Interdisciplinary Journal of Information, Knowledge, and Management

An Official Publication of the Informing Science Institute InformingScience.org

IJIKM.org

Volume 13, 2018

PRATO: AN AUTOMATED TAXONOMY-BASED REVIEWER-PROPOSAL ASSIGNMENT SYSTEM

Basem Y. Alkazemi | Department of Computer Science, Umm Al-Qura University, Makkah, Saudi Arabia | bykazemi@uqu.edu.sa

ABSTRACT

Aim/Purpose: This paper reports our implementation of a prototype system, namely PRATO (Proposals Reviewers Automated Taxonomy-based Organization), for automatic assignment of proposals to reviewers based on categorized tracks and partial matching of reviewers' profiles of research interests against proposal keywords.

Background: The process of assigning reviewers to proposals tends to be a complicated task as it involves inspecting the matching between a given proposal and a reviewer based on different criteria. The situation becomes worse if one tries to automate this process, especially if a reviewer partially matches the domain of the paper at hand. Hence, a new controlled approach is required to facilitate the matching process.

Methodology: Proposals and reviewers are organized into categorized tracks as defined by a tree of hierarchical research domains which correspond to the university's colleges and departments. In addition, reviewers create their profiles of research interests (keywords) at the time of registration. Initial assignment is based on the matching of categorized sub-tracks of proposal and reviewer.

Where the proposal and a reviewer fall under different categories (sub-tracks), assignment is done based on partial matching of proposal content against reviewers' research interests. Jaccard similarity coefficient scores are calculated of proposal keywords and reviewers' profiles of research interest, and the reviewer with highest score is chosen.

Evaluation: The system was used to automate the process of proposal-reviewer assignment at the Umm Al-Qura University during the 2017-2018 funding cycle. The list of proposal-reviewer assignments generated by the system was sent to human experts for voting and subsequently to make final assignments accordingly. With expert votes and final decisions as evaluation criteria, data system-expert agreements (in terms of "accept" or "reject") were collected and analyzed by tallying

Accepted by Editor Tharrenos Bratitsi │ Received: May 28, 2018 │ Revised: July 26, August 27, October 1, October 11, 2018 │ Accepted: October 15, 2018.

Cite as: Alkazemi, B. Y. (2018). PRATO: An automated taxonomy-based reviewer-proposal assignment system. *Interdisciplinary Journal of Information, Knowledge, and Management*, *13,* 383-396. https://doi.org/10.28945/4134

frequencies and calculating rejection/acceptance ratios to assess the system's performance.

Contribution: This work helped the Deanship of Scientific Research (DSR), a funding agency at Umm Al-Qura University, in managing the process of reviewing proposals submitted for funding. We believe the work can also benefit any organizations or conferences to automate the assignment of papers to the most appropriate reviewers.

Findings: Our developed prototype, PRATO, showed a considerable impact on the entire process of reviewing proposals at DSR. It automated the assignment of proposals to reviewers and resulted in 56.7% correct assignments overall. This indicates that PRATO performed considerably well at this early stage of its development.

Recommendation for Practitioners: It is important for funding agencies and publishers to automate reviewing process to obtain better reviewing quality in a timely manner.

Recommendations for Researchers: This work highlighted a new methodology to tackle the proposal-reviewer assignment task in an automated manner. More evaluation might be needed with consideration of different categories, especially for partially matched candidates.

Impact on Society: The new methodology and knowledge about factors influencing the implementation of automated proposal-reviewing systems will help funding agencies and publishers to improve the quality of their internal processes.

Future Research: In the future, we plan to examine PRATO's performance on different classification schemes where specialty areas can be represented in graphs rather than trees. With graph representation, the scope for reviewer selection can be widened to include more general fields of specialty. Moreover, we will try to record the reasons for rejection to identify accurately whether the rejection was due to improper assignment or other reasons.

Keywords: reviewers matching, taxonomy-based, Jaccard index, proposals auto assignment

INTRODUCTION

Selecting the most appropriate reviewers for proposals is a widely known challenging task that requires careful consideration and knowledge of the involved domains. Many factors must be taken into consideration prior to selection: the specialty of the reviewer and their research interest, the number of assigned proposals to a certain reviewer, and transparency and integrity in selecting reviewers. These factors require a sophisticated system that can manage the assignment of proposals to reviewers. Moreover, the assignment process should be done automatically to speed up the process and, more importantly, to eliminate human bias as much as possible.

Despite the availability of several solutions to address the problem of reviewer-proposal (or paper) matching (Wang, Chen, & Miao, 2008), none has been thoroughly customized for utilization by other parties, and they mostly rely on manual assignment processes, which tend to be very time-consuming and solely dependent on human judgment. Furthermore, in addition to making appropriate decisions for perfect matches, the automated solution needs to account for partially matching reviewers with proposals (or papers) in cases where a perfect match is not found.

This paper documents a prototype system developed for the Deanship of Scientific Research (DSR) at Umm Al-Qura University (UQU) to help assign reviewers to research proposals with consideration of the requirements and regulations involved. In this paper, we report a methodology to address the problem of proposal-reviewer assignments, which we implemented at UQU to improve the reviewing process within the DSR and to obtain better review quality with minimal human intervention.

Special attention was paid to addressing the problem of assigning proposals to reviewers based on partial matches or matching those belonging to different clusters. Thus, the PRATO tool was developed to automate the proposal-reviewer assignment process, especially for partially-matched candidates who fall within different categories in the taxonomy.

The remainder of this paper is organized as follows. The next section describes related work within the domain. After that, a third section provides an overview of Umm Al-Qura University's (UQU) funding scheme. The fourth section discusses the overall specifications of PRATO system. The evaluation of the prototype and experimental results are highlighted in the fifth section of this paper. Finally, the sixth section concludes the paper with brief discussion of intended future work.

RELATED WORK

Several conference management systems, such as Easy Chair, CMOS, and Microsoft Conference Management Tool (CMT), are commonly used. However, these systems manage purely the proposal submission process, manual assignment to reviewers, and tracking of the review process. Selection criteria remain beyond their capabilities.

It was not until 2010 that the first paper-reviewer auto assignment system, the Toronto Paper Matching System (TPMS), was built for a conference, which was upgraded to integrate with the Conference Management Tool (CMT) developed by Microsoft and introduced in 2012 (Charlin & Zemel, 2013). Nevertheless, the idea of auto paper-reviewer assignment was initially proposed almost ten years earlier by Dumais and Nielsen (1992). In this system, the paper-review task was approached from an information retrieval point of view. However, the problem with this approach was the retrieval of more than one reviewer qualified to work on the paper as the system treated the paper as a query, based on the keywords present in the document. Other available approaches are Semantic Indexing, Vector Spacing, and Mixture Language (Hettich & Pazzani, 2006; Mimno & McCallum, 2007).

One of the models used for expert finding (i.e., reviewers) is the Latent Dirichlet Allocation (LDA) (Blei, Ng, & Jordan, 2003). The LDA model functions by matching a group of data sets on the reviewer-scoring side to another group of data sets on the paper-scoring side and then assigning a match based on probabilistic matrix factorization. Another model is the C-Value model (B. Li & Hou, 2016), which works by extracting text from a document in three steps: POS tagging, statistical ranking, and linguistic filtering (Junior, Magalhães, Caseli, & Zangirolami, 2015). Page Rank (Misale & Vanwari, 2017) is another model that is used to create and rank profiles with the same algorithm as Google Search does. The technique used by this model is to count the number of links to a document to rank its importance. Some systems have been proposed to identify the level of expertise of a subject, one of which is the Expert Finding System (Anongnart, 2012). The system uses all four models: c-value to identify the domain, LDA to determine the topic, Page Rank to assign numerical weights to expert profiles, and Lambda to find the ranking list of experts. Another system, proposed by Eto (2015), combines word-based models such as LDA and extends co-citation such as the Page Rank model. Zhang, Ma, and Zhong (2009) propose a system based on Wikipedia, which builds profiles based on publications submitted by reviewers. Goldsmith and Sloan (2007) discuss the problem of assigning conference papers to the most appropriate reviewers. They utilize a bidding approach to assign submitted papers to program committee members. In addition, they utilize expert preferences and keywords provided in the paper to find a potential match among available reviewers on the list. X. Li and Watanabe (2014) propose a new method to assign papers to reviewers automatically based on preference and topic matching. They conclude that their proposed algorithm can outperform the Hungarian algorithm (Kuhn and Yaw (1955)) as it significantly reduces the time consumed by the comparison process. Editorial Manager (2018) provides a sophisticated matching capability based on a hierarchal classification of keywords to describe papers and reviewers, hence finding better matching candidates by identifying reviewers and papers that belong to the same cluster. The ACM Computing Classification Scheme (2012) describes an n-level tree of specialties for the computing field to organize subjects by area. This classification scheme is mostly used by some journals and conferences to organize contributions and candidate reviewers. The tree varies in depth and is not restricted to 3 levels as in the case of our work.

The named approaches and classification schemes might work in environments where the journal or the conference is restricted to certain specialty fields. The case of UQU is different, because it attempts to cover all possible areas within the university so that over a period of five years the research strength within different colleges can be identified. Additionally, many of the solutions lacked the capability to process partially matching candidates automatically, especially if reviewers belong to different tracks. Consequently, UQU needs a new, yet effective, methodology to process and manage its funding activities.

Brief Overview of the Proposals and Funding Scheme at UQU

Through the DSR, the university calls for proposals at the start of every academic year (i.e., October). Approximately 8,000 faculty members from 27 colleges and their corresponding departments can apply through its in-house research management system (RMS) (Research Management System, 2018). DSR is the funding agency within the university that manages the entire process starting from the calls for proposals until project completion. DSR sends invitations to reviewers around the world asking them to join the UQU reviewers panel. Interested reviewers can register their research interests and other information into the system for future assignments when proposals that match their specialty are received.

DSR applies the regulations of Saudi universities, defined by the Ministry of Education in Saudi Arabia, for funding and managing research projects. There are some restrictions regarding the number of proposals that can be submitted at one time during a round and the total quota for every researcher. These regulations impact the proposal submission, reviewing, and funding processes. Readers can consult the Regulations of Scientific Research in Saudi Universities (2018) for more information about the regulations that every Saudi university funded by the government must comply with. After reviewers show their interest, a list of their publications is obtained, and they are filtered based on their h-index as an indicator of their qualifications as a reviewer.

The PRATO System

Details of the classification methodology, reviewer-proposal matching, assignment process, and the design considerations of PRATO are described in the following sub-sections.

Methodology

Identifying the research interests of every reviewer is important as it requires processing from CVs or profiles on the internet and extracting relevant keywords that match the proposals' research domain. Although this approach might be usable, it is not always reliable due to its lack of accuracy. Thus, we adopted a simple, yet more controlled, approach to accomplish this task.

A 3-level tree of research specialties was defined as per the specialties in the colleges of the university. The three levels are:

- Level 1: general field of specialty
- Level 2: track
- Level 3: sub-track

The general field represents the colleges within the university, the tracks represent departments, and the sub-tracks represent different specialties within departments. The tree was built this way to allow recording of necessary statistics about all faculty members in the university and apply them easily. The values in the tree are read from the UQU academic system, which defines colleges and departments. However, sub-tracks are defined utilizing the ACM classification scheme for computing and engineering fields. Additionally, based on surveying the available specialty interests within the university for areas that lack a referenced categorizing scheme (e.g., History, Geography, and the Arabic Language), several workshops have been conducted to introduce the concept of the system so that representatives from every college can complete the sub-tracks that faculty members are currently interested or working in.

This tree is coded using the numeric notation system of Dewey Decimal Classification. For example, of the notation “10.1.2”, the “Computing” general field is represented by the number “10”, the “Computer Science” track is represented by the number “1”, and the “Software Engineering” sub-track is represented by the number “2”. Proposals are classified according to this specialty tree. Researchers need to select all three levels in the tree, so that their research proposals can be classified accordingly. Additionally, reviewers who show interest in reviewing proposals for the university must pick the correct combinations that match their research specialty to construct a profile at the time of registration. They can then select the most appropriate categories that fit into the tree. In addition, they can provide reviewing interests, which we use to match reviewers with other sub-tracks in the tree that may belong to different tracks or even general fields of specialty. Figure 1 below illustrates the classification scheme of proposals and reviewers.

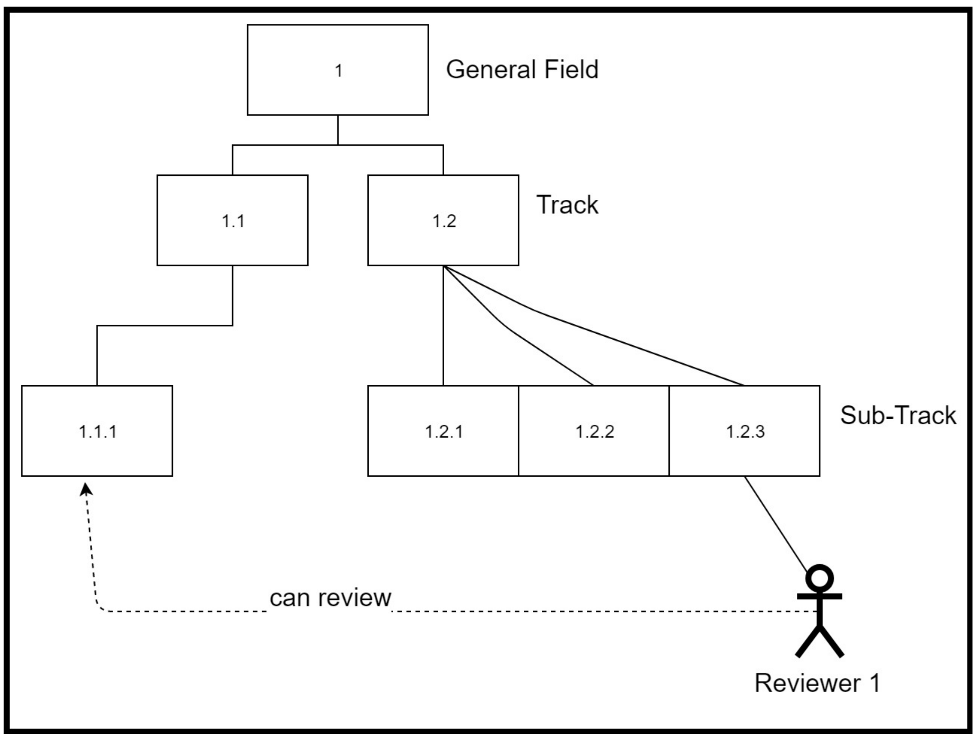

Figure 1. Proposal-reviewer classification scheme

This classification scheme covers most selections that can be made automatically by the system. The remaining candidates fall into other categories, in which proposals and reviewers belong to different leaf nodes (i.e., sub-tracks). Matching a proposal with a reviewer from a different category is not trivial and needs more detailed analysis. We have also added a stage where we encode reviewers’ research interests as per our tree of specialty. For example, a reviewer who falls into the “Software Engineering” sub-track can review proposals within the “Information Science” sub-track because of the presence of shared keywords such as “big data” and “data mining” in both the reviewer’s interests and the proposal text. We have automated sub-track matching to reviewer research interests by computing the highest co-occurrence of words in reviewers’ research interests and the proposal submitted by a researcher. More co-occurrence means greater relevance to a proposal within the same general field. We defined the threshold of acceptance as a required minimum number of co-occurring phrases or words. For example, a reviewer in sub-track 10.2.1 can review proposals in sub-track 10.3.4 if 10 or more keywords or phrases from the proposals are also present in the reviewer's research interest or publications. Based on that, we generate a code for each research interest that a reviewer has registered in the system.

It was noticed that reviewers who are classified under a certain general field can still review proposals that belong to different general fields. For instance, we have witnessed cases where a reviewer in applied science-biology can review a proposal in medicine. At the current stage of this project, we limited our work to encoding reviewers' specialties only to the sub-tracks that are leaf nodes under the same general field.

MODELLING REVIEWER-PROPOSAL MATCHING

A reviewer (R) can review a proposal (P) if the reviewer's sub-track R_t falls in the proposal's sub-track P_t. For cases in which a reviewer belongs to a different node than the proposal, we attempted to formally model that to identify the nearest match between them. When they fall under different sub-tracks, we calculate the strength of the link (aka relevance) between the reviewer and the proposal based on the number of shared keywords between the reviewer's stated interests and the proposal. So, the relevance between R and P can be calculated by the *Jaccard similarity coefficient* (Jaccard, 1901) between R_k that represents the extracted keywords from the reviewer's profile and P_k that represents the provided keywords of the proposal as follows:

$$J(R,P) = \frac{|\mathrm{R}_k \cap P_k|}{|\mathrm{R}_k| + |P_k| - |\mathrm{R}_k \cap P_k|} \leq 1$$

Then, the overall strength S_{rp} is calculated as:

$$\mathrm{S}_{rp} = J(R,P) + \lambda$$

Where

$$\lambda = 1 \; if \; R_t = P_t, 0 \; if \; R_t \neq P_t$$

The overall result can be between 0 and 2. If the vertex weight is 0, a very weak or no link exists between R and P. If the weight is 2, there is a very strong link between R and P. For example, if

P = {Data Mining, OWL, NLP, Formal Methods, process models}

R = {Data Mining, NLP, AI, IoT, Formal Methods, XUML, SOA}

Then, J(P, R) = 0.33; If $R_t = P_t$, then the overall S_{rp} = 1.33.

Figure 2 illustrates the strength between a proposal and candidate reviewers by means of the weight.

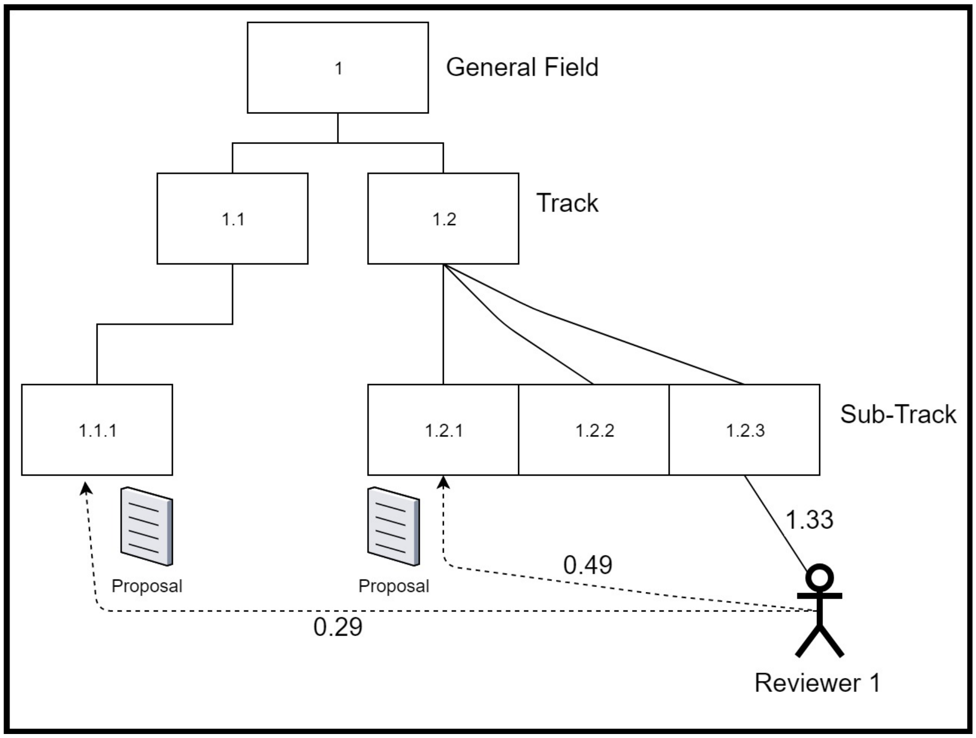

Figure 2. Proposal-reviewer-weighted matching

Overview of Proposals-Reviewers Assignment Cycle

The cycle of assigning reviewers to proposals is done in three main stages, namely, matching, voting, and linking. The matching stage starts by simply scanning through the tree of hierarchy to explore for reviewers who fall in the same sub-track as the proposal and flagging them as the most appropriate candidates for selection as reviewers for the proposal. Then, the list of reviewers those belong to different sub-tracks than the proposal and having a matching score above the threshold are listed. After the selection is made automatically and every proposal is matched with a candidate reviewer, we employed a voting mechanism on the results to prioritize the list, as well as to evaluate the accuracy of the selection process at this very early stage of the prototype. DSR has a scientific committee made up of various specialties. This committee can check the suitability of the reviewer candidates prior to making the final assignment and voting on their validity if needed. The highest overall score that a match has obtained will be ranked higher in the list. Once the list of matched proposals/reviewers is refined, it is sent to the linking stage, during which a proposal is pre-assigned to a candidate reviewer after checking the conditions, as explained in the next section. Once the linking stage is completed, all records are submitted for final assignment through RMS. These activities define one complete assignment cycle. A new cycle may then take place if the selected reviewers decline the invitation to review the allocated proposals.

PRATO Design Considerations

As highlighted in the previous section, the three stages defined by the PRATO system are matching, voting, and linking. At every stage, a number of conditions are checked to ensure that a proposal is assigned to the most suitable reviewer. The general conditions that PRATO examines throughout the cycle are:

- Reviewers must not be one of the team members of the proposal (i.e., co-investigators or consultants), as that will invalidate the review process.

- The selection must consider the number of completed assignments for each reviewer; the maximum number of completed reviews allowed for any reviewer is five.
- The reviewer must not be assigned to review more than three proposals at once. Based on successful completion, a new assignment can be made until the reviewer's total of completed reviews reaches five.
- Proposals must not be assigned to more than two reviewers at any time. Some proposals may need a third reviewer to confirm a decision of accepting or rejecting a proposal, but that is carried out after both reviewers have completed their reviews.
- The assignment must be randomly set in cases where several reviewers have a matching score above the threshold. Thus, no human intervention is needed, to achieve maximum transparency and integrity. Notably, the voting stage is an optional stage that was defined to validate the output of the matching stage. We expect to omit the voting stage at the advanced stages of PRATO usage.

Figure 3 illustrates the high-level structure of PRATO using a data-flow diagram, showing various functional components of the system and their interactions.

At the matching stage, each proposal is matched to reviewers, according to the algorithm in Listing 1, where P_{st} and R_{st} represent the sub-tracks that a proposal and a reviewer were classified under. R_c represents the total number of reviews completed by a reviewer. P_{id} and R_{id} are the proposal and the reviewer IDs, respectively. At first, PRATO checks whether a reviewer has already completed five reviews. Every reviewer must not exceed the limit of five proposals due to financial restrictions imposed by the DSR. The completed reviewing tasks are validated by the DSR team, and only the accepted reviews are counted as completed for a reviewer. There are cases where a reviewer has completed the assigned task but the DSR rejected the review due to inconsistency or other problems, such as a lack of detail. In such cases, the reviewer's count of completed reviews is then not incremented. After that, sub-tracks are matched by checking if they belong to the same category. If the conditions are fulfilled, then a proposal is matched with priority $_{p1}$. If the proposal falls into a different category other than the reviewer category, PRATO checks if both the reviewer and proposal belong to the same general field. If the check passes, PRATO attempts to calculate the similarity (relevance) between the reviewer and the proposal S_{rp}. We defined the threshold value of similarity for acceptance as 0.25. Once these conditions are fulfilled, the proposal is matched with a reviewer with priority $_{p2}$. Differentiating between priorities (i.e., $_{p1}$ and $_{p2}$) is necessary, to give reviewers who fall in the same sub-track a privilege over others in different sub-tracks.

```
If Rc < 5 then
       If PST = RST then
              Match_p1 Pid -> Rid;
       else if Pgf = Rgf then
              if Srp >= 0.25 then
                     Match_p2 Pid -> Rid;
              end if
       end if
end if
```

Listing 1. Matching algorithm

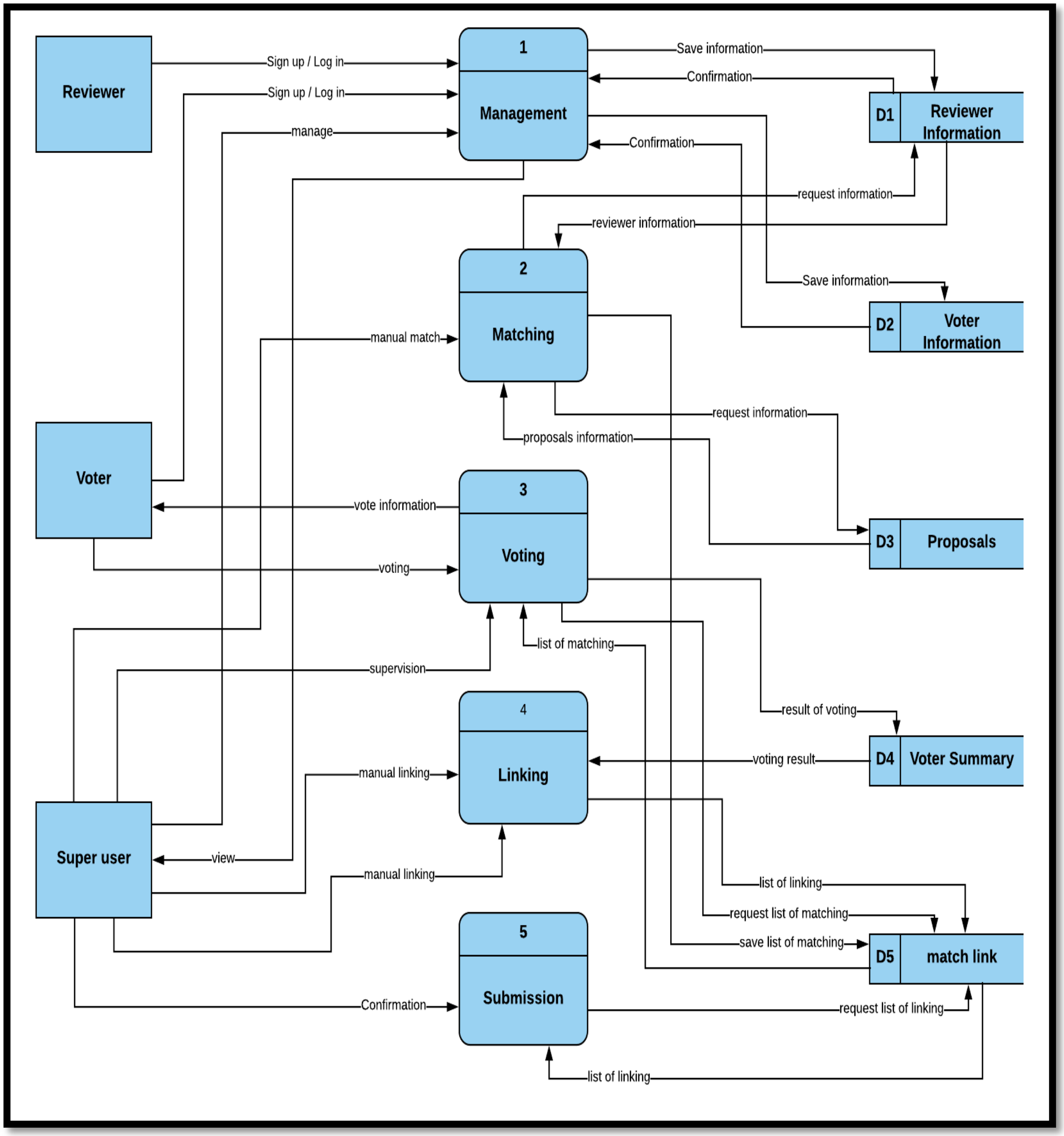

Figure 3. PRATO DFD

After a proposal is matched with a reviewer, the generated list is sent to the voting module in which voters can examine proposal-reviewer matching and prioritize the list based on their judgment. Every voter can inspect the title of the proposal and abstract, as well as the reviewer's name, specialty, and key words of research interest. If voters need more information, they can communicate with PRATO administrators to obtain necessary detail including any part of the proposal, e.g., the introduction and results sections as well as the reviewer's CV.

The selection of voters is done manually by the PRATO administrator, a super user, who is in charge of managing the entire review process at DSR. Every voter can vote with *Yes*, *No*, or *Abstention*. Voting *Yes* or *No* prompts the voter to enter an integer score to indicate his or her level of confidence in voting. The scoring integers are 1, 2, and 3, with "1" meaning weak, "2" medium, and "3" strong. For each proposal-reviewer match, the totals of confidence scores recorded for votes *Yes* and *No* are calculated respectively. Ultimately, the matches are divided into two groups based on the calculated totals of confidence scores, i.e., the *Accept* group or the *Reject* group, and the list of matched proposals and reviewers is re-ordered accordingly within each.

The underlying process behind the voting system works as follows:

- If the total score for *Yes* is higher than the total score for *No,* then the overall vote on the match will be *Yes*. And reversely, *No* if the total score for *No* is higher.
- In the case where both scores are equal, the final decision will be voted as *No*, indicating that the match should have less priority calculated by the overall score of *No* obtained.
- Finally, PRATO calculates the highest scores for the matches that obtain *Yes* and merges them into a single list in a descending order.

The list is then sent over to the linking module of PRATO. In this version of PRATO, we considered only the *Accept* group. The *Reject* group was considered only manually under some circumstances, e.g., none of the reviewers in the *Accept* group responded. However, we counted the percentage of records in the *Accept* and *Reject* groups as an initial evaluation of PRATO's accuracy. This is reported in the evaluation section of the paper.

During the linking stage, the assignments of proposals to reviewers are finally determined, after checking that all conditions are satisfied. We use the algorithm described in Listing 2, where X represents the matched reviewer/proposal record in the list passed from the voting module. X_{status} represents the current status of the record. We defined five values for record status, namely:

- Linked: when a proposal is already assigned to a reviewer
- Un-linked: when a proposal is withdrawn from a reviewer
- Rejected: when a reviewer rejects the assigned proposal
- Completed: when the reviewer completes the reviewing task and the review is approved by the DSR
- Not Responded: when a reviewer does not respond to the assignment request.

```
Read x matches from the list starting at top;
If x_status != "Linked" or "Completed" or "Rejected" then
       If R_cl < 3 && P_cl < 2 then
              If x_score(i) = x_score(j) then
                     Randomize selection of x;
              End if
              Link R_x with P_x;
       End if
End if
```

Listing 2. Linking algorithm

R_{cl} represent the number of concurrent reviewing tasks in which a reviewer must currently review at an instance of time, and P_{cl} represents the number of concurrent proposals already assigned to reviewers. R_x and P_x represent the actual instances of proposal and reviewer respectively. X_{score} represents the score of a proposal matched to a reviewer after the voting stage.

PRATO reads the list of matched reviewers and proposals, starting with the highest scores and proceeding to the lowest. It then checks the status of the match record to see if it is "Linked", "Completed", or "Rejected". If yes, it means that no assignment is needed. Next, the number of concurrent assignments of each reviewer is checked to ensure that it does not exceed the maximum of three allowed at any time, as per DSR regulations. Additionally, all proposals must not be assigned to more than two reviewers at a time. There are cases in which a proposal must be submitted to a third reviewer, such as when previous assignments result in different decisions and a third review is necessary to arbitrate. Once the total concurrent assignments have been checked, PRATO examines the list to determine whether it contains records with equal overall voting scores. If such cases are found, then a randomization function is called to select a record from the list. Afterwards, the proposal is assigned to the identified reviewer and the status of the record is changed to "Linked".

The output from the linking stage is submitted to the RMS, which is responsible for sending the actual proposal documents to the reviewers and providing them with evaluation forms.

EVALUATION

We evaluated the system during the last round of matching proposals with reviewers in 2017. In total, 125 proposals were submitted to the DSR, and 542 reviewers registered in our database, willing to join the UQU scientific committee to participate in the proposal-reviewing process. By running the PRATO system on the sets of proposals and reviewers, we obtained 983 records of proposal-reviewer matching/linking. The list of matches was then sent over to the voting stage to examine their accuracy and prioritize them appropriately. Each matching link was evaluated by two voters; if their votes were different, a third voter was added to arbitrate the decision. This served as a form of human evaluation of PRATO's matching accuracy, as voters can determine whether the allocated candidates are qualified and able to review a given proposal, based on their experience in the field. Thus, this helped us to evaluate the performance of PRATO at this early stage of development.

Of the 983 records generated by PRATO, 13.2% of the records were voted *No* (86.8% voted *Yes*), which is a very good indicator of PRATO's effectiveness. We reviewed the group of *No* votes to understand why they were rejected and found that most of them were either voted with low confidence by voters or involved a third voter, which means that at least one voter made their decision with acceptance. The group of *Yes* votes was passed over to the linking stage, and the first round of processing resulted in 250 links committed and submitted to the RMS for the actual assignment.

A follow-up analysis of reviewers' decisions of accepting/declining proposal assignments and the quality of completed reviews was conducted to gain additional insights about the system's performance. We recorded the status of every reviewer-proposal assignment as being either accepted or rejected by the reviewer. The most important status to consider was "rejected", and subsequently the number of reviewers who rejected reviewing a proposal was counted. For evaluation purpose, we considered the "rejected" status as an indicator that the assigned proposal does not match the reviewer's interest; however, that was not always the case since reviewers sometimes rejected due to being too busy at the time of the assignment. We considered the "completed" status as confirmation that the PRATO link was valid. We checked the quality of review reports submitted by each reviewer and flagged only the ones with critical and detailed comments as "completed". If a report was found to be of low quality, we returned it to the reviewer, requesting for a more complete review. If the reviewer failed to provide a quality review, the report was omitted and the reviewer was not compensated for the task. Based on this, we gathered data for the reviewing period, which lasted for four months and six rounds. The results of tallied frequencies are presented in Table 1.

Specifically, Table 1 shows the various linking rounds during the reviewing stage of the proposal. We needed six rounds to finish the entire review cycle. In each round, we linked the available proposals to randomly selected reviewers that were identified by the system. The first round included 250 links, among which seven were rejected (i.e., a reviewer rejected the proposal), 151 did not get a response from the assigned reviewer, possibly due to not receiving the notification email, and 92 were completed. The ratio of rejected to completed tasks was about 1:13 in the first round. We then started the second round and so forth until completion in round six.

Table 1. Proposals-reviewers linking status

Linking round	Number of proposals linked	Number of links rejected	Not responded	Completed	Rejection ratio	Completion ratio
1	250	7	151	92	2.8	36.8
2	158	3	71	84	1.89	53.1
3	74	8	24	42	10.8	56.75
4	32	1	15	16	3.12	50
5	16	0	9	7	0	43.75
6	9	0	0	9	0	100

Overall, the results appear very promising at this early stage of development, as we managed to finish successfully the entire reviewing cycle in six rounds without involving human decisions in the selection stage. The rejection ratio (approximately 3.1% on average) was minimal relative to the completion ratio (56.7% on average), and this result indicates that PRATO performed well in allocating reviewers to proposals automatically.

It is worth noting that we did not examine the impact of voting on the generated result, i.e., to compare the system's performances with and without voting. At this stage, we needed to make sure that we complete the reviewing cycle and deliver quality results within the allocated time frame specified by the DSR.

CONCLUSION AND FUTURE WORK

This paper reported our methodology of automatic proposal-reviewer matching and assignment implemented as part of the PRATO system. The system, a taxonomy-based tool, can help assign proposals to reviewers automatically with consideration of some special restrictions at hand. The key contribution of PRATO is its ability to handle the assignment of proposals to reviewers in an automated manner, especially in the case of partially matching candidates. This automated feature is not available in other systems designed for the same purpose, which renders them inadequate to meet the needs of UQU, especially to meet the requirements imposed by the university's ecosystem and funding policies.

The PRATO specialty tree is currently limited to three levels, and we did not have a chance to investigate if it can still function equally well with a specialty tree of four or more levels. Moreover, PRATO was built on the assumption that proposals and reviewers were classified according to certain controlled taxonomy. We believe that its partial-matching capability might be utilized when a random list of reviewers and proposals need to be processed, although this possibility is yet to be determined empirically. Acceptance and rejection criteria were limited to reviewers' responses to the assignments, and selection decisions were made without considering their justification.

Notably, one factor contributing to PRATO's success is the controlled method of categorizing proposals and reviewers according to the research specialty tree. In the future, we plan to examine PRATO's performance with research specialties based on different classification schemes, which allow specialty areas to be represented in graphs rather than trees. Therefore, the scope for reviewer selection can be widened to include more general fields of specialty. Moreover, we will try to record the reasons for rejection to identify accurately whether it is due to improper assignment or other reasons. The setting of threshold value of link strength for acceptance will be tested at different levels to investigate its impact on the accuracy of the selection. We also intend to test different matching algorithms and measures of similarity between a reviewer's profile of interests and a proposal, e.g. cosine similarity, to determine the most appropriate measure. We believe that there is still some room for further improvements to be identified by measuring the accuracy of assignment.

REFERENCES

The ACM Computing Classification System. (2012). *Advancing computing as a science & profession*. Retrieved from http://www.acmorg/publications/class-2012

Anongnart, S. (2012). Building Fexpert: System for searching experts in research university using K-MEANS algorithms. In *IEEE Symposium on Computers & Informatics* (pp. 176-179). Penang, Malaysia: IEEE. https://doi.org/10.1109/ISCI.2012.6222689

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

Charlin, L., & Zemel, R. S. (2013). The Toronto paper matching system: An automated paper-reviewer assignment system. In *Proceedings of ICML Workshop on Peer Reviewing and Publishing Models*.

Dumais, S. T., & Nielsen, J. (1992). Automating the assignment of submitted manuscripts to reviewers. In *Proceedings of the 15th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 233-244). New York, NY: ACM. https://doi.org/10.1145/133160.133205

Editorial Manager. (2018). *Online manuscript submission and peer review*. Retrieved from http://www.editorialmanager.com/homepage/home.htm

Eto, M. (2015). Combination effects of word-based and extended co-citation search algorithms. In *Proceedings of the 15th ACM/IEEE-CS Joint Conference on Digital Libraries* (pp. 245-246). New York, NY: ACM. https://doi.org/10.1145/2756406.2756957

Goldsmith, J., & Sloan, R. H. (2007). The AI conference paper assignment problem. In *Proceedings of the Workshop on Preference Handling for Artificial Intelligence* (pp. 53-57). Vancouver, BC: AAAI.

Hettich, S., & Pazzani, M. J. (2006). Mining for proposal reviewers: Lessons learned at the national science foundation. In *Proceedings of the 12th International Conference on Knowledge Discovery and Data Mining*, (pp. 862–871), New York, NY: ACM. https://doi.org/10.1145/1150402.1150521

Jaccard, P. (1901). Étude comparative de la distribution florale dans une portion des Alpes et des Jura [Comparative study of floral distribution in a portion of the Alps and Jura]. *Bulletin de la Société Vaudoise des Sciences Naturelles, 37*, 547-579.

Junior, A., Magalhães, C., Caseli, H., & Zangirolami, R. (2015). Topic modeling for keyword extraction: Using natural language processing methods for keyword extraction in Portal Min@s. *Journal of Language Studies, 23*(3), 695-726.

Kuhn, H. W., & Yaw, B. (1955). The Hungarian method for the assignment problems. *Naval Research Logistics Quarterly*, *2*(1/2), 83-97. https://doi.org/10.1002/nav.3800020109

Li, B., & Hou, Y. T. (2016). The new automated IEEE INFOCOM review assignment system. *IEEE Network, 30*(5), 18-24. https://doi.org/10.1109/MNET.2016.7579022

Li, X., & Watanabe, T. (2014). Paper-to-reviewer assignment, based on expertise degree of reviewers and relevance degree between reviewers and papers. *International Journal of Knowledge and Web Intelligence, 5*(1), 1-20. https://doi.org/10.1504/IJKWI.2014.065062

Mimno, D., & McCallum, A. (2007). Expertise modeling for matching papers with reviewers. In *Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 500-509). San Jose, California, USA: ACM. https://doi.org/10.1145/1281192.1281247

Misale, M., & Vanwari, P. (2017). A survey on recommendation system for technical paper reviewer assignment. In *International Conference of Electronics, Communication and Aerospace Technology (ICECA)* (pp. 329-331). Coimbatore, India: IEEE. https://doi.org/10.1109/ICECA.2017.8212826

Regulations of Scientific Research in Saudi Universities. (2018). *Consolidated regulations for scientific research in universities unified regulation*. Retrieved from https://drive.uqu.edu.sa/_/dsr/files/Grants/Scientific_Research_Regulations.pdf

Research Management System. (2018). Retrieved from http://dsr.uqu.edu.sa/rms/index.php

Wang, F., Chen, B., & Miao, Z. (2008). A survey on reviewer assignment problem. In *Proceedings of the 21st International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems* (pp. 718-727). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-540-69052-8_75

Zhang, W., Ma, J., & Zhong, Y. (2009). Incorporating similar people for expert finding in enterprise corpora. In *International Conference on Computational Intelligence and Software Engineering* (pp. 1-4). Wuhan, China: IEEE. https://doi.org/10.1109/CISE.2009.5362743

BIOGRAPHY

Basem Alkazemi is an Associate Professor in the collage of Computer and Information System, Umm Al-Qura University, Saudi Arabia. He obtained his PhD degree from Newcastle University, UK. His current research interests include Software Engineering and Data Mining. He is involved, as a PI, in a number of funded research projects in the area of WSN, Bigdata and NLP. Basem served as a reviewer in several international conferences and journals. He supervised a number of postgraduate students who conducted their thesis in software CD model, BPM, IoT and Bigdata for retails. Basem is the head of the software engineering research group in Umm Al-Qura University. At the moment, he is also holding the position of vice dean for research projects and grants affairs in the deanship of scientific research (DSR).