A Topic-based Reviewer Assignment System

Ngai Meng Kou *1 Leong Hou U *2 Nikos Mamoulis #3 Yuhong Li *4 Ye Li *5 Zhiguo Gong *6

*Department of Computer and Information Science, University of Macau, Macau SAR

#Department of Computer Science, University of Hong Kong, Hong Kong SAR

{1yb27406,2ryanlhu,4yb27407,5yb47438,6fstzgg}@umac.mo 3nikos@cs.hku.hk

ABSTRACT

Peer reviewing is a widely accepted mechanism for assessing the quality of submitted articles to scientific conferences or journals. Conference management systems (CMS) are used by conference organizers to invite appropriate reviewers and assign them to submitted papers. Typical CMS rely on paper bids entered by the reviewers and apply simple matching algorithms to compute the paper assignment. In this paper, we demonstrate our Reviewer Assignment System (RAS), which has advanced features compared to broadly used CMSs. First, RAS automatically extracts the profiles of reviewers and submissions in the form of topic vectors. These profiles can be used to automatically assign reviewers to papers without relying on a bidding process, which can be tedious and error-prone. Second, besides supporting classic assignment models (e.g., stable marriage and optimal assignment), RAS includes a recently published assignment model by our research group, which maximizes, for each paper, the coverage of its topics by the profiles of its reviewers. The features of the demonstration include (1) automatic extraction of paper and reviewer profiles, (2) assignment computation by different models, and (3) visualization of the results by different models, in order to assess their effectiveness.

1. INTRODUCTION

In conference management systems (CMS), one of the most important tasks is the fair assignment of submissions to reviewers. The most widely used CMS (i.e., Conference Management Toolkit[1] and EasyChair[2]) assign the papers based on reviewer bidding preferences [3]. However, there are certain drawbacks of this methodology. First, some reviewers may be too lazy to go through the complete list of paper titles and abstracts, so their bidding may not precisely reflect their actual preferences. Second, the preferences of a reviewer are not essentially consistent with her expertise, therefore she may get papers for which she is not an expert.

[1]http://cmt.research.microsoft.com/cmt/

[2]http://www.easychair.org/

Proceedings of the VLDB Endowment, Vol. 8, No. 12

In this paper, we demonstrate our Reviewer Assignment System (RAS), which supports the *automatic* assignment of papers to reviewers (without the need of a bidding phase). RAS employs a similarity model between the expertise of reviewers and the topics of the submissions. RAS includes the following features: (1) profile extraction and profile tuning interface, (2) assignment computation by different models, (3) visualization and quality assessment of assignments.

Given a program committee (PC), the offline topic extraction model employed by RAS automatically extracts a set of research topics and the *profiles* (i.e., topic vectors) of PC members from their publication records. Based on the extracted topics, RAS generates the profiles of submissions by an Expectation-Maximization (EM) process. The profile of a reviewer (paper) is a weighted vector, which captures the expertise (relevance) of the reviewer (paper) to each of the research topics. RAS also offers a user interface that allows reviewers to manually change the weights in their profiles.

The appropriateness of pairing reviewer r to submission p can be measured by applying a similarity function $c(\overrightarrow{r}, \overrightarrow{p})$ (e.g., histogram intersection) on their profile vectors $\overrightarrow{r}$ and $\overrightarrow{p}$, respectively. Given a global assignment objective (for all papers and reviewers) and load balancing constraints (e.g., each paper gets three reviews and the review load is equally distributed to reviewers), our target is to find an assignment $\mathbb{A}$ (i.e., a set of reviewer-paper pairs) such that $\mathbb{A}$ maximizes the objective. For instance, the *optimal assignment* [5] $\mathbb{A}$ maximizes the overall assignment quality, i.e., $\text{argmax}_{\mathbb{A}} \sum_{(r,p)\in\mathbb{A}} c(\overrightarrow{r}, \overrightarrow{p})$. However, pairings are only individually considered in this objective; thus, an interdisciplinary paper could be assigned to a group of reviewers, who are expert to only one of the paper topics.

In our recent work [4], we argue that the quality of reviewers for a paper should be measured based on *how well the topics of the paper are covered by the expertise of the reviewers.* Thereby, we propose to aggregate the topic vectors of a reviewer group $\{r_i, \ldots, r_j\}$ into a *group vector* $\overrightarrow{g}$ and measure the review quality of a paper by $c(\overrightarrow{g}, \overrightarrow{p})$. However, finding the optimal assignment based on this definition becomes hard, as we search for combinations of reviewers for each paper which collectively maximize the scoring function. As shown in [4], a special case of this problem can be reduced to the maximum coverage problem [1], which is NP-hard. Thus, in [4], we propose a greedy algorithm that computes a 1/2-approximate solution.

The goal of this demonstration is to allow users, who could be potential conference organizers or reviewers, to assess the quality of (i) automatic profile extraction for reviewers and papers and (ii) alternative assignment algorithms based on different objectives. Our RAS includes an interface, which allows users to simulate the assignment process of a conference management system. The user can select from a set of potential reviewers to be the PC members or manually include additional reviewers. The user can also choose from a pool of papers published in database conferences to be the submissions. The assignment is then computed based on a set of assignment algorithms (i.e., optimal assignment, stable marriage, and our group-based assignment) and assignment parameters (e.g., the workload of reviewers). RAS also includes an interface that simulates a journal editorial system, where the user (journal editor) is looking for qualified reviewers for a single submission. This problem can be viewed as a special case of our group-based assignment and it is possible to solve it exactly and fast, with the help of a branch-and-bound algorithm [4]. For journal paper assignment, RAS computes and returns the top-k reviewer groups to a given paper, giving flexibility to the ultimate selection by the journal editor. Finally, RAS offers an easy-to-use interface to navigate the assignment results computed by the various assignment objectives. Thus, attendants view the details of each assignment group by clicking on the paper title and judge the quality of the assignment result.

The rest of our paper is organized as follows. We first give the global picture of RAS in Section 2. Then, we briefly describe its main components in Sections 3 and 4. The demonstration scenarios are discussed in Section 5.

2. SYSTEM OVERVIEW

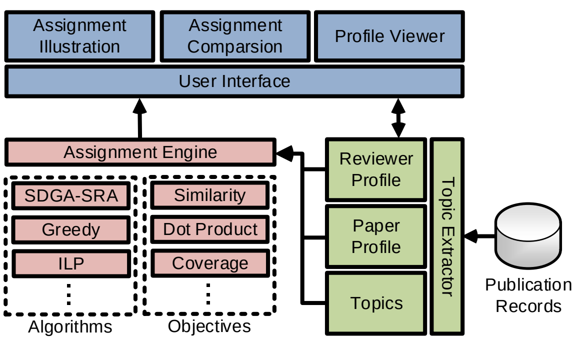


Figure 1: System overview

Figure 1 shows an overview of our RAS, which consists of a User Interface, an Assignment Engine, and a Topic Extractor. The Topic Extractor applies an Author Topic Model (ATM) [7] on the abstracts of the publications of all candidate reviewers. The Assignment Engine offers a set of assignment algorithms and objectives. The web interface allows users to screen the assignment results and to compare the assignments of different assignment objectives.

RAS simulates the assignment processes in a conference management system (CMS) or a journal editorial system (JES). The basic workflow of a CMS includes the following steps. First, the system asks the user (i.e., the PC chair) to form the program committee and to select the set of paper submissions.[3] The Topic Extractor then defines a profile (topic vector) for each reviewer and paper. The user can browse the reviewer profiles by clicking on their names and fine-tune them as necessary. After setting the assignment parameters (e.g., the workload of reviewers, the assignment model and algorithm), the Assignment Engine computes the assignment and shows it. The simulation of a JES is similar to that of a CMS, but it only allows to select one paper submission. In this case, the system computes and returns the top-k reviewer groups for the input submission.

3. TOPIC EXTRACTOR

We represent the profiles of reviewers and papers by T-dimensional vectors of topics $\mathbb{T}$ (i.e. subjects), where each topic $t \in \mathbb{T}$ is a distribution of words taken from the corpus (titles and abstracts of published papers by the reviewers). Table 1 shows the top-5 words of 3 topics selected from our test dataset. From these, we can understand that the 3 topics are motif discovery (t_1), xml queries (t_2), and database security (t_3). To extract the topic set $\mathbb{T}$, we adapt ATM [7], a statistical model specialized to discover topics from titles and abstracts of academic papers. In the same statistical process, the profiles of reviewers $\mathbb{R} = \{\overrightarrow{r_1}, \ldots, \overrightarrow{r_n}\}$ are generated based on the topic set $\mathbb{T}$. Given the topic set $\mathbb{T}$, the profile of a paper $\overrightarrow{p_i}$ can be computed by the Expectation Maximization (EM) algorithm [9] as follows.

$$\overrightarrow{p} = \underset{\overrightarrow{p}}{\operatorname{argmax}} \prod_{i=1}^{W_p} \sum_{j=1}^{T} \mathcal{P}(w_i|t_j)\overrightarrow{p}[t_j] \qquad (1)$$

where W_p indicates the set of words in paper p and $\mathcal{P}(w_i|t_j)$ is the probability of word w_i in topic t_j.

t_1	mining, distance, indexing, classification, motif, ...
t_2	xquery, xpath, query, tree, structure, ...
t_3	access, control, security, database, attack, ...

Table 1: An example of topics

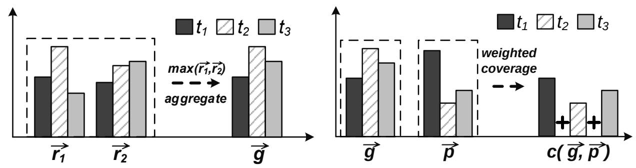
4. ASSIGNMENT ENGINE

The Assignment Engine is the core part of RAS. To conduct a fair assignment, one objective is to constrain the assignment workloads: *every paper should be assessed by* δ_p *reviewers and every reviewer should only review at most* δ_r *papers.* Workload constraints can be determined from the requirements (e.g., if each paper should get $\delta_p = 3$ reviews, to determine δ_p, we have to divide the total number of required reviews by the number of reviewers to determine δ_r).

Finding a good assignment given the workload constraints is clearly an optimization problem, which can be modeled as an optimal assignment problem [5] or a stable marriage problem [2]. However, the assignment pairs are only individually considered in these models; thus, an interdisciplinary paper could be assigned to a group of reviewers, who are expert to only one of the paper topics. This issue was also raised in a recent study [6], where the authors propose to measure the quality of the reviews on a paper based on topic

[3] For the sake of demonstration, candidate reviewers are chosen from the PCs of recent database conferences and retrieve the abstracts of their publications in 2000-2009 from [8]. We also imported a set of papers published in data management venues as submission candidates. Still we provide an interface that allows adding more submissions by manually typing their titles and abstracts.

set coverage of the reviewers on the paper. However, in [6], all topics of a paper are assumed to have identical importance, which is not typical and consistent with our extracted data by ATM and EM.



(a) Group topic vector (b) Weighted coverage

Figure 2: Group topic vector and weighted coverage

In our recent work [4], we overcome this problem by estimating the quality of a group assignment by its *weighted coverage* of paper topics instead of set coverage. We first define the expertise of a group $\vec{g}$ on topic t as the maximum expertise of any reviewer in the group on t. Formally,

$$\vec{g}[t] = \max_{r \in g} \vec{r}[t], \forall t \in \mathbb{T} \tag{2}$$

where $\vec{g}[t]$ and $\vec{r}[t]$ is the value of $\vec{g}$ and $\vec{r}$ on topic t. Figure 2(a) shows an example of a group with 2 reviewers. The topic vector of this group is $\vec{g} = (\vec{r_1}[1], \vec{r_1}[2], \vec{r_2}[3])$. Then, the coverage of paper $\vec{p}$ by $\vec{g}$ can be measured by:

$$c(\vec{g}, \vec{p}) = \frac{\sum_{t \in \mathbb{T}} \min\{\vec{g}[t], \vec{p}[t]\}}{\sum_{t \in \mathbb{T}} \vec{p}[t]}, \tag{3}$$

where the denominator normalizes $c(\vec{g}, \vec{p})$ to take values in [0,1]. Figure 2(b) illustrates the computation of weighted coverage of reviewer group $\vec{g}$ on paper $\vec{p}$.

Problem - WGRAP. Based on the above, we define the Weighted-coverage Group-based Reviewer Assignment Problem (WGRAP) in [4]. The objective of WGRAP is to find an assignment $\mathbb{A} \subseteq \mathbb{P} \times \mathbb{R}$ such that the sum of weighted coverages of all papers is maximized, subject to the group size and workload constraints. For the ease of discussion, we use $\mathbb{A}[x]$ to denote the assignment pair(s) of x. For instance, if $\mathbb{A} = \{(r_1, p_1), (r_2, p_1)\}$, then $\mathbb{A}[r_1] = \{(r_1, p_1)\}$ and $\mathbb{A}[p_1] = \mathbb{A}$. WGRAP is formally defined as follows:

$$\begin{array}{lll} \max & \sum_{p \in \mathbb{P}} c(\vec{g}, \vec{p}) & \\ \text{where} & g = \{r | (r,p) \in \mathbb{A}[p]\} & \\ s.t. & |\mathbb{A}[r]| \le \delta_r & \forall r \in \mathbb{R} \\ & |\mathbb{A}[p]| = \delta_p & \forall p \in \mathbb{P} \end{array}$$

Algorithm - SDGA-SRA. After showing the NP-hardness of WGRAP, in [4], we proposed an approximation algorithm, Stage Deepening Greedy Algorithm (SDGA), that achieves a 1/2-approximation ratio compared to the exact WGRAP solution. Furthermore, we enhance the assignment quality by a post-processing Stochastic Refinement Algorithm (SRA) which improves the SDGA result closer to the optimal. SRA's approximation quality is higher than 98.5% in some experiments.

Algorithm - BBA. The assignment process of a journal submission is a special case of a conference assignment; the editor is looking for δ_p qualified reviewers for just one submission. Even though finding the best group of reviewers based on weighted coverage is still NP-hard, for typical problem sizes, it is possible to compute the exact solution efficiently by our proposed a Branch-and-Bound Algorithm (BBA) [4]. Our experiments show that BBA can return the best reviewer group in one second for typical reviewer pools (e.g., 500 reviewers).

Objective functions. Besides the weighted coverage function (cf. Equation 3), our solutions (SDGA and BBA) seamlessly work with other submodular objective functions (e.g., dot-product). For the sake of comparison, we include two objective functions (weighted coverage and dot-product) in this demonstration.

5. DEMONSTRATION SCENARIOS

Our demo system (RAS) consists of four modules: (1) Reviewer Profiler, (2) Paper Profiler, (3) Conference Management System, and (4) Journal Editorial System.

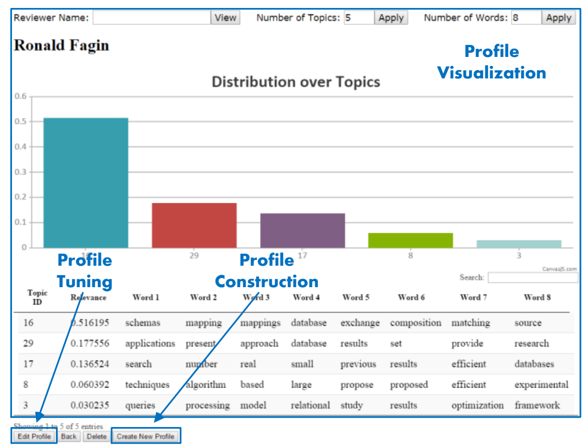


Figure 3: Reviewer profile management

Reviewer Profiler. It allows users to view the extracted profiles of reviewers by ATM. In addition, users can add a new reviewer profile or adjust existing ones. Figure 3 shows the interface, which consists of three parts, (1) profile visualization, (2) profile tuning, and (3) profile construction.

Paper Profiler. It allows users to view the profiles of existing papers and to also add new submissions. Given a new title and paper abstract, RAS can instantly generate the paper profile based on the topic set $\mathbb{T}$.

Conference Management System. CMS first asks users to form the program committee by selecting a set $\mathbb{R}$ of reviewers (determined by the Reviewer Profiler) and a set $\mathbb{P}$ of paper submissions (determined by the Paper Profiler). For the ease of use, the profile of the reviewers and papers are popped up if users double click on their names. After deciding the workload of papers δ_p, the system automatically determines the workload of reviewers δ_r to the minimum possible value (i.e., $\delta_r = \lceil |\mathbb{P}| \cdot \delta_p / |\mathbb{R}| \rceil$). CMS also allows users to select an objective function (either weighted coverage or dot product). The Assignment Engine then proceeds to compute the results by several assignment meth-

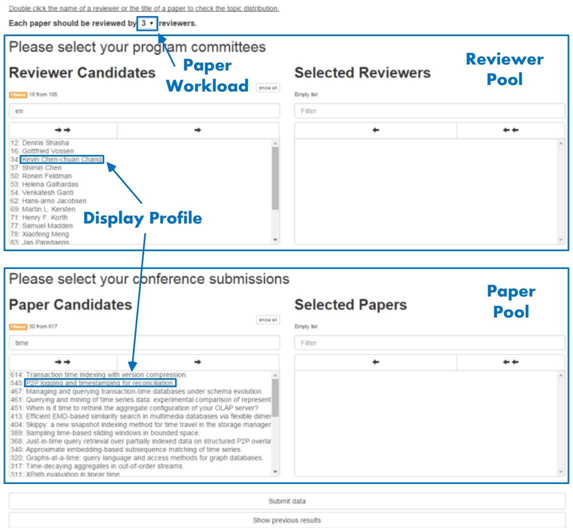


Figure 4: Conference Management System

ods, including optimal assignment [5], stable marriage [2], Greedy [6], BRGG [4], and SDGA-SRA [4].

Figure 5: Result page and assignment comparison

The assignment results are listed in a table, where each row includes the reviewers assigned to a paper. Users can double click on a paper assignment and the topic profiles of the paper and the corresponding group of reviewers are visualized by a set of topic bars (cf. Figure 5).

Journal Editorial System. Figure 6 is the user interface of the Journal Editorial System. Users can select a paper from the paper pool. The assignment engine returns the best k groups of reviewers by running BBA. Note that we do not include any other (approximate) methods, as BBA already returns the best k groups of reviewers based on the given objective function and the profiles.

6. CONCLUSION

By this demonstration, we show the functionality of our advanced Reviewer Assignment System (RAS), which simulates the reviewer assignment processes for conferences and

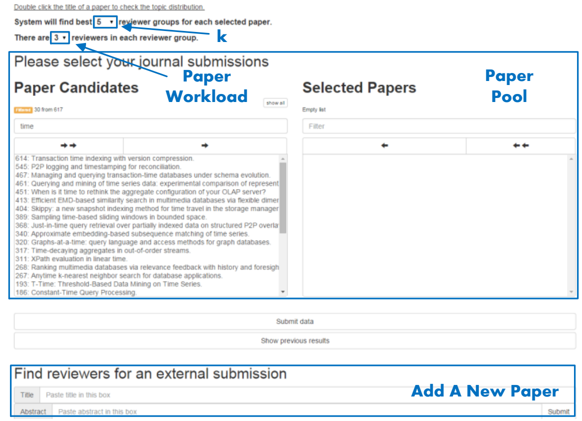


Figure 6: Journal Editorial System

journal submissions. The primary goals of the demonstration is to illustrate the benefits of (1) automatically extracting profiles of reviewers based on their publication records instead of asking reviewers to bid for papers and (2) using the group weighted coverage of the paper topics by the expertise of reviewers as the assignment objective instead of simply accumulating the quality of individual reviewer-paper pairs. The demonstration offers realistic features to the user (selection of reviewers and papers, fine-tuning of the reviewer expertise to the extracted topics, comparing assignments by alternative models and approximation algorithms). The feedback from VLDB attendants will definitely help us to move forward toward integrating our prototype system into real systems (e.g., CMT and EasyChair).

Acknowledgments. This project was supported by grant MYRG109(Y1-L3)-FST12-ULH and MYRG2014-00106-FST from UMAC RC, grant FDCT/106/2012/A3 from Macau FDCT and grant HKU 715413E from Hong Kong RGC.

7. REFERENCES

[1] U. Feige. A threshold of ln n for approximating set cover. *J. ACM*, 45(4):634–652, July 1998.

[2] D. Gale and L. S. Shapley. College admissions and the stability of marriage. *The American Mathematical Monthly*, 120(5):386–391, 2013.

[3] N. Garg, T. Kavitha, A. Kumar, K. Mehlhorn, and J. Mestre. Assigning papers to referees. *Algorithmica*, 58(1):119–136, 2010.

[4] N. M. Kou, L. H. U, N. Mamoulis, and Z. Gong. Weighted coverage based reviewer assignment. In *SIGMOD*, 2015.

[5] H. W. Kuhn. Variants of the hungarian method for assignment problems. *Naval Research Logistics Quarterly*, 3(4):253–258, 1956.

[6] C. Long, R. C.-W. Wong, Y. Peng, and L. Ye. On good and fair paper-reviewer assignment. In *ICDM*, 2013.

[7] M. Rosen-Zvi, T. Griffiths, M. Steyvers, and P. Smyth. The author-topic model for authors and documents. In *UAI*, 2004.

[8] J. Tang, J. Zhang, L. Yao, J. Li, L. Zhang, and Z. Su. Arnetminer: extraction and mining of academic social networks. In *KDD*, 2008.

[9] C. Zhai, A. Velivelli, and B. Yu. A cross-collection mixture model for comparative text mining. In *KDD*, 2004.