Fair and Accurate Review in Publication Process: A learning-based Proactive Approach for Assigning Reviewers to Manuscripts

[1]Aboli H. Patil, [2]Parikshit N. Mahalle

[1]Research Scholar, Research Centre at Department of Computer Engineering, SKN College of Engineering, Pune, India

[2]Professor, Department of Computer Engineering, SKN College of Engineering, Pune, India

ABSTRACT: Peer review is one of the most crucial and important tasks that are associated with academic conferences, journals and grant proposals; and assignment of an appropriate reviewer plays vital role for accurate and fair review process. This paper presents a learning based proactive system that assigns reviewer(s) whose expertise matches with the domain(s) of the paper satisfying constraints. The assignment of reviewer to paper needs to satisfy various constraints such as maximum number of papers per reviewer, minimum number of reviewers per paper and conflict of interest. he core challenge in reviewer paper assignment is to make the computer understand the subject domain of experts and papers. In proposed system, features are extracted from title, abstract and introduction section of publications of reviewer and submitted papers. These features help the model learn the domain features of experts and submitted papers more accurately. Once the training set is built utilizing the inherent correlation between abstract and title, the model is trained and the similarity between reviewers and papers is predicted. The experimental results on test data set of AAAI 2014 and NIPS 2019 demonstrate the effectiveness of the proposed system.

INTRODUCTION

Quality research publications are on utmost impotence in academics. Conferences and journals provide a strong platform to researchers and academicians for publishing and getting recognition to their research work. One of the most multifaceted and essential task associated with conferences and journals is assigning appropriate reviewer to submitted manuscript. Here appropriate reviewer refers to the expert whose topic of expertise and topic of the paper match and there is no conflict of interest. This assignment problem is popularly known as Reviewer Paper Assignment (RAP). In review process, each submitted paper isto be reviewed by most accurate and unbiased reviewers for determining the quality of scientific knowledge of paper. Further the reviews are forwarded to author for improvement of paper quality and decision is made whether the paper is accepted or rejected [1]. For quality publications, the review process plays an important role and to support the review process, it is important to assign papers to the most appropriate reviewers as the incorrect reviews directly affect the quality of publications [2].

In RAP, assigning reviewer to submitted paper need to focus on two important parameters as - relevance and fairness. Relevance is computed by determining the topic similarity between the expertise domains of reviewers and submitted manuscripts [3]. For fair reviews, one needs to assure that assignment of papers and reviewers have no Conflict –of –Interest (COI). Typical conflict of interests includes- working in same institute or university or industry, a reviewer and an author through intermediate co-authors, and an author –coauthor as researchers and doctoral supervisor[1]. Literature study reveals that there is a need of RAP system that provides accurate and fair reviews by accurately assigning the reviewer to the paper.

The core contributions are as follows:

- It is noticed that new manuscript is quite different from other papers in words because of the existence of plagiarism detection software. But use of terminology, concepts, and logic are common in academia, so text processing methods can capture the field differences caused by these features. In proposed system the field relationship between the reviewer and the manuscript through textual information is well exploited.
- Often the domains that are research topics are extracted by processing the publications of reviewer and submitted manuscripts separately. The proposed system trains model to learn extracting topics of papers and expertise of reviewers together; utilizing manuscripts and published papers as one corpus.

ORGANIZATION OF PAPER

The paper is organized as follows: section 2 presents work related to RAP, Section 3 provides details of the proposed system architecture, Section4 presents experimental results and Section 5 provides concluding remarks.

MOTIVATION

Multi-disciplinary domain research leads to increasing count of publications submitted at conferences and journals for review by well-known experts. Reviewers are assigned to paper either manually or using semiautomatic systems. The core challenge is simulating the human domain expertise in identifying research fields of submitted papers and domain of expertise for reviewers. The fair and accurate review process depends highly on appropriate reviewer assignment. One more problem is, rejection with reason that the submitted paper is out of the scope of the journal. Popular journal finder systems too yield the abrupt results. Need is to match the proficiency of the expert to the multidisciplinary domain of submitted paper. Lots of application areas motivate to provide solution to reviewer assignment problem. The proposed approach with the help of inherent correlation among various fields like title, abstract, introduction and keywords as the topic extraction fields, extracts the topics of reviewer expertise and manuscript accurately.

RELATED WORK

An Adequate number of research publications are available related to reviewer assignment problem. Literature survey reveals that, reviewer assignment problem plays an important role and is thrust domain for research. RAP process mainly consists of four phases as- building reviewer's profile, building paper profile, computing similarity between these profiles and assigning appropriate reviewer to a paper satisfying the constraints. Constraints associated with assignment process are COI, maximum number of papers per reviewer and minimum reviewers per paper. Sufficient literature is available elaborating these four phases of Rap process. For building reviewers and papers profile, the useful information from title of paper, abstract of paper and keywords can be extracted. There are two methods for computing paper-reviewer similarity as- explicit methods and implicit methods. [9-12]
• Explicit methods: In explicit method, it is necessary to authors and reviewers to provide additional information regarding their papers and the competences.
 • Implicit methods: In implicit methods, it is not necessary for authors and reviewers to provide additional information regarding their papers and competences. Similarities are calculated based on content analysis of publications of expert.
Various researches have presented their work in RAP domain; some of them are referred here. Price and Flach proposed an explicit method of similarity computation that is based selection of keywords / topics from pre-defined list [4]. This idea is a very simple as in this technique, when the author submits his paper for any conference, he selects the area/ topic where his paper belongs, and reviewer also follows the same step while registering for the conference. Further by using Jaccards similarity, the similarity index is computed and based on similarity index paper is assigned to reviewer [4]. This method has limitation as when author has not selected keywords/ topic, then topic extraction accuracy decreases; it further affects the accuracy of assignment of reviewer to a paper. In another technique, along with keywords, the abstracts of papers are given to the reviewers to know their willingness to review a paper. If count of submitted papers to a conference is high, then reviewers are not likely to browse all papers and read their abstracts. Thus, collected bid (or preferences) will be sparse and incomplete. For avoidance of such situation, the constraint of maximum number of papers per reviewer as per his preference is to be satisfied and the papers assigned should match his/her expertise. And second, the missing preferences could be "guessed" by applying collaborative filtering techniques as suggested in by Conry et al. and Rigaux [5].The recommendation of papers to reviewers could be based on paper-reviewer similarities calculated in any method.
The common method used for RAP is a feature-based matching proposed by Kalmukov in which keywords or topics are hierarchically organized in taxonomy [2]. It considers the count of number of common keywords. Along with this, it also determines the semantically closeness of non- matching keywords. A non-zero similarity is calculated even if the paper and the reviewer do not share any keyword in common. Reviewer is given with bidding facility in terms of providing his willingness to review specific papers. Andreas Pesenhofer et al. presented a paper for computing paper-reviewer similarities based Euclidian distance between the titles of the submitted papers and the titles of all reviewers' publications [5]. Stefano Ferilli et al. use Latent Semantic Indexing (LSI) for the automatic extraction of paper topics, titles and abstracts of a manuscript submitted by author. From DBLP dataset the titles of reviewer's publication are extracted [6].   Zemel and Charlin et al proposed a paper assignment system known as "The Toronto Paper Matching System (TPMS)" [7]. This system extracts reviewer's previous publications from

Google Scholar and builds the reviewers profile. Further by using Latent Dirichlet Allocation (LDA), TPMS finds the research topics. Along with this the system also supports reviewer's self-assessment of expertise in respect to the submitted papers. For building profile of paper Latent Dirichlet Allocation (LDA) is applied to the submitted manuscript.

For matching the similarity between reviewer and author, Liu et al. proposed a recommender system that calculates paper reviewer similarities based on three aspects of the reviewer: expertise, authority, and diversity. Authority refers to public recognition of the reviewer in the scientific community, while diversity— whether he/she has diverse research interests and background [8]. Latent Dirichlet Allocation (LDA) is applied over the sets of submitted papers and reviewers' publications to extract their topics. Then cosine similarity is used to calculate the relevance between the topic vectors of each paper and each reviewer's publication [14]. Authority is determined by constructing a graph that consists of the paper being processed and all of its candidate reviewers. Two reviewers are connected with an edge if they have co-authored at least one paper. The weight of the edge depends on the number of papers they co-authored. The intuition behind this is that if a reviewer is well connected which means it has many co-authors; he or she would be considered as having higher authority. It is observed that to incorporate the expertise, the diversity and the authority, the researcher has used a Random Walk with Restart (RWR) model on the graph.

For more accurate results, for building reviewer's profile, the various fields are to be considered such as his/ her expertise research domain, recency of his/ her published papers in particular research domain and quality of publication. The quality of publication is computed using citations of published papers, authored as books, book chapters and number of PhD students Supervised. This information is collected using global sources like DBLP, Aminer, Google scholar, and research gate. By considering all these factors and based on the reviewer's publication and Latent Dirichlet Allocation (LDA), reviewer's profile is built.[14-17]

RESEARCH GAP

It is observed that recent publications are multi and inter disciplinary, and needs the reviewer who can expertise in multiple domains. For accurate reviewing of such papers, we need to assign set of reviewers to such papers so that all topics of paper are covered. The other key issues related to RAP process is the expertise extraction of reviewer using his/her publications. These publications of reviewer are spread across long span of his/her career. Better accuracy can be achieved if we assign more weighing to domains of recent publications. It is also noticed that for most of these publications, the reviewer is second or third author. [2, 8, 10]

Though enough attention has been paid by researchers on RAP, Literature survey discloses that the existing systems for reviewer assignment has some challenges as – a uniform distribution of papers to reviewers, coverage of all topics of paper across the set of assigned reviewers, and high accuracy of assignment. The proposed system addresses these issues and challenges and provides more accurate results.[11-13]

THE REVIEWER PAPER ASSIGNMENT PROBLEM

The reviewer assignment problem is defined as –for a set of papers and a set of reviewers, the main objective is to assign the appropriate reviewer to a paper. Expected outcome is assignment of paper to reviewer with high relevance and low conflict of Interest satisfying the constraints.

Let us consider, *P*is a set of submitted papers.

$P = \{p_1, p_2, p_3, \ldots p_n\}$.

R is a set of reviewers $R = \{r_1, r_2, r_3, \ldots r_m\}$ *are given.*

For each paper $p_i \in P$, $r_j \in R$, $M(p_i, r_j)$ *is presented as a match if* r_j *is assigned to* p_i.

PROPOSED SYSTEM ARCHITECTURE

System architecture for Reviewer Assignment Problem is as shown in figure 1. The conference paper is input to the system that is collection of research papers. After pre- processing of each paper, the system generates topic list and relationship among topics, and topic dictionary. The system mainly uses Gibbs sampling with Latent Dirichlet Allocation algorithm for topic discovery. The system architecture of proposed methodology consists of major phases as data pre-processing, Gibbs Sampling with Latent Dirichlet Allocation, topic analysis, and Similarity Measure Relevance, and ranking and assigning a reviewer to a paper.

Data Pre-processing-The input to the data processing phase is conference manuscript papers. By applying language processing techniques nouns, adjectives, and adverbs of the statements are extracted. The extracted words are further lemmatized. This word set is used for further processing for topic modeling.

Gibbs Sampling with Latent Dirichlet Allocation-Extracted words from the data preprocessing phase are processed with Gibbs sampling and Latent Dirichlet Allocation for finding topics. This algorithm is an iterative method to find maximum likelihood or maximum posteriori (MAP) estimates of parameters.

Topic Analysis-Topic analysis deals with finding the relationship among title, abstract, introduction, keywords and conclusion as supervisory fields to understand the research area of paper. The topic (research area) relationship between title and abstract is used to train the model. This model is used to calculate the similarity between manuscript and reviewer's publications.

Relevance Measure-To find most appropriate reviewer for a paper we need to calculate relevance between reviewer and paper. Relevance between reviewer and paper signifies the similarity level of research interest of reviewer and research area of paper. In order to do so, weights obtained during topic analysis for each manuscript are sorted in descending order and considered top 5 topics having maximum weights as most relevant topics for a paper. Similarly, we obtained top 5 most relevant topics for each reviewer using his/her publications. Using these top topic weights similarity index between each reviewer and paper is calculated.
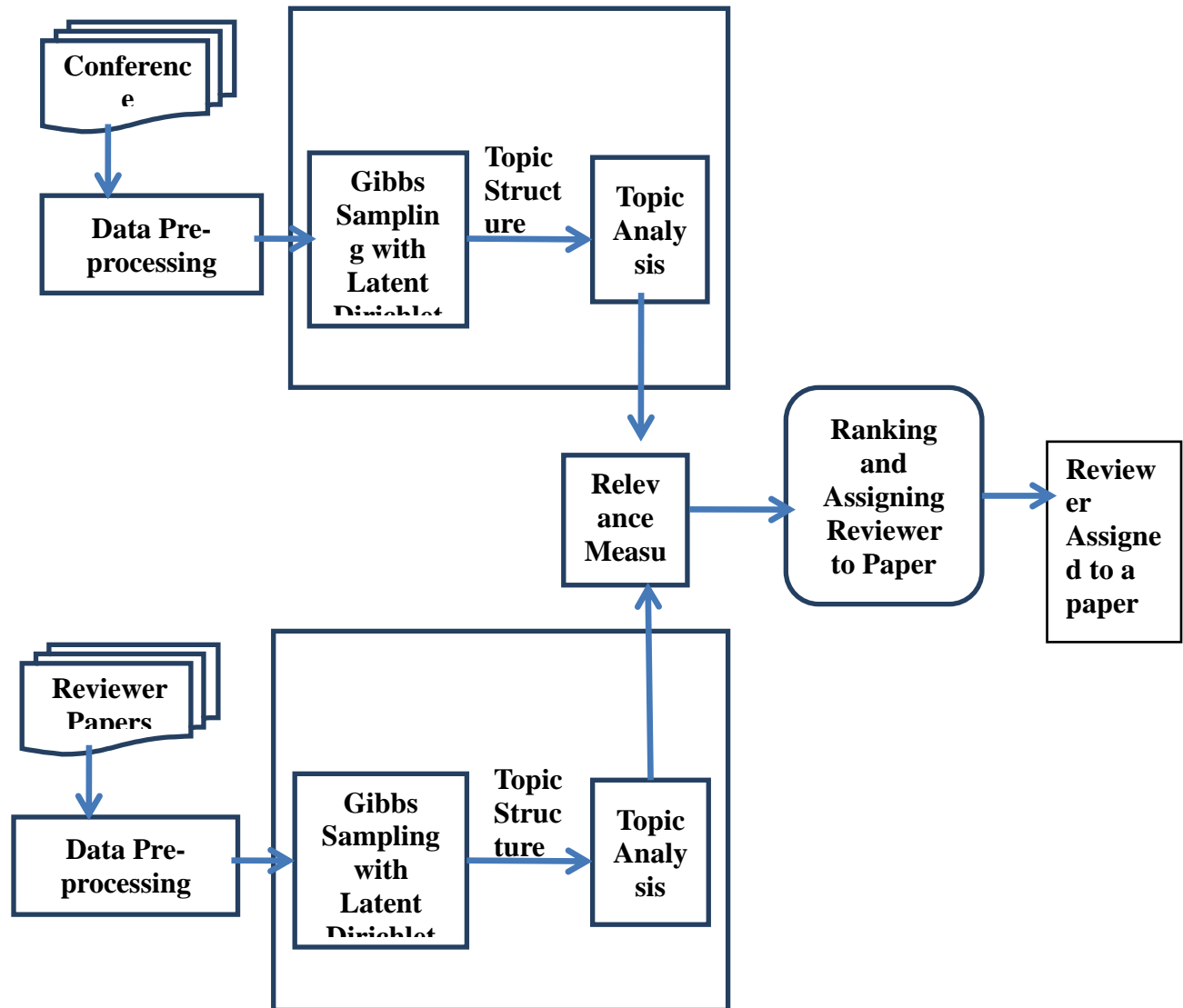

Figure 1: Proposed System Architecture for RAP

Ranking and assigning reviewers to paper-Similarity index that is computed in previous step and Conflict of Interest parameters are used ranking the reviewer for particular paper. Based on ranking, appropriate reviewer is assigned to a paper.

DATASETS

The reviewer assignment problem system is implemented using python. We have implemented and tested our algorithms on 1GHz, single-core CPU; 8 GB RAM. Performance of proposed algorithm is computed on InterSpeech, NIPS 2019 and AAAI 2014 conferences dataset [9]. From collected data, 679 papers and 220 reviewers were used to evaluate our proposed method. For building reviewer profile, we collected data from academic resources such as the DBLP Computer Science Bibliography, ResearchGate, and CiteSeer, that provide easy user access. Each reviewer was distinguished based on their name, Affiliation, publications, and network [10-12]. Unique identification numbers are assigned to all these papers.

Table 1: Data set for papers

Dataset(Papers)			
Name of Conference	Number of Papers	Fields Used	Format
AAAI 2014	398	Paper-id, Title, authors, Keywords, abstract	CSV format where each row is a paper and each column an attribute.
Available at UCI Machine Learning Repository https://archive.ics.uci.edu/ml/datasets/AAAI+2014+Accepted+Papers			
NIPS 2019	1425	Paper-id, Title, authors, abstract, Introduction, conclusion	We created using available papers
InterSpeech	145	Paper-id, Title, authors, abstract, Introduction	We created using available papers
https://papers.nips.cc/book/advances-in-neural-information-processing-systems-32-2019			

Table 1 provides details for papers to be reviewed and table 2 indicates details for reviewer's database.

Table 2: Data set for Reviewers

Dataset (Reviews)							
Number of Reviewers	220						
Reviewers count of publications	760						
Publications fields	Publication-ID, Year, Title, Abstract, Keywords, Introduction, Conclusion						
Publications Years Span	2007-2020						
Sample Reviewer's Details							
ID	Name	Affiliation	Research Interest	Publications	h-index	i-10 index	Citations

R00004	Andrea Tacchetti	Research Scientist - DeepMind	Artificial Intelligence, Statistical Learning, Computer Vision	35	11	12	1292
R00005	Aaron Klein	Amazon Research Berlin	Machine Learning, Bayesian Optimization, Deep Learning	31	15	17	1960
R00006	Ahmed M. Alaa	University of California, Los Angeles (UCLA)	Machine Learning, Deep Learning, Statistics, Artificial Intelligence, AI for Medicine	86	15	21	545
R00009	Alexander Ritchie	University of Michigan	Machine learning, algorithmic fairness	3	1	1	119
R00104	Dr. Parikshit Mahalle	University of Pune	Internet of Things, Identity Management, Access Control, Security in wireless communication, Algorithms	20	15	15	1283
R00106	Dr. Varsha Patil	University of Pune	Parallel Computing, Evolutionary Computing, Image Processing	67		7	324
R00058	Konstantin Mishchenk	KAUST	Optimization, Machine Learning, Deep Learning	19	8	7	197
R00063	Marco Fraccaro	Technical University of Denmark	Machine learning	17	9	9	472
R00105	Dr. Sachin Sakhare	University of Pune	Operating System, Data Mining, IOT	15	3	2	136

RESULT AND ANALYSIS

Experimentation results are presented for testing performance using the 679 papers and 220 reviewers as dataset. To restrict overlap among topics, 25 topics were selected. The number of iterations in the experiment was 3000, and further to debug the model parameters, initial 1000 iterations were used. The table 3 shows the Identified top 4 expertise as topic and their weights for sample reviewers. These are computed using their publications.

Table 3: Identified Top 4 Topics and associated weights

Reviewer		Identified Top 4 Topics and associated weights			
Reviewer_id	Name	1st Top Topic/ Weight	2nd Top Topic/ Weight	3rd Top Topic/ Weight	4th Top Topic/ Weight
R00004	Andrea Tacchetti	7	4	3	1
		0.461	0.242	0.171	0.072
R00005	Aaron Klein	3	7	4	--
		0.691	0.138	0.134	
R00006	Ahmed M. Alaa	7	3	4	5
		0.373	0.317	0.127	0.067
R00009	Alexander	7	3	4	5

	Ritchie	0.557	0.302	0.074	0.044
R00051	Jackson Gorham	3	7	4	2
		0.745	0.184	0.042	0.013
R00058	Konstantin Mishchenk	3	7	4	--
		0.612	0.213	0.151	
R00063	Marco Fraccaro	3	4	7	5
		0.437	0.235	0.209	0.094
R00104	Dr. Prikshit Mahalle	4	7	3	5
		0.492	0.339	0.113	0.031
R00105	Dr. Sachin Sakhare	4	7	3	--
		0.412	0.308	0.257	
R00106	Dr. Varsha Patil	4	7	2	3
		0.369	0.278	0.187	0.158

Topics are extracted from all papers and some of them with paper identification number and their respective paper topic weight; 10 topics are extracted from each paper are considered here. Similar to paper, the prominent topics and weight per topic is computed for experts too. These topics are computed using their publications. The similarity is computed using manuscripts and reviewer's publications. Once similarity is computed, after sorting top 5 as the most relevant experts matching the paper topics are listed in table 4 for few papers. Table 5 shows the title of paper and topic number with maximum topic weight.

Table 4: Sample Papers and Reviewers

Paper	Identified Reviewers				
paper_id	1	2	3	4	5
8296	R00005	R00006	R00009	R00004	R00106
8297	R00004	R00006	R00105	R00106	R00104
8298	R00006	R00005	R00009	R00004	R00106
8299	R00006	R00005	R00004	R00009	R00106
8300	R00005	R00006	R00009	R00058	R00004
8301	R00004	R00006	R00009	R00106	R00105
8302	R00006	R00009	R00004	R00106	R00105
8303	R00006	R00009	R00004	R00106	R00005
8304	R00006	R00004	R00104	R00105	R00106
8305	R00005	R00006	R00009	R00004	R00106
9724	R00104	R00106	R00006	R00105	R00004
9725	R00105	R00106	R00104	R00006	R00004
9726	R00004	R00104	R00106	R00006	R00105
9727	R00006	R00004	R00105	R00106	R00104
9728	R00004	R00006	R00009	R00106	R00105
9729	R00006	R00004	R00106	R00104	R00009
9730	R00006	R00106	R00004	R00105	R00009

9731	R00006	R00004	R00009	R00106	R00105
9732	R00006	R00106	R00105	R00004	R00005
9733	R00006	R00005	R00106	R00004	R00009
9734	R00006	R00005	R00106	R00004	R00009

Table 5: Topic number with maximum Paper Topic Weight

Paper id	Title of Paper	Topic having maximum topic weight	Weight of Topic
8298	Stochastic Shared Embeddings: Data-driven Regularization of Embedding Layer	Topic #3	0.462
8302	Stand Alone Self Attention in Vision Models	Topic #7	0.440
9725	Big Data in IoT	Topic #4	0.444
9726	Automated Enhanced Learning System using IOT	Topic #3	0.385
9733	A Diffusion Process for Wavelet-Transform-based Image Denoising and, its Application to Document Image Binarization	Topic #3	0.937

Reviewer identification number and their respective paper topic weight; when 10 topics are extracted from each paper are as shown in table 6.

Table 6: Reviewer ID with Maximum Paper Topic Weight

Reviewer id	Name of Reviewer	Topic Number having maximum weight	Weight of Topic
R00104	Dr. Parikshit N. Mahalle	Topic #4	0.492
R00105	Dr. Sachin Sakhare	Topic #4	0.412
R00006	Ahmed M. Alaa	Topic #7	0.373

Topic weights similarity index of top 5 reviewers and paper is as shown in table 7.

Table 7: Topic weights similarity index of top 5 reviewers and paper

paper Id	R1	Similarity for Paperweight	R2	Similarity for Paper weight	R3	Similarity for Paper weight	R4	Similarity for Paper weight	R5	Similarity for Paper weight
8298	R00004	0.9083	R00005	0.8378	R00009	0.7352	R00004	0.7198	R00106	0.6581
8302	R00006	1.3372	R00009	1.2183	R00004	1.2096	R00106 R00105	0.9832	R00105	0.9801
9725	R00105	0.8343	R00106	0.7728	R00104	0.7646	R00006	0.7216	R00004	0.6711
9726	R00004	0.97735	R00104	0.9750	R00106	0.91035	R00006	0.8683	R00105	0.8617
9733	R00006	1.3040	R00005	1.1940	R00106	1.1346	R00004	1.1033	R00009	1.0097

From table 3, it is observed that, for paper id 8298, 8302, 9725, 9726, 9733, reviewer Andrea Tacchetti (id-R00004), Ahemad M. Alla (id-R00006), Dr. Sachin Sakhare(id-R00105), Dr. Parikshit N. Mahalle (id- R00104), and Ahemad M. Alla (id-R00006), are respectively assigned. Results were presented to experts to seek opinion on grade of 5, (5-stong, 4- moderate, 3- average, 2-low, 1-poor) for correlation between generated topics with manually classified topics. It is observed that proposed system generated topics are strongly correlated.

CONCLUSION

In this paper, the problem of paper-reviewer assignment is addressed satisfying constraints. Among recent manuscripts, most of the papers are multi-domain and reviewer's expertise is in multiple domains. The proposed system, extraction of research topics is done by processing the publications of reviewer and submitted manuscripts jointly and trains model to learn extracting topics of papers and expertise of reviewers together. Experimental results demonstrate that the proposed approach assigns reviewer to a paper more effectively and more efficiently by covering all domains. Further the proposed algorithm can be used for various applications such as funding proposals, grouping digital library documents and similar.

ACKNOWLEDGMENT

REFERENCES

1. Sixing Ya, Jian Jin, Qian Geng, Yue Zhao, and Xirui Huang(2017), "Utilizing Academic- Network – Based Conflict of Interests for Paper Reviewer Assignment", International Journal of Knowledge Engineering, Vol. 3 No. 2, pp 65—73
2. Kalmukov, Y(2012). "Describing papers and reviewers' competences by taxonomy of keywords", Computer Science and Information Systems, 9(2), pp.763–789.
3. M. Karimzadehgan, C. X. Zhai, and G. Belford (2008), "Multi-aspect expertise matching for review assignment", in Proc. ACM conference on Information and Knowledge Management, pp 1113-1122.
4. Price, Simon, & Flach, Peter A.(2017), "Computational support for academic peer review: A perspective from artificial intelligence", Communications of the ACM, 60(3), pp. 70–79.
5. Pesenhofer, A., Mayer, R., &Rauber, A.(2006), "Improving Scientific Conferences by Enhancing Conference Management System with information mining capabilities", Proceedings IEEE International Conference on Digital Information ,pp. 359–366, ISBN: 1-4244-0682-x;
6. Ferilli, S., Di Mauro, N., Basile, T. M. A., Esposito, F., & Biba, M.(2006), "Automatic topics identification for reviewer assignment", 9th international conference on industrial, engineering and other applications of applied intelligent systems, IEA/AIE, pp. 721–730, Springer LNCS.
7. Charlin, L., & Richard Z.(2013), "The Toronto paper matching system: An automated paper-reviewer assignment system", Proceedings of the 30th international conference on machine learning, Atlanta, Georgia, USA, Vol. 28.
8. Liu, X., Suel, T., & Memon, N., "A robust model for paper reviewer assignment"(2012), In Proceedings of the 8th ACM conference on recommender systems, pp. 25–32.
9. Dinh Tuyen Hoang, Ngoc Thanh Nguyen, and Dosam Hwan(2019), "Decision Support System for Assignment of Conference Papers to Reviewers", @Springer Nature, Switzerland AG, pp. 441- 450.
10. Ivan Stelmakh, Nihar B. Shah, Aarti Singh., ArXiv (2019) "PeerReview4All: Fair and Accurate Reviewer Assignment in Peer Review", Proceedings of Machine Learning Research, 828-856
11. Yichong X, Han Z, Xiaofei S, Nihar Shah. (2018) On Strategy-proof Conference Peer Review
12. Hongwei Peng, Haojie Hu, Keqiang Wang, and Xiaoling Wang. (2017) Time-Aware and Topic-Based Reviewer Assignment Book Chapter @ DASFAA 2017 International Workshops: BDMS, Springer International Publishing, 145–157.
13. Hua Zhao, Wei Tao, Ruofei Zou, and Chunming Xu. (2018) Construction and Application of Diversified Knowledge Model for Paper Reviewers Recommendation", Springer Nature Singapore
14. https://archive.ics.uci.edu/ml/datasets/NIPS+Conference+Papers+1987-2015
15. http://dblp.uni-trier.de/xml/.
16. https://www.researchgate.net/.
17. http://csxstatic.ist.psu.edu/downloads/data