
Reviewer Assignment Using Weighted Matching and Hungarian Algorithm

Ali Qays Abduljaleel¹, Mohammed Abdullah Naser², Safaa O. Al-mamory³

¹Dept. of Computer Science, University of Babylon, Babylon, Iraq. aliqais85@gmail.com

²Dept. Of Computer Science, University of Babylon, Babylon, Iraq. wsci.mohammed.abud@uobabylon.edu.iq

³College of Business Informatics, University of Information Technology and Communications, Baghdad, Iraq. salmamory@uoitc.edu.iq

Article History: Received: 10 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 16 April 2021

Abstract: After closing submission system of the conference, the organizing committees' duties are huge and among these tasks are assigning at least three reviewers (at least) for each paper. This task involves matching a set of papers against reviewers in order to assign most suitable group of reviewers to specific paper to be reviewed. In this paper, a proposed weighted similarity measure is computed between each pair of (reviewer, author) work. This can be done by keeping a sample of published papers for each reviewer in order to extract the field of expertise of the reviewer and measure the relevance with the author's paper. The proposed weighted similarity measure is based on dividing each paper of authors and reviewers into five sections (i.e. Title, Abstract, Keywords, References, and the rest text in the paper) to compute the matching degree; each paper part is based on its importance thus TF-IDF and cosine similarity techniques are adopted to calculate the degree of correspondence between each pair of coincidence parts. Hereafter, Hungarian optimization algorithm is employed to assign each paper to the most relevance reviewers using computed similarity measure. The experimental results on **NTICT 2018** dataset showed that the proposed method achieved its goal.

Keywords: Reviewer Assignment Problem, Text Mining, TF-IDF, Cosine Similarity, Hungarian Algorithm.

1. Introduction

The most important thing in organizing scientific conference is the assignment of paper to proper reviewer. It plays a crucial role in reputation of conference and hence the quality of the evaluation. Even a small inaccuracy in the assignment may cause serious misjudgments. Manually assignment for large number of papers and reviewers is hard to Program Committee and time consuming to assign each paper to appropriate reviewers due to the many constraints (conflict of interest, specific specialization, load balancing, etc.). In addition, the manual method decreases its accuracy as the number of research papers increases so finding a way to do this automatically is necessary for conference quality and to achieve fairness without bias. Automatic review assignment was appeared first time by the study of Dumais and Nielsen [1]; it is continuing to develop by researchers until recently. In the literature, there are many strategies for Reviewer Assignment problem (RAP) which use full text or specific part of paper to compute relevance with some publications of reviewer.

In this paper, new approach is proposed which can be summarized in three steps. First, specifying the research area of each reviewer by considering specific number of last published papers (for each reviewer) where each paper represents research area of that reviewer separately (without combining them); this step will give a reviewer number of different chances to get most relevance author's paper. Second, calculating matching degree between each pair (reviewer, author) papers by dividing each paper into five parts (title, abstract, keywords, references, and the rest text in the paper) and gives each part a weight represent its importance to compute end matching result. Third, apply Hungarian algorithm[4] to get optimized results.

The calculating of matching degree is very important part in RAP because the assignment step depends on matching results. TF-IDF[2] is used to give weight for each term in specific document, which represents the importance of each term, most representative terms in specific document given high weight whereas low important terms that repeated several times in most documents will take low weight value; this technique was used by Hettich et al [3]. After using TF-IDF and converting the text of all documents to numeric vector, cosine similarity with weight for each part mentioned above was applied to measure how two specific documents are similar to each other taking into consideration no matching degree computed between author's paper and reviewer's paper having conflict of interest between them and substitute zero value instead. After constructing matching degree matrix (where columns represent authors' paper and rows represent reviewers' papers), preparing matching degree matrix by repeating its columns to specific number depending on constraint value (number of reviewers assigned to each paper) and make the matrix equal dimension to apply assignment stage. This approach gave a higher performance

when tried it in our datasets. Then matching degree was used between each paper and reviewer to assign most relevance paper to a proper reviewer by using Hungarian algorithm[4].

This paper is organized as follows. Section 2 presents related work with our work. Section 3 describes proposed system. Section 4 explains experimental results. Finally, Section 5 presents conclusions.

2. Related Work

Reviewer Assignment problem has been extensively studied; it is very common in academic journals and conferences. Reviewer assignment problem can be categorized into two main techniques which are preference-based and automatic relevance dedication.

The first direction of research try to solve reviewer assignment problem by Preference-Based approach or bidding data from reviewers[5]. In most preference-based approaches, it is required from reviewers to bid papers to see whether they have their interest for the papers or not. Goldsmith et al.[6] use Network Flow, and apply any polynomial time algorithm to compute the maximum flow through the network. Rigaux [7] use collaborative filtering techniques to grow the preference by asking users to bid on most papers in a given topic. The basic assumption of collaborative filtering techniques is that reviewers who bid similarly on a number of the same papers have likely the similar preference for other papers. The inadequacy of bidding information and consume additional time and hand-work are the main drawbacks of this technique. The proposed system is different form these approaches because it adopts automatic relevance dedication technique between authors and reviewers using text mining and optimality assignment.

The second direction of research employ automatic relevance dedication technique to solve reviewer assignment problem. This automated method can be classified according to the form of model for document comparison into the following groups.

- **Models that are feature-based:**The first research presented in this field was by Dumais et al. [1] where they use automatic feature extraction by latent semantic analysis to extract features. Jin et al.[8] use ordered weighted averaging to aggregate and rank candidate reviewers. Li et al.[9] use latent semantic indexing to measure the suitability score between a paper submitted and representative papers of a reviewer. Mimno et al.[10] use machine learning techniques to help predict reviewer expertise using Author-Persona-Topic. Tang et al. [11] employ optimization framework (i.e. convex cost flow problem) which guarantees an optimal solution under given constraints. Hettich et al.[3] presents empirical experiences (at the National Science Foundation) and uses the term frequency-inverse document frequency (TF-IDF) vector space model to annotating the proposals submitted with a vector of the top 20 representative words.
- **Probabilistic models:** Laurent et al. [12] and Campbell et al. [13] use Latent Dirichlet Allocation (LDA). LDA require a large corpus to accurately identify the topics and topic distribution in each document, which can be problematic when applied to short documents, such as abstracts.
- **Models Focused on Embedding:**The use of Word Mover's Distance (WMD) is another method to capturing similarities between documents. WMD depends on the alignment of word pairs from two texts, and the textual dissimilarity is calculated as the total distance between the word pair vectors[14]. Another work use neural network and word embeddings instead of word distributions [15].
- **Models based on Graphs:**All the above models only presume access to the texts of the submissions and publications of the reviewers. Liu et al.[16] Using a graph model to catch academic ties between reviewers and demonstrate that such information improves the matching output. Each node in the graph is a reviewer. If the respective reviewers have co-authored articles, there is an edge between two nodes and the edge weight is the number of publications. LDA does this work to determine the similarities between the submission and the reviewer. Cagliero et al. [17] use weighted association rules to recommend of external additional reviewers based on the extracted co-author relationships. Li et al. [18] suggest to calculate the correlation between the research field of the reviewer and paper and use a new approach via the reference information. Long et al.[19] research the impact of conflict of interest on fairness when topics have been used to match reviewer papers.

3. Conference Constraints

Various conferences have different constraints regarding reviewing policy like reviewing form, number of reviewers per paper, number of papers assigned for each reviewer, etc. Subsection 3.1 presents data acquisition process and data representation. Subsection 3.2 states how to address conference constraints related to reviewers and papers.

3.1 Data Acquisition and Representation

Two datasets were used in this paper. The first dataset is obtained from the organizing committee of 3rd *International Conference on New Trends in Information and Communications Technology Applications (NTICT 2018)*¹ which is held by the University of Information Technology and Communications from Iraq. The number of received papers (i.e., valid papers) is 87 including papers from different continents. Few of these papers were out of the conference scope.

The second dataset is about reviewers. This dataset is built manually by downloading three papers for each reviewer which reflect reviewers experience field(s). We select three papers since most conferences used this number of reviewers for any paper at minimum. It is assumed here that the number of downloaded papers for each reviewer is equal to the number of assigned reviewers for each paper. The number of reviewers for the NTICT’18 conference was 127. This dataset contains information about 381 papers (i.e. 127×3) which were retrieved from Google Scholar². Three Conflict of Interests were detected after going through each paper in the two datasets.

Based on these two datasets, we went through each paper and retrieved the title, abstract, keywords, references, and rest of paper except the other retrieved information. For the abbreviation purposes, the papers from first and second datasets will be denoted P and RP, respectively. The extracted information from these two datasets is represented by a 2D *matching degree matrix* in which the columns represents papers to be reviewed (P) and rows represent reviewers’ papers (RP). Table (1) states the mentioned information where (d) is number of reviewers, (m) is number of published papers for each reviewer.

Table 1: Matching degree Matrix

	P1	P2	Pi	Pn
R1	R1,1	M1,1,1	M1,1,2	Mj,k,i		
	R1,2	M1,2,1	M1,2,2	Mj,k,i		
	R1,3	M1,3,1	M1,3,2	Mj,k,i		
R2	R2,1	M2,1,1	M2,1,2	Mj,k,i		
	R2,2	M2,2,1	M2,2,2	Mj,k,i		
	R2,3	M2,3,1	M2,3,2	Mj,k,i		
	Rj,k	Mj,k,1	Mj,k,2	Mj,k,i		
		
	...					
	Rd,m	Md,m,1	Md,m,2	Md,m,i	...	Md,m,n

3.2 Reviewers vs Papers

In this paper, we assume that each paper (P) is reviewed by three reviewers and each reviewer shouldn’t review more than three papers (i.e. reviewer workload). Achieving these constraints with Hungarian algorithm [4] is a challenge task because this algorithm is designed to assign each column (i.e. a paper to be reviewed) to specific row (i.e. a reviewer). In other words, assigning each unique row to each unique column will not guarantee achieving these conference constraints.

The reviewers’ dataset includes three papers for each reviewer in order to guarantee assign at most three papers for each reviewer. The representation in the last subsection will be used in assignment stage to achieve conference constraint about (reviewer workload). Another advantage of this formulation is that it takes into consideration if reviewer has multidiscipline in his research area. In other words, this method gives the reviewer three opportunities to get best match with specific author paper. As shown in matching degree matrix, each reviewer has three consecutive rows and each author paper has one column.

From another point of view, At least three reviewers should review each article. The proposed approach to overcome on this challenge is to repeat each column in matching degree matrix three times. This operation will make Hungarian algorithm[4], in assignment stage, assign same paper “repeated paper” three times to different reviewers. Hence, every single paper will be assigned to most relevance three different reviewers.

¹ <http://www.uoitc.edu.iq/conf/index.html>

² <https://scholar.google.com>

4. The Proposed System

In this section we describe in detail our proposed system. In the first part of this section, we explain how Hungarian algorithm works. In the second part describe proposed Matching Degree Computation method. At the end of this section, we explain our Proposed (RAP) Technique.

4.1. Hungarian Optimization Algorithm

The Hungarian method is an algorithm for combinatorial optimization that Harold Kuhn developed and published in 1955. Originally, this technique was built for the best assignment of a group of people to a set of jobs. It is a special example of the transportation problem. The algorithm identifies an optimal allocation for a given "n x n" cost matrix. It consists of the following four steps. The first two steps are carried out once, while steps 3 and 4 are repeated until an optimum solution is found. The algorithm input is an "n x n" square matrix with non-negative elements only. The steps given below are applied to the n×n cost matrix to find an optimal assignment.

Step 1: All elements in the row are subtracted from the smallest element in each row.

Step 2: In each column, the smallest element is subtracted from all elements in the column.

Step 3: A minimum number of lines in the cost matrix are drawn to cover all 0.

Step 4: Optimality Test:

a) If n is the minimum number of lines, an optimum assignment and we are completed.

b) An optimal assignment is not yet feasible if the lines are less than n. Go to Step 5 in that case,

Step 5: The smallest entry not marked by any line is calculated. Subtract this entry from each uncovered row, and then add it to all elements that are covered twice. Return to Step 3.

If the assignment problem is maximization problem then must subtract each element in the matrix from maximum element in the same matrix before applying these steps above.

4.2 Matching Degree Computation

After aggregate previous publication of reviewers, matching degree computed between each paper of specific reviewer and authors papers using the proposed method, weighted similarity (WS). First, we dividing each paper for authors and reviewers to five parts (title, keywords, abstract, references, rest of the text) then Equation (1) is used to compute matching degree between each reviewer's paper and all authors' papers. This will be resulting matching degree matrix; its column represents author's papers and the rows represent reviewer's paper as in Table (1).

4.3 The Proposed RAP Technique

First of all, we take author's papers and put them in list and same thing we do with reviewer's papers. As mentioned in Section 3, the number of reviewers for each paper and number of papers for each reviewer is specified. The next step is dividing each paper to five parts (title, keywords, abstract, references, rest of the text paper). Hereafter, preprocessing and cleaning text is conducted on the results obtained from previous step, including remove stop words, punctuation and numeric values, convert to lower case letters and apply lemmatization operation. A basic step is to check if conflict of interest exists and remove it. Then, the weights are Calculated for each part of a paper. The next step is to compute matching degree matrix using Equation (1). After constructing matching degree matrix, take the largest value among three values resulting for each reviewer and make other two values equal zero. After that, each column of matching degree matrix is repeated number of times equal to number of reviewers we want to assign to each paper as we mentioned for constraints reason related to conference policy. Since Hungarian algorithm works on square matrix, dummy columns are added to matching degree. The steps of the main algorithm are proposed in Figure 1.

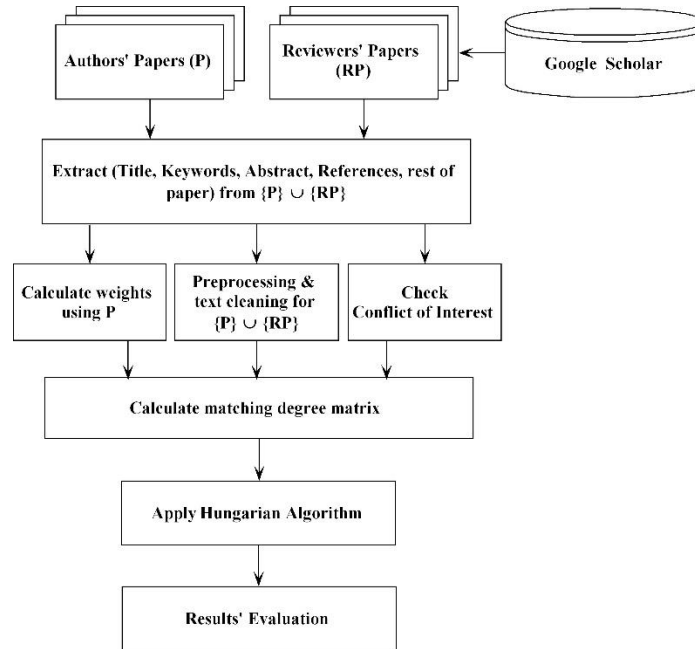


Fig 1. The main steps of the proposed algorithm

A new method is proposed to compute Matching degree, we called this method **Weighted Similarity (WS)**; it is described in Equation (1) where each cell in the matching degree is computed using the mentioned equation. In this way mentioned above, matching degree Matrix was constructed. This method makes matching degree more effective by using each segment of paper with its important representation and will not depend on specific part of paper because sometimes one part gives high relevance degree in some document and low in others so the proposed method balance relevance degree value using weights commensurate with importance of each part. Synchronize with this operation, conflict of interest is avoided using the data available about authors and reviewers (name, organization) and substitute with zero instead of computing matching degree ($M_{j,k,i}$) in case there is conflict of interest between (P_i) and ($R_{j,k}$) in order to avoid assignment between them.

$$M_{j,k,i} = (\sum s(seg_h(P_i), seg_h(RP_{j,k})) \times w_q) \dots\dots (1)$$

Where:

$M_{j,k,i}$: is matching degree be

S : is cosine similarity .

seg_h : is segment h of specific paper.

w_q : is a weight represent the importance for that segment (part) of paper.

5.Evaluation and Results

This section aims to present experiments conducted to evaluate the proposed system effectiveness by using subjective and objective evaluation measures. The dataset is obtained from 3rd International Conference on New Trends in Information and Communications Technology Applications (NTICT 2018)³ while the reviewers' metadata is gathered from Google Scholar⁴ by searching and keeping last three published papers for each reviewer. All experiments were developed using Python⁵ programming language. Table 2 presents few statistics about the mentioned dataset before and after assignment process.

³ <http://www.uoitc.edu.iq/conf/index.html>

⁴ <https://scholar.google.com>

⁵ <https://www.python.org>.

Table 2: Summarizes papers and assignment results

Inspection Stage	Criteria	Result
Before Assignment	Number of Conflict of Interest	3
	Number of author’s papers (valid papers)	87
	Number of reviewers	127
After Assignment	Number of Assignments	261
	Mean Number of Papers Assigned to each Reviewer	2.211
	Mean Number of reviewers Assigned to each paper	3
	Number of reviewers have no assigned papers	9
	Number of reviewers get 3 papers	50
	Number of reviewers get 2 papers	43
	Number of reviewers get 1 paper	25

Several preprocessing steps are applied on authors and reviewers’ papers. The first step is paper segmentation to produce five parts (i.e. title, Abstract, Keywords, Reference, rest of the text paper) for each paper. The second step is to clean texts from stop words and punctuation to exclude these tokens from matching operation. In addition, this step is needed because the proposed model is based on TF-IDF and Cosine similarity. Another step is converting all letters in text as lower case to avoid latter case difference between terms. Finally, lemmatization operation was applied to integrate the words that differs in form and similar in meaning by remove inflectional endings and to return the base or dictionary form of a word.

Moreover, TF-IDF [2] is computed for all documents that represent papers of authors and publication papers of reviewers. This is to obtain importance terms that most representative in each document preparing to use these factors in cosine similarity computation between documents to give us the relevance degree between compared documents.

Each part of the paper has its contribution when the similarity is computed between any pair of papers. These contributions are depicted in Figure 2 which, in addition, includes the proposed average similarity. It can be noted that the reference part achieved the highest value while the keyword part has the lower value. However, depending on average is, statistically, more stable than single value. The proposed value achieved third highest value which is excellent stable results.

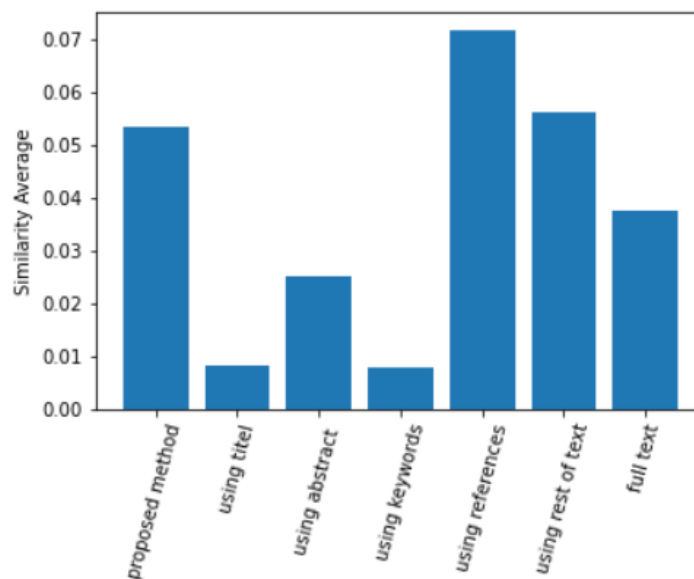


Fig. 2: Average similarity for each part of paper

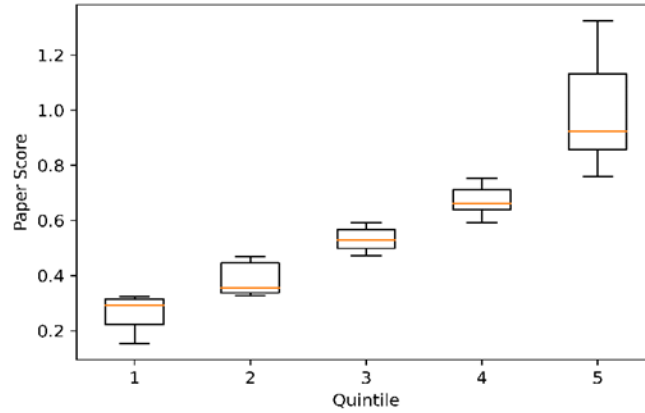


Fig. 3: Matching profile of papers' scores

The score quintiles for P papers are depicted in Fig. 3 as a *matching profile*. Paper score is summation of matching degree for three assigned reviewers to specific paper. In order to build paper profile, the paper score is computed for set P. Hereafter, the scores for all papers are sorted ascendingly and binned into five equal width bins. In the first box plot, the smallest horizontal line is described by the smallest paper score while the highest horizontal line is defined by the largest paper score in the last boxplot. The orange line in each boxplot represents the median paper score. This plot gives a visual overview of the paper score distribution, including the best and worst paper score.

A subjective evaluation of the proposed system is conducted. The results of the assignment process are reviewed by experts in order to ensure the correctness and quality of the assignment. Each expert scales every assignment from 1 to 5 where 5 is excellent assignment, 4 is very good, 3 is good, 2 is acceptable, and 1 is weak assignment. Hereafter, the results are averaged for each paper and for all assignments. The average scaling result is excellent for 37% of assignments, very good for 32%, good for 25%, acceptable for 5%, and weak for 1% of assignments. The Experts emphasized that the assignment process was wonderful. This result proves the validity of the proposed system.

The assigned reviewers to papers are classified to three groups after assignment stage is finish. First group is reviewers assigned to them three papers; second group contains reviewers assigned to them two papers for each one; one paper is delegated to each reviewer in the third group. After this separation we take average of matching degrees for each reviewer and it is assigned papers except third group in which we take whole matching value between each reviewer and it's assigned paper because they get one paper for each one. The result of these computation is depicted in Figure 4. This figure represents the relevance degree between author's paper assigned to each group of reviewers mentioned above. Note that the curve represents the third group of reviewers is higher than other two groups because one reviewer $R_{2,2}$ has high matching degree value with paper P_{18} and low matching degree with other papers. After inspection, the research area for $R_{2,2}$ and P_{18} are specialized in mathematics.

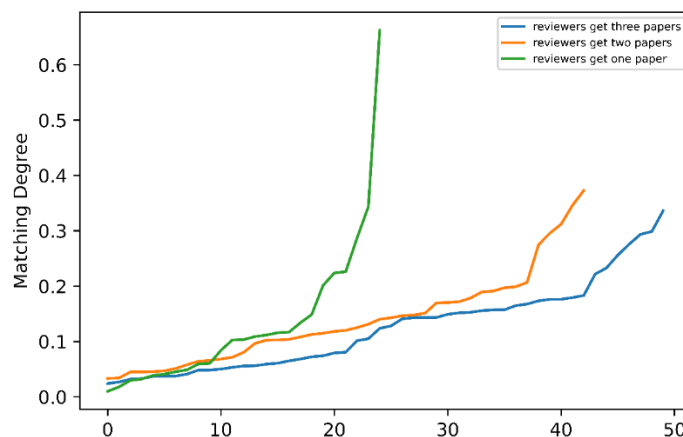


Fig. 4: Matching degree for each group of reviewers

Diversity factor is proposed to inform the conference committee how the submitted papers are close to each other or have different research area. To compute diversity factor, matching degree was computed between author's paper with each other using same method (WS) with one difference that is the rows and columns of matching degree matrix were author's paper. A unique matching between each paper and other papers is considered without self-matching; then compute the average value of these values to get end diversity factor value between (0 to 1) representing how submitted papers are different or close in research area. Figure 5 below explains the results of computing diversity factor for author's papers using different methods. As we see proposed method gives feasible results compared with other methods. The diversity factor has low value since the conference topics include papers from different fields in computer science which are computer networks, system and network security, machine learning, intelligent control system, communication applications, computer vision, and e-learning.

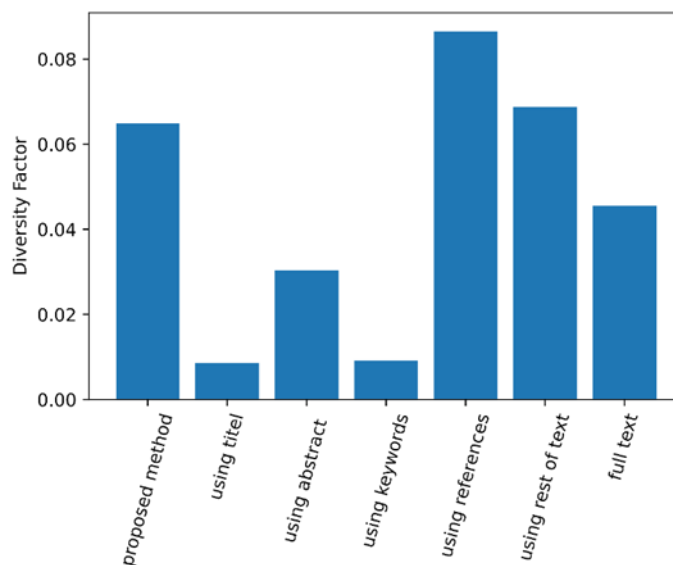


Fig. 5: Diversity factor of NTICT'18 conference papers

6. Conclusions

The problem of automatic assigning papers to reviewers is examined. The proposed system is based on using TF-IDF and cosine similarity with Hungarian algorithm in assignment stage to assign specific paper to optimal three reviewers depending on matching degree matrix. Choosing TF-IDF in our proposed system is to scales up rare and more informative terms in compared documents. This gives us relevance between documents not just naïve similarity between them. The proposed system showed strong and efficient performance.

References

1. S. T. Dumais and J. Nielsen, "Automating the assignment of submitted manuscripts to reviewer," SIGIR Forum (ACM Spec. Interes. Gr. Inf. Retrieval), no. June, pp. 233–244, 1992, doi: 10.1145/133160.133205.
2. G. Salton, "1988 Salton, G. and Buckley, C., 1988. Term-weighting approaches in automatic text retrieval_7896.pdf." 1988.
3. S. Hettich and M. J. Pazzani, "Mining for proposal reviewers: Lessons learned at the national science foundation," Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min., vol. 2006, pp. 862–871, 2006.
4. J. Robinson and S. Assignment, "14-16, 1951.," 1950.
5. N. Garg, T. Kavitha, A. Kumar, K. Mehlhorn, and J. Mestre, "Assigning papers to referees," Algorithmica (New York), vol. 58, no. 1, pp. 119–136, 2010, doi: 10.1007/s00453-009-9386-0.
6. J. Goldsmith and R. H. Sloan, "The AI conference paper assignment problem," AAAI Work. - Tech. Rep., vol. WS-07-10, pp. 53–57, 2007.
7. P. Rigaux, "An iterative rating method: Application to web-based conference management," Proc. ACM Symp. Appl. Comput., vol. 2, pp. 1682–1687, 2004.

8. J. Nguyen, G. Sánchez-Hernández, N. Agell, X. Rovira, and C. Angulo, “A decision support tool using Order Weighted Averaging for conference review assignment,” *Pattern Recognit. Lett.*, vol. 105, pp. 114–120, 2018, doi: 10.1016/j.patrec.2017.09.020.
9. 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>.
10. D. Mimno and A. McCallum, “Expertise modeling for matching papers with reviewers,” *Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, pp. 500–509, 2007, doi: 10.1145/1281192.1281247.
11. W. Tang, J. Tang, and C. Tan, “Expertise matching via constraint-based optimization,” *Proc. - 2010 IEEE/WIC/ACM Int. Conf. Web Intell. WI 2010*, vol. 1, pp. 34–41, 2010, doi: 10.1109/WI-IAT.2010.133.
12. R. S. Z. Laurent Charlin, “The Toronto paper matching system: an automated paper-reviewer assignment system,” *Icml*, vol. 28, 2013.
13. J. C. Campbell, A. Hindle, and E. Stroulia, “Latent Dirichlet Allocation: Extracting Topics from Software Engineering Data,” *Art Sci. Anal. Softw. Data*, vol. 3, pp. 139–159, 2015, doi: 10.1016/B978-0-12-411519-4.00006-9.
14. Matt J. Kusner, Yu Sun, Nicholas I. Kolkin, and Kil-ian Q. Weinberger. 2015. From word embeddings to document distances. In *Proceedings of the 32nd International Conference on International Conference on Machine Learning - Volume 37*.
15. O. Anjum, H. Gong, S. Bhat, W.-M. Hwu, and J. Xiong, “PaRe: A Paper-Reviewer Matching Approach Using a Common Topic Space,” pp. 518–528, 2019, doi: 10.18653/v1/d19-1049.
16. X. Liu, T. Suel, and N. Memon, “A robust model for paper-reviewer assignment,” *RecSys 2014 - Proc. 8th ACM Conf. Recomm. Syst.*, pp. 25–32, 2014, doi: 10.1145/2645710.2645749.
17. L. Cagliero, P. Garza, A. Pasini, and E. M. Baralis, “Additional reviewer assignment by means of weighted association rules,” *IEEE Trans. Emerg. Top. Comput.*, vol. X, no. APRIL, 2018, doi: 10.1109/TETC.2018.2861214.
18. X. Li and T. Watanabe, “Automatic paper-to-reviewer assignment, based on the matching degree of the reviewers,” *Procedia Comput. Sci.*, vol. 22, pp. 633–642, 2013, doi: 10.1016/j.procs.2013.09.144.
19. C. Long, R. C. W. Wong, Y. Peng, and L. Ye, “On good and fair paper-reviewer assignment,” *Proc. - IEEE Int. Conf. Data Mining, ICDM*, vol. 1, pp. 1145–1150, 2013, doi: 10.1109/ICDM.2013.13.