic so that the researcher gets to know about the different research trends and problems in the given area. Therefore, diversity is a necessary evaluation criterion. In other words, a diverse set of papers have a high level of novelty as papers deal with different sub-topics. In traditional RS studies, measurement of novelty and diversity metrics is based on distance between the recommended resources [15]. Longer distance means high diversity while shorter distance means low diversity. In the scientific research domain, since there is usage of references and citations network for the papers, subgraph properties could be used to infer the level of diversity. If $G(V,E)$ is a graph built with references and citations of papers about a particular topic, $G1(V1,E1)$ is a subgraph built with just the recommended papers from the coverage measurement technique. The number of edges $E1$ from $G1$ is an indication of level of diversity in the recommended papers. If there is more number of edges, it means there are many inter-referencing connections between the papers, thus implying a less diverse recommendation list and vice-verse for high diversity.

#### 2) Survey papers inclusivity

Literature survey and systematic literature review papers provide an overview of the existing research performed in the particular research area. Researchers generally read this type of papers as a starting point in their literature review. These papers can be considered to have high research coverage as they cite most of the research papers in the particular research topic. Therefore, the coverage measurement technique is expected to identify such papers with high coverage values. The criterion can be measured in a straightforward way by counting the number of survey papers in the final recommendation list.

#### 3) Popular papers inclusivity

Popular papers are papers which have a seminal status in the particular research area. Generally, these papers have very high citation counts, indicating their popularity. These papers can be considered to have high research coverage as they are cited by many papers. We use the term 'coverage' to indicate the ability of a paper to lead the reader to other interesting set of papers both in backward direction (using references) and forward direction (using citation). In this sense, seminal papers have high coverage since they lead to many other papers by following citations. Setting the citation count threshold is the key factor for this criterion, as it is dependent on the discipline. Similar to survey papers, this criterion can be measured by counting the number of papers in the final recommendation list, which are above the citation count threshold.

#### 4) Recent papers inclusivity

One of the observed issues with the existing methods used for building reading lists is the neglect of recently published papers. Highly cited seminal papers are regarded as the key papers with high coverage values. On the other hand, few recently published papers that cite many other important papers should also be considered as papers with decent level of coverage. Therefore, inclusion of recent papers should also be a key requirement for a coverage measurement technique. Similar to popular papers, the key factor for this criterion is setting the threshold value for identifying the recent papers. In general, papers published three years prior to the current date can be considered to be recent. The final measurement is the count of such recent papers in the final recommendation list of papers with high coverage values.

## IV. EXPERIMENT

In order to analyze the performance of the coverage techniques at a detailed level, an experiment was conducted with a restricted set of papers which are about a single broad topic. *Information Retrieval (IR)* was chosen as the research topic for this study as it is a broad area with 71 sub-topics in the ACM 2012 taxonomy [16]. From a base extract of the ACM DL covering papers published between 1951 and 2010, 1473 IR papers were initially shortlisted. 976 papers were finally chosen as the sample set as these papers had complete information in various attributes such as keywords, abstract, bibliography and full text. 21,243 references and citations data of the sample set papers were used to build the references and citations network. The coverage values for the three techniques were measured for all the sample set papers and the top 20 papers were shortlisted for comparison purposes since most users tend to select resources within the top 20 results that are displayed within the two pages of the search results [17]. The HITS scores for the papers was calculated using the JUNG java library [18] while the TPC and TC measurement was performed using custom java code with data retrieved from MySQL database. The generation of subgraphs for the diverse papers inclusivity evaluation was also performed using the JUNG library. For the recent papers inclusivity criterion, papers published between 2008 and 2010 were considered as recent papers for this experiment as the ACM DL collection was obtained as of 2011. The citation count is calculated based on the citations within the ACM DL extract used in the current study.

## V. RESULTS & DISCUSSION

The experiment results for the four evaluation criteria are provided in Table I for the three techniques. The top 20 papers from the three techniques are displayed in Table II, III and IV respectively where the columns Ref. Ct and Cit. Ct refer to references count and citations count of the papers. In these

three tables, survey papers and also recent papers published between 2008 and 2010 are highlighted in italics.

TABLE I. CRITERIA COUNTS FOR THE THREE TECHNIQUES

| Criteria Count | HC | TPC | TC |
|---|---|---|---|
| No. of Edges in Subgraph (Diversity) | 14 | 5 | 18 |
| No. of Literature Survey Papers | 6 | 1 | 3 |
| No. of Papers with Citation Count above 100 | 9 | 16 | 16 |
| No. of Recent Papers | 3 | 7 | 5 |

TABLE II. TOP 20 PAPERS FROM HC TECHNIQUE

| Rank | Title (year) | Ref. Ct | Cit. Ct |
|---|---|---|---|
| 1 | *Inverted files for text search engines (2006)* | 205 | 900 |
| 2 | *Information retrieval on the web (2000)* | 231 | 735 |
| 3 | *Information storage and retrieval: a survey and functional description (1977)* | 238 | 19 |
| 4 | *Web mining research: a survey (2000)* | 128 | 1742 |
| 5 | Information science in a Ph.D. computer science program (1969) | 200 | 9 |
| 6 | Information retrieval on the semantic web (2002) | 28 | 225 |
| 7 | Building efficient and effective metasearch engines (2002) | 83 | 413 |
| 8 | Refinement of TF-IDF schemes for web pages using their hyperlinked neighboring pages (2003) | 26 | 68 |
| 9 | *A survey of Web clustering engines (2009)* | 113 | 277 |
| 10 | Collection synthesis (2002) | 47 | 64 |
| 11 | Using information scent to model user information needs and actions and the Web (2001) | 26 | 401 |
| 12 | Social network document ranking (*2010*) | 36 | 20 |
| 13 | An indexing model of HTML documents (2003) | 25 | 12 |
| 14 | A Hybrid Technique for English-Chinese Cross Language Information Retrieval (*2008*) | 49 | 20 |
| 15 | Efficient on-line index maintenance for dynamic text collections by using dynamic balancing tree (2007) | 15 | 28 |
| 16 | Superimposing codes representing hierarchical information in web directories (2001) | 16 | 10 |
| 17 | Information filtering and information retrieval: two sides of the same coin? (1992) | 33 | 1664 |
| 18 | *Is this document relevant? probably?: a survey of probabilistic models in information retrieval (1998)* | 90 | 270 |
| 19 | Microsearch: A search engine for embedded devices used in pervasive computing (*2010*) | 36 | 21 |
| 20 | Merging techniques for performing data fusion on the web (2001) | 27 | 30 |

TABLE III. TOP 20 PAPERS FROM TPC TECHNIQUE

| Rank | Title (year) | Ref. Ct | Cit. Ct |
|---|---|---|---|
| 1 | Information retrieval on the semantic web (2002) | 28 | 225 |
| 2 | Methods and metrics for cold-start recommendations (2002) | 31 | 1024 |
| 3 | StuffI've seen:a system for personal information retrieval and re-use (2003) | 32 | 813 |
| 4 | Designing a digital library for young children (2001) | 20 | 156 |
| 5 | *Information retrieval on the web (2000)* | 231 | 735 |
| 6 | A cluster-based resampling method for pseudo-relevance feedback (*2008*) | 35 | 142 |
| 7 | A case for interaction: a study of interactive information retrieval behavior and effectiveness (1996) | 13 | 398 |
| 8 | Query dependent pseudo-relevance feedback based on wikipedia (*2009*) | 34 | 135 |
| 9 | Scatter/gather browsing communicates the topic structure of a very large text collection (1996) | 11 | 245 |
| 10 | Answer Garden 2: merging organizational memory with collaborative help (1996) | 25 | 399 |
| 11 | Using information scent to model user information needs and actions and the Web (2001) | 26 | 401 |
| 12 | A knowledge-based search engine powered by wikipedia (2007) | 21 | 153 |
| 13 | Evaluation over thousands of queries (*2008*) | 21 | 79 |
| 14 | Sources of evidence for vertical selection (*2009*) | 20 | 126 |
| 15 | Interactive textbook and interactive Venn diagram: natural and intuitive interfaces on augmented desk system | 21 | 117 |

| | | | |
|---|---|---|---|
| | (2000) | | |
| 16 | Categorizing web queries according to geographical locality (2003) | 21 | 171 |
| 17 | Automatic query generation for patent search (*2009*) | 12 | 67 |
| 18 | Probabilistic query expansion using query logs (2002) | 18 | 415 |
| 19 | Extending average precision to graded relevance judgments (*2010*) | 27 | 51 |
| 20 | Learning in a pairwise term-term proximity framework for information retrieval (*2009*) | 19 | 47 |

TABLE IV. TOP 20 PAPERS FROM TC TECHNIQUE

| Rank | Title (year) | Ref. Ct | Cit. Ct |
|---|---|---|---|
| 1 | A vector space model for automatic indexing (1975) | 7 | 5780 |
| 2 | *Information retrieval on the web (2000)* | 231 | 735 |
| 3 | Information retrieval on the semantic web (2002) | 28 | 225 |
| 4 | Cross-language information retrieval based on parallel texts and automatic mining of parallel texts from the Web (1999) | 19 | 320 |
| 5 | A Hybrid Technique for English-Chinese Cross Language Information Retrieval (*2008*) | 49 | 20 |
| 6 | What the query told the link: the integration of hypertext and information retrieval (1997) | 37 | 117 |
| 7 | Geospatial mapping and navigation of the web (2001) | 22 | 282 |
| 8 | Stuff I've seen: a system for personal information retrieval and re-use (2003) | 32 | 813 |
| 9 | *Is this document relevant? Probably? :a survey of probabilistic models in information retrieval (1998)* | 90 | 270 |
| 10 | *Inverted files for text search engines (2006)* | 205 | 900 |
| 11 | A case for interaction: a study of interactive information retrieval behavior and effectiveness (1996) | 13 | 398 |
| 12 | Biterm language models for document retrieval (2002) | 10 | 107 |
| 13 | Discovering key concepts in verbose queries (*2008*) | 34 | 197 |
| 14 | Statistical transliteration for english-arabic cross language information retrieval (2003) | 22 | 156 |
| 15 | Categorizing web queries according to geographical locality (2003) | 21 | 171 |
| 16 | Query term disambiguation for Web cross-language information retrieval using a search engine (2000) | 19 | 60 |
| 17 | Modeling and visualizing geo-sensitive queries based on user clicks (*2008*) | 14 | 16 |
| 18 | Integration of news content into web results (*2009*) | 34 | 105 |
| 19 | An improved markov random field model for supporting verbose queries (*2009*) | 32 | 47 |
| 20 | Extended Boolean information retrieval (1983) | 28 | 1048 |

For the diverse papers inclusivity criterion, the least number of edges was for the subgraph formed with the top 20 papers from the TPC technique *(n=5)* (refer Table I*)* while the highest was for TC technique *(n=18)*. The inference is that TPC technique is better at providing a diverse set of papers about different sub-topics in IR. This result is validated by the inclusion of non-IR keywords in the base set formation for TPC technique, thereby increasing the scope for a heterogeneous mix of papers. On the other hand, TC and HC subgraphs had more connections as the top 20 papers are of IR-exclusive nature. Literature survey/review papers inclusion was clearly the highest for the HC technique *(n=6)*. The finding vindicates the working mechanism of HITS algorithm as it is expected to find hub nodes which point to many authority nodes. The references count is evidently higher for survey papers as these papers review the state-of-the-art in the particular research area. The number of survey papers in the top 20 papers of TC (*n=3*) and TPC *(n=1)* was comparatively lower as these techniques do not necessarily give exclusive importance to papers with high reference counts, at the conceptual level. The survey paper titles are highlighted in italics in the Tables II, III and IV.

For the popular papers inclusivity criterion, both TPC and TC techniques were good at providing high coverage values to highly cited papers (*n=16*) while interestingly for HC technique, the count was lower (*n=9*) even though the technique identified more number of survey papers. This is due to the tendency of the HC technique in giving higher coverage scores to papers with high reference counts. In the case of recent papers inclusivity, the least number of recent articles *(n=3)* was for HC technique while the count was better for TPC *(n=7)* and TC *(n=5)*. None of the techniques have temporal preferences in the conceptual model, therefore the results for this criterion need to be validated with papers for different research topics. It is assumed that if a recent paper has a higher references count, it is expected to get good coverage scores with the HC technique. In the current experiment, only one recent survey paper 'A survey of Web clustering engines' published in 2009 received higher coverage score (rank 9 in HC top 20 papers).

TABLE V. RANKS OF THE THREE TECHNIQUES

| Rank | HC | TPC | TC |
|---|---|---|---|
| No. of Edges in Subgraph (Diversity) | 2 | 1 | 3 |

| | | | |
|---|---|---|---|
| No. of Literature Survey Papers | 1 | 3 | 2 |
| No. of Papers with Citation Count above 100 | 3 | 1 | 1 |
| No. of Recent Papers | 3 | 1 | 2 |

Based on the performance of the three techniques with the four evaluation criteria, ranks have been assigned in Table V. TPC technique performs the best for diverse, popular and recent papers while HC technique performs the best for survey papers. The coverage metric can be implemented in citations databases, digital libraries and academic search engines for ranking research papers. The most specific usage of this metric is in ranking papers towards building a reading list of papers for literature review. Even though, it could be argued that the reading list generation technique should be standardized for all researchers, special preferences could also be given. For instance, when recommending research papers to novice researchers, a substantial amount of survey articles could be recommended. Alternatively for experienced researchers, a diverse set of papers about different sub-topics of a particular research area could be recommended. Therefore, the TPC technique and HC technique could be employed as per the specific needs of the users, for providing better results.

## VI. CONCLUSION AND FUTURE WORK

In this paper, we have proposed two novel research coverage measurement techniques topical coverage (TC) and topical and peripheral coverage (TPC) for scientific papers. These techniques are based on author-specified keywords in research papers. These two techniques were evaluated with an existing technique which makes use of HITS algorithm for scoring the importance of a paper in scientific networks. A simple experiment was performed with a set of papers addressing a single broad topic in the ACM DL extract. The evaluation was carried out with four key inclusivity criteria. Results showed that the proposed techniques TPC and TC were better at providing higher coverage scores to diverse, recent and popular papers. The HITS approach was better at scoring literature survey papers. Since the experiment was conducted in the form of a case study, the data was limited to just one research topic. Therefore the results are to be validated with more research topics. Nevertheless, the underlying reasons for most of the current results have been identified in this study. The proposed techniques provide scope for implementation in academic search systems irrespective of the type of retrieval technology used.

Coverage scores are apt for ranking papers for the task of generating a reading list of papers on a particular research topic. We are currently using the TPC technique for calculating coverage scores for a bigger set of papers from an ACM DL collection. The intention is to use the coverage score for ranking research papers which are the output recommendations of a hybrid scientific paper recommender system [19] which is under the evaluation phase. We are currently performing offline evaluations of the recommender system for the reading list generation task. Future user studies will also be conducted to ascertain user satisfaction of the generated recommendations in the system.

## ACKNOWLEDGEMENTS

We wish to thank ACM for providing us with an extract of the ACM DL indexed papers. This research is 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.

## REFERENCES

[1] S. Lehmann, B. Lautrup, and A. D. Jackson, "Citation Networks in High Energy Physics," *Phys. Rev. E*, vol. 68, no. 2, 2003.

[2] F. Ricci, L. Rokach, and B. Shapira, *Introduction to recommender systems handbook*. 2011.

[3] D.-H. Bae, S.-M. Hwang, S.-W. Kim, and C. Faloutsos, "On Constructing Seminal Paper Genealogy," *IEEE Trans. Cybern.*, vol. 44, no. 1, pp. 54–65, Mar. 2014.

[4] M. D. Ekstrand, P. Kannan, J. a. Stemper, J. T. Butler, J. a. Konstan, and J. T. Riedl, "Automatically Building Research Reading Lists," in *Proceedings of the fourth ACM conference on Recommender systems - RecSys '10*, 2010, p. 159.

[5] J. G. Jardine, "Automatically generating reading lists," 2014.

[6] C. L. Giles, K. D. Bollacker, and S. Lawrence, "CiteSeer: An Automatic Citation Indexing System," in *Proceedings of the third ACM conference on Digital libraries*, 1998, pp. 89–98.

[7] J. Beel and B. Gipp, "Google Scholar's Ranking Algorithm: The Impact of Citation Counts (An Empirical Study )," in *Third International Conference on Research Challenges in Information Science, 2009. RCIS 2009.*, 2009, no. April, pp. 439–446.

[8] R. K. Merton, "The Matthew Effect in Science," *Science*, vol. 159, no. 3810. pp. 56–63, 1968.

[9] L. Page, S. Brin, R. Motwami, and T. Winograd, "The PageRank Citation Ranking: Bringing Order to the Web," 1999.

[10] J. M. Kleinberg, "Hubs , Authorities , and Communities," *ACM Comput. Surv.*, vol. 31, no. 4, 1999.

[11] Y. Liu and Y. Lin, "Supervised HITS algorithm for MEDLINE citation ranking," in *Proceedings of the 7th IEEE International Conference on Bioinformatics and Bioengineering, BIBE*, 2007, pp. 1323–1327.

[12] J. Jardine and S. Teufel, "Topical PageRank : A Model of Scientific Expertise for Bibliographic Search," in *Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics*, 2014, pp. 501–510.

[13] A. P. Singh, K. Shubhankar, and V. Pudi, "An Efficient Algorithm for Ranking Research Papers Based on Citation Network," in *2011 3rd Conference on Data Mining and Optimization (DMO)*, 2011, pp. 88–95.

[14] S. M. Mcnee, *Meeting User Information Needs in Recommender Systems*. Proquest, 2006.

[15] P. Castells, S. Vargas, and J. Wang, "Novelty and Diversity Metrics for Recommender Systems: Choice, Discovery and Relevance," in *Proceedings of International Workshop on Diversity in Document Retrieval (DDR)*, 2009, pp. 29–37.

[16] ACM, "The 2012 ACM Computing Classification System," 2015. [Online]. Available: https://www.acm.org/about/class/2012. [Accessed: 14-Aug-2015].

[17] A. J. A. M. Van Deursen and J. A. G. M. Van Dijk, "Using the Internet: Skill related problems in users' online behavior," *Interact. Comput.*, vol. 21, no. 5, pp. 393–402, 2009.

[18] JUNG, "JUNG - Java Universal Network/Graph Framework," 2015. [Online]. Available: http://jung.sourceforge.net/. [Accessed: 14-Aug-2015].

[19] A. Sesagiri Raamkumar, S. Foo, and N. Pang, "Rec4LRW – Scientific Paper Recommender System for Literature Review and Writing," in *Proceedings of the 6th International Conference on Applications of Digital Information and Web Technologies*, 2015, pp. 106–120.