"An Intelligent data analysis in big data for peer review"

Bhadane Ashwini Dilip

Department of Computer Engineering Matoshri College of Engineering, Nashik

Dr. Prof. Varsha H. Patil

Department of Computer Engineering Matoshri College of Engineering, Nashik

Abstract

The assignment problem is a fundamental problem when it comes to assigning reviewers to research proposals. Matching of reviewers and research proposals affects the review quality. The number of proposals are continuously increasing day by day and hence they fail to satisfy the practical needs. This paper proposes an approach where the proposals will be grouped first and then will be assigned to appropriate reviewers. The basic idea is to identify valid proposals and reviewers, classify them, partition the proposals into groups and assign reviewers to proposal groups. A system is been developed based on the proposed approach for assigning the proposals to reviewers.

Index Terms—*Proposals assignment, Reviewers assignment, Proposals grouping, Decision making*

I. INTRODUCTION

Selecting appropriate research project for funding is an im- portant task in government funding agencies. Research project selection is also a complex one which often involves many sub-tasks. It generally begins with a call for proposals (CFP) which describes funding opportunities and requirements of the funding agencies, and then the CFP is distributed to relevant communities, such as universities and research institutions. Researchers in relevant communities who are interested in CFP-related topics can then submit their proposals to the corre- sponding funding agencies. Submitted proposals are validated, compiled and then assigned to field experts who are invited as reviewers to provide their opinions. These reviewers comment on the proposals based on their professional knowledge and with reference to the specific criteria issued by the funding agencies. The reviewing results are aggregated to determine which proposals would be funded. Assigning proposals to reviewers is one of the most important and challenging tasks. It must be done appropriately because assigning proper reviewers to a research proposal ensures reviewers have enough expertise to judge the quality of the proposal.

II. REVIEW OF LITERATURE

Paper 1: How to Choose Appropriate Experts for Peer Review

One of the major factors of construction of a university faculty is talent introduction. The academic level of a university as well as its comprehensive strength depends on the quality of talent. Hence, the universities must take effective steps to control the quality of applicants in the process of their talent introduction. In general, universities use peer review to assess applicants. They select a number of experts who have similar research areas as the applicant and then review their application documents. It is very important to choose appropriate reviewers as it will have great impact on the assessment of applicants because an appropriate reviewer will help universities to select excellent talent. On the other hand, a reviewer who is not suitable may result in the loss of talent. Hence, choosing appropriate reviewers becomes one of the key factors of peer review process. Currently, the process of selecting reviewer is manual. In this process, a university staff member first collects applicants application documents, personal information and then identifies his/her features. Then they try to find those research areas that are similar to the applicants with the help of database expert professionals. After that they send invitations to the selected experts by phone or e-mail. Finally, they decide whether the proposal should be accepted or not.

Paper 2: The Conference Paper Assignment Problem: Using Order Weighted Averages to Assign Indivisible Goods

Assigning indivisible items to multiple agents is a fundamental problem in many elds including computer science, economics and operations research. Algorithms for matching and assignment are used in a variety of application areas including allocating runways to airplanes, residents to hospitals, kidneys to patients, students to schools, assets to

individuals in a divorce, jobs to machines, and tasks to cloud computing nodes. Understanding the properties of the underlying algorithms is an important aspect to ensuring that all participating agents are happy with their allocations and do not attempt to misrepresent their preferences; a key area of study for computational social choice. An area that is near to many academics hearts is the problem of allocating papers to referees for peer review. The results of grant, journal, and conference reviewing can have signicant impact on the careers of scientists. Ensuring that papers and proposals are reviewed by the most qualied/interested referees most is part of ensuring that items are treated properly and all participants support the outcome of the processes. Making Sure these processes work for both the proposers and the reviewers is important and methods for improving peer review have been. There are a number of ways one can improve the quality of peer review. First is to ensure that reviewers are not incentivized to misreport their reviews for personal gain. Along this line there has been signicant interest recently in strategy proof mechanisms for peer review. Another way is to ensure that reviewers are competent to provide judgments on the papers they are assigned. The Toronto Paper Matching System is designed to improve the process from this paper centric model. A third alternative, and the one we focus on in this study, is ensuring that reviewers are happy with the papers they are asked to review. This is fundamentally a question about the optimization objectives of the assignment functions used.

Paper 3: Proposal reviewer recommendation system based on big data

A national research management organization supports different research areas by supporting business to foster learning, constructing the national base of knowledge production and pursuing greater competitiveness in academic research. This has led to the production of a variety of research proposals in each area. This research has been utilized as basic data of various forms for the development of related areas and subsequent study. The number of SCI (Science Citation Index) journal papers in science and technology in Korea has been increasing every year, from 10,000 in 2002 to 60,000 in 2013. Accordingly, the Korean government and Korean research management institute have been preparing an institutional system to select researchers engaged in advanced research to create a national research development organization to respond to domestic and foreign changes in the research environment and to improve the productivity of the organization responsible for national research development. Driven by the acceleration of globalization, there has been keen competition for up-to-date knowledge and high-value- added source technology. Recent years have seen a sharp rise in the number of papers and patent applications as the outcomes of research projects. Accordingly, when performing a research project, the importance of conducting a survey and analysis of preceding research based on the use of relevant keywords has been increasing. A research project proposal can only be improved when suitable researchers are selected. Thus, the construction of a system capable of supporting researchers based on their research outputs is very important for academic development. When selecting researchers by investigating existing research outputs, for example by performing a keyword search, it is possible to prevent investment in overlapping research projects. Our analysis of the trends in research projects led us to introduce a mandatory system, which automatically submits patent search results. In Korea, many aspects of systematic project selection in the national research and development business need to be improved. This requires a convenient method to search for research projects and to recommend researchers by utilizing the data of the national research management organization. Accordingly, it is necessary to design a tool to search the database efficiently and to construct a researcher database for research recommendation purposes. By extending existing database research techniques, it is possible to apply the research database to each area. In particular, the selection of a standard group from which an expert could be chosen for evaluating research proposals in related research areas, and the presentation of a reviewer recommendation model, are require.

Paper 4: A Conflict of Interest Declaration and Detection System for Peer Review Management

A fair peer-review process is key point for a successful academic event. From an authors point of view, fairness of view process is paramount to her research. From the academic events point of view, fairness has direct impact on its reputation. Fairness is affected by many factors, some of them are expertise of reviewers, review comments quality, review form design, etc. However, the most important factor is the relationships between authors and reviewers. We present reviewer suggestion system that focuses on declaration and detection of conflicts of interest in the peer-review process, an issue that has received attention despite its significance in upholding quality and fairness of an academic event. In academic peer-review, we can categorize conflicts of interest into two types, definite conflicts of interest and latent conflicts of interest. As a common practice in existing conference management systems the definite conflicts of interest can be collected by a set of declaration rules. These rules cannot cover all conflicts of interest. For instance, an author and her academic siblings may have conflicts of interest but this is not required to be declared according to the rules. One possible reason is that not every academic sibling relationship has conflicts of interest. These conflicts of interest can be determined but they could influence the quality of a peer review process.

Paper 5: A decision support approach for assigning reviewers to proposals

Traditional assignment methods rely heavily on a single decision maker (e.g., panel chair) to manually analyst the title, abstract, keywords and other parts of the proposals and then identify a set of reviewers who are most likely to review the proposals. Furthermore, constraints that should be considered when assigning proposals to reviewers make the assignment problem more complicated. For example, proposals should be evaluated by reviewers who are knowledgeable in the

corresponding research area, and relationships between applicants and reviewers that could affect the justice of review process (e.g., co authorship) need to be avoided. It presents a great challenge for managers who are responsible for the assignment task, especially when there are large amounts of proposals and reviewers.

Paper 6: Context-Aware Reviewer Assignment for Trust Enhanced Peer Review

The research conferences has an aim to share research progress and findings among students and support progress of academic disciplines of the year. The quality of publications of peer review systems assures the quality of publications. This is the reason why peer review systems are being used in funding applications. Below are the stages of a peer review process- 1. Submission: Authors submit their proposed submissions to the peer review system. 2. Proposal allocation and review: Before peer review process starts, each submission is reviewed and allocated to some reviewers who will then decide whether to accept or reject the proposal.

3. Organizers Decision: From the collected peer reviews, a decision should be made by organizers, on the acceptance of the reviewed submissions. Peer review process is very significant to the quality of conferences, and even to the development of academic disciplines.

III. PROPOSED METHODOLOGY

An integrated approach assisting in assigning proposals to reviewers where proposals need to be partitioned into groups. The proposed approach facilitates the reviewer assignment through the below steps: Step 1. To identify valid proposals and reviewers, Step 2.To classify proposals and reviewers accordingly, Step 3. To partition proposals into groups and assign reviewers to proposal groups.

A. Architecture

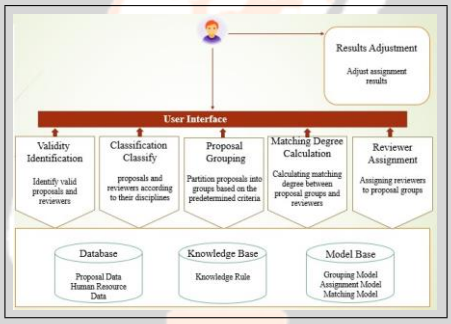

Fig. 1. Proposed System Architecture

B. Algorithms

The basic idea of our proposed approach is to group proposals first and then assign reviewers to each proposal group.

1) **Step 1- Identifying valid proposals and reviewers** The aim of this step is to identify valid proposals and reviewers.
Rules for general program
//Check number of proposal to proceed

- IF Number of on-going projects + Number of Submitted proposals= 2
THEN Status = Undetermined

//Check here for the organization where it is registered

- IF Unregistered (PI) or Unregistered (affiliated or- ganization)
THEN Status = incomplete

//Check PI status as below

- IF Position of PI = Full-Time Graduate Student OR Position of PI = Part-Time Graduate Student OR Position of PI = Retired
THEN Status = Invalid

//Check PI in Reputation list which is stored in record

- IF PI in Bad Reputation List Record THEN Status = undetermined

Rules for identifying valid reviewers

- IF reviewer not active in recent five years THEN Status = undetermined
- IF reviewer in Bad Reputation Record THEN Status = invalid

2) Step 2- Classifying proposals and reviewers

Assume that there are K discipline areas, and Dk is a field that denotes discipline area k (k = 1,2, . . . ,K); Assume that there are J reviewers; Rj denotes reviewer j (j = 1,2,. . . , J); Rnj denotes the nth research area of reviewer Rj, Sk denotes the reviewer sets of discipline area k, then Sk can be calculated as follows:
For k = 1 to K // nth no of discipline area loop count For j = 1 to J // Reviewer Count
If Rj R Categorized reviewers Then //Check reviewer category to Identify the Category
N = Number of discipline areas of reviewer Rj For n = 1 to N
If Rnj Dk Then
Rj will be added to Sk End Loop
End If Next
If n = N
//To check category of matched reviewers and proposals Then Rj will be added to set Categorized reviewers End If
End If Next Next

IV. RESULT ANALYSIS AND DISCUSSION

A. Results

The proposed approach facilitates the reviewer assignment through the below steps: Step 1. To identify valid proposals and reviewers, Step 2.To classify proposals and reviewers accordingly, Step 3. To partition proposals into groups and assign reviewers to proposal groups.

B. System Requirements

*1) **Software Requirement**:*

- Operating System : Microsoft Windows 7
- IDE : Visual studio
- Language : Java script
- Database : SQLite

*2) **Hardware Requirement**:*

- Processor : Core Intel 3 or Above
- RAM : 1 GB
- Hard Disk : 250 GB

V. CONCLUSION

This paper presents a novel approach to solve the proposal assignment problem in funding agencies where the number of proposals is large. In the proposed approach, knowledge rules and decision models complement each other to improve the assignment process. Based on this approach, a decision support system has been designed to enable program directors or department managers in funding agencies to complete the proposal assignment task efficiently and effectively. The major benefits of our method are: (1) It is in compliant with the current workflow and standards of the proposal assignment process of NSFC (2) It frees program directors or department managers from manual and time-consuming task of the pro- posal assignment (3)The web-based DSS enables the users to access it at any place and time; (4) The system can be easily incorporated with the existing ISIS to support the entire project selection process in NSFC.

REFERENCES

[1]Jing Wu Lian1, Nicholas Mattei2, Renee Noble1, and Toby Walsh3, "The Conference Paper Assignment Problem: Using Order Weighted Averages to Assign Indivisible Goods", 19 May 2017.

[2]Lei Li1*, Yan Wang2, Guanfeng Liu3, Meng Wang1, Xindong Wu1,4, "Context-Aware Reviewer Assignment for Trust Enhanced Peer Re- view", 19 June 2015.

[3]G. Sena Das a,*, Tolunay Gken b, "A fuzzy approach for the reviewer assignment problem", 12 March 2014.

[4]Ho Sun Shon, Sang Hun Han,Kyung Ah Kim, Eun Jong Cha, Keun Ho Ryu, "Proposal reviewer recommendation system based on big data for a national research management institute", 26 April 2016.

[5] Duanduan Liu1, Wei Xu1, Wei Du1,2 and Fuyin Wang1 , "How to Choose Appropriate Experts for Peer Review", 2015.
[6] Mingliang Yue1, Kailin Tian2 Tingcan Ma1 , "An Accurate and Impartial Expert Assignment Method for Scientific Project Review", 6 November 2017.
[7] Stacy L. Janak, Martin S. Taylor, and Christodoulos A. Floudas* ,Maria Burka and T. J. Mountziaris, "Novel and Effective Integer Optimization Approach for the NSF Panel-Assignment Problem: A Multiresource and Preference-Constrained Generalized Assignment Problem", 2006.
[8] Yunhong Xu a,b,*, Jian Ma a, Yonghong Sun a, Gang Hao c, Wei Xu d, Dingtao Zhao b, "A decision support approach for assigning reviewers to proposals", 2010.
[9] Siyuan Wu,Leong Hou U, Sourav S Bhowmick,WolfgangGatterbauer, " PISTIS:A Conflict of Interest Declaration and Detection System for Peer Review Management", 2010.
[10] David, M., Andrew, M., "Expertise modeling for matching papers with reviewers. In Proceedings of the 13th ACM SIGKDD international conference on knowledge discovery and data mining", 2007.
[11] Deep, K., Das, K. N. , "Quadratic approximation based hybrid genetic algorithm for function optimization. Applied Mathematics and Compu- tation", 2008.
[12] Amendola, G., Dodaro, C., Leone, N., Ricca, F, " On the application of answer set programming to the conference paper assignment problem", 2016.
[13] Aziz, H., Lev, O., Mattei, N., Rosenschein, J.S., Walsh, T, "Strategyproof peer selection: Mechanisms, analyses, and experiments", 2016.
[14] Price, S., Flach, P.A, "Computational support for academic peer review: A perspective from articial intelligence", 2016.