Multi-label Reviewer Profile Building and Ranking based on Expertise, Recency, Authority and h-index: Vital Module of Reviewer Paper Assignment

Aboli H. Patil[a]*, Parikshit N. Mahalle[b]
[a] Research Scholar, SKN College of Engineering, Pune, India, [b] Professor, SKN College of Engineering, Pune, India
[a]aboleee.patil@gmail.com, [b]aalborg.pnm@gmail.com

Article History: Received: 10 November 2020; Revised 12 January 2021 Accepted: 27 January 2021; Published online: 5 April 2021

Abstract: Reviewers serve as the key executors and prove to be the essential resources of the review process. The work aims to address and solve one of the problems associated with reviewer paper assignment problem that is identifying an expert with relevant knowledge of given research topic domain. We propose an LDA based model building for reviewer profiles with multiple domain labels using expert's existing publications. Further reinforcement of ranking them within the specific research domain is done based on number of publications in topic domain, recency/freshness, and h-index. We provide test results for a dataset of 107 reviewers with their 900 publications collected from DBLP and demonstrate that average Precision for top@5 topics with reference to manual annotations is 81% for 10 topic LDA model.

Keywords: LDA, reviewer profile, precision, topic model

1. Introduction

Recent decades have witnessed exponentially increased count of submissions at conferences and journals. Publications are an important part of research and academics; thousands of scientific conferences are organized each year and it encompasses the tedious and complex task of assigning reviewers to papers and getting the review process completed in time with the most accurate and fair reviews[24, 27]. It leads to the urgent need of providing a solution for the laborious and tedious task of domain wise clustering submitted manuscripts and assigning the most appropriate reviewers to these manuscripts. It is expected that each paper must be reviewed by a minimum N number of reviewers-Coverage, each reviewer should be assigned maximum M papers- load balance, and reviewer assignment should satisfy conflict of interest-CoI [16, 21, 34].

Most of the research these days are interdisciplinary and include more than one research domain. Interestingly it is noticed that experts too possess expertise in multiple domains. Challenge is to assign reviewers to manuscript so that at least one reviewer is assigned per research domains of manuscript for accurate reviews. Assigning reviewers to papers can be done manually but for a higher count of papers it is not a feasible task. Performing this task manually requires prior knowledge about expertise of each experts' expertise and recent research interest domains and recent publications in those domains [18, 26]. It is observed that based on bidding and registered interests of experts, the TPC chairs perform reviewer assignment tasks as an 'one-by-one-assign-as-you-go' manner, studying each paper and trying to match reviewers in a sequential way[5, 16 , 28, 39]. For a huge count of papers and dynamically updating the field of researcher's expertise domains, this task is too complex and time consuming [33].

Literature study reveals that the expertise and interests of experts can be extracted using sensibly selected domain specific keywords [3, 8, 29] . Reviewer profile building is a research domain subtle process that needs significant expertise and thoughtful efforts by the TPC chair. Study divulges that research papers include keywords, title and abstract that can help in extracting domains [17, 32, 40].

The conference management systems can request the reviewers to select from a list or add keywords to build their own profiles. But it is noticed that people may select the recent trends or may ignore their past expertise and over-emphasize recent interests [24, 37]. Study indicates that there is substantial work involved in creating fair and accurate reviewer and paper profiles. Bidding is prone to the same bias as self-profiling, for obvious reasons; and it needs each reviewer to look through an extensive list of submitted manuscripts [5, 24]. Popularly a person with sufficient knowledge and skills in specific research domains is referred as expert. Challenge is what basis we should utilize to measure and extract the expertise of person. One of the solutions is to process and identify the digitally available scientific publications of these persons. Though these publications portray the expertise, collecting and processing these publications manually is not feasible. One of the most widely used techniques to process the document corpus that is collections of scientific publications is information retrieval technique. We in our research work present a technique to build a profile of a reviewer that will help to match and assign the experts to papers easily and automatically without human expert intervention. Along with publications, it is obvious that we should refers to experts' CV, but it is noticed that often CVs remain not updated and become outdated, follow diverse formats making them difficult to process often missing with research domains in CV; so, we planned to use DBLP and Google Scholar for building reviewer profiles. We extracted the details as publication titles, year of publication, citations and h-index for each expert from Google Scholar.

Machine learning (ML) techniques are very efficient in solving real world problems so having vast applicability in numerous domains like health care, medical image processing, language modeling and many more [7]. One of the popular domains is information retrieval based techniques such as language modeling with Latent Dirichlet Allocation. The Latent Dirichlet Allocation (LDA) model is utilized for extracting research domains of experts. Experts with these extracted labels can be matched with paper for more accurate reviewer assignments. The key contributions of our research work include-

- As more than one expertise domains are determined, one can assure diversity of multiple domains reviewer assignments for each domain of manuscript for interdisciplinary and multidisciplinary domain submitted manuscripts.
- Among the most similarity matching reviewers further our work permits to rank the reviewers based on expertise, recency, number of papers in specific domain, and h-index.
- Proposed technique when used for experts' publications and manuscripts as one combined corpus, extracts the topics with higher accuracy.
- The outcome is support for the most accurate reviewer assignment of reviewer to manuscript.

2. Organization of paper

Rest of this paper is organized in six sections as- section 'motivation' presents motivation behind the work, literature related to research work in presented in section 'related work', details of proposed technique are described in section 'building profile of reviewer', experimental results are presented in following section and conclusions are presented in last section followed for acknowledgement and references.

3. Motivation

For academic conferences, agencies that provide research funding, and editorial boards of academic journals, fair and accurate reviews contribute for assuring equality [33, 38]. For fair and accurate reviews, appropriate selection and assignment of reviewers is necessary. The criteria for reviewer selection are based on relevant recent expertise knowledge. Publications of experts and databases of experts are not labeled with research domains. For reviewer paper assignment, the main task is to identify the expertise of researchers for matching with manuscripts. Most of the reviewer paper assignment computing depends upon human intervention. Ambient Computing is an environment that executes computations without a direct intervention of human. We propose technique that will serve as heart of reviewer assignment problem. Reviewer assignment serves as the key of conference and journal management systems. Our work will work behind the scenes as one of the features of ambient computing that will proactively recommend experts in anticipation of the paper reviewing process needs. Once the profiles of reviewers are built, it will serve as an inference engine in building experts' knowledge base and will also able predict probable knowledge domain changes [14]. Vast set of applications need accurate expert finding techniques as funding agencies, trainers for courses, reviewing papers, thesis and many more.

4. Related Work

Several studies have investigated and are available as published papers for reference for reviewer paper assignment. Available work can be classified in two broad classes as- Information retrieval based techniques and Optimization models based techniques.

Von Bakanic et al. reported the first paper in 1987 titled 'the manuscript review and decision-making process' stating that there is a need for techniques for automatic reviewer assignment. The work involved study of review processes and decisions of editorial regarding the manuscripts submitted to the American Sociological Review between 1977 and 1981 [37]. Researchers studied professional features of reviewer, characteristics of manuscript, procedures involved, and most importantly the recommendations of reviewer and the study revealed higher than 58 percent of the variance in the decisions. The key observation reported in about reviewer assignment and researchers reported that the assignment of referee had high influence and selection of referees and the count of revisions done had the greatest variance in the final decisions[29, 37]. These researchers were pioneered to report the need of automatic reviewer paper assignment for fair and accurate reviews.

Most of the popular research uses the information retrieval-based techniques that mainly work on the principle of computing the matching degree called as relevance believing in their level of knowledge and characteristics related to review ability. In other categories of information retrieval techniques, selection of reviewer is a multi-criteria decision analysis-based process that computes integrated score based on more than two features such as relevance, recency, authority and interest [11, 39, 40].

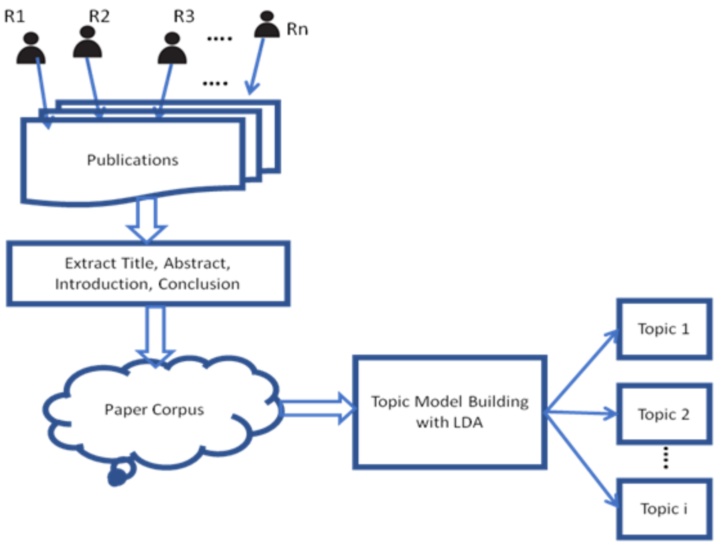
A comprehensive algorithm framework is proposed by Protasiewicz et al [13] to support information retrieval based techniques. This framework helps in retrieving classification of publication, keywords, and abstracts and it is further utilized to recommend from the blend of cosine similarity between full-text index [16]. For appropriate reviewers' selection for the given query with better performance, researchers have used TF-IDF (term frequency– inverse document frequency) [17] or a language model [5] and also a topic model [21] for building the experts' profile and profile of query documents. It is noticed that the research interests of the reviewers change over the time period. Hence the researchers Peng et al. have proposed a time– perception and topic-based reviewer assignment technique [11] .They have built archives for reviewers' professional knowledge using topic models as per recent publications of experts and by seeing the TF–IDF statistical characteristics of the submitted manuscript along with the quality earlier review tasks [17].

Often authors recommend the reviewer while submitting manuscripts expecting that they are relevant and experts of repute. In addition, experts are invited to serve as reviewers stating their domain of expertise and interest. Looking at this, Jin et al. [15] have proposed an integer linear programming technique for recommending reviewers based on the relevance and authority. Most of the published related work have utilized the domains/tracks/keywords selected by experts as priority preferences for review [9, 38]. The study reveals that very few of the work have given priority to the semantics of concepts while selecting the experts. Conventional optimization algorithms face difficulty in comparing huge textual data. Researchers in [30] state that the computing time is high for computing the relevance of reviewers with a proposal. Most of the researchers have used the LDA model for information processing [8, 12, 27].

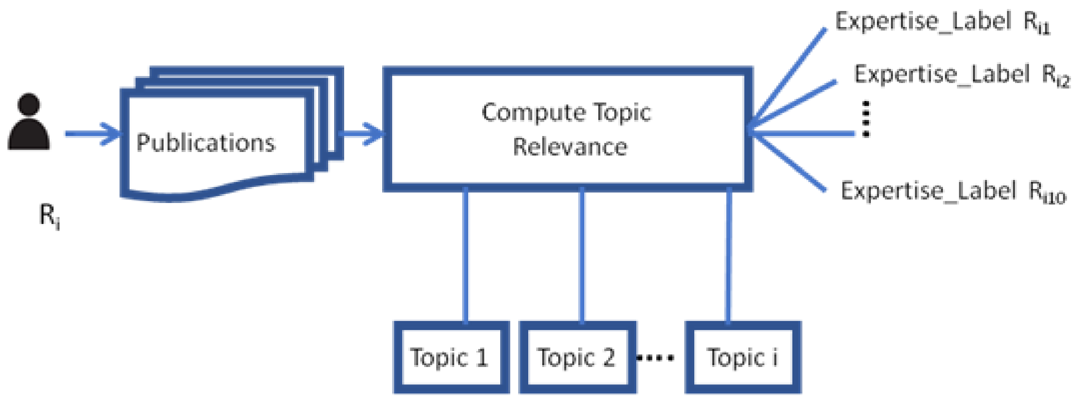
In this paper we propose a technique to build a reviewer profile using the LDA model. Automatic building of the reviewer profile that we propose here can be observed as a transitional step to optimized reviewer assignment assuming that there exists proximity/affinity between reviewers and manuscripts.

5. Building the Profile of Reviewer

LDA is one of the popular unsupervised machine learning techniques and it is used to find original topic evidence from a large corpus of documents [29]. LDA is a Bayesian model with three level hierarchical that is document–topic–word. LDA is based on two hypotheses: bag of words and bag of documents. The fundamental principle is that documents are made of infinite blends over latent topics and a topic is characterized by a spread of words [31, 35, 40].



(a) Extracting Topics from Corpus of Publications



(b) Labeling Expert with expertise domain labels

Figure 1. Process Flow diagram for reviewer profile building

Each domain specific notion is a combination of relatively few raw keywords. In other words, we can say that a given notion should not be a combination of the most keywords. Further each notion is expected to be sparse in keywords space.

One of the observations is person is expert in limited domains and sub-domains, the set of experts and set of manuscripts together possess a big count of keywords for fitting specificity [23, 31,36]. Assumption is expert may that each expert is related to a limited latent notion and even though the be ideally sparse in notion space which means a linear combination of a few notions; and it leads to higher sparsely that is to be exploited towards reducing latent space. Based on this we propose a technique to build a reviewer profile using Latent Dirichlet Allocation (LDA) model [27, 32].

We use LDA to extract the topics domains from publications of experts to compute the expertise of reviewers and with LDA lesser computational complexity by reducing dimensionality. Consider a reviewer who published P documents and each of these documents is made up of K keywords, and then a D dimensional standard corpus is built in which all unique keywords are aggregated. Further a linear vector represents reviewers' expertise, where each member of the vector indicates the total number of the corresponding keyword that appears in publications. This way the profile of experts is mapped to different weights indexed by the corpus. In LDA model-

- Total number of publications is 'M'
- Total number of keyword in document is 'N'
- The parameter of the Dirichlet prior for topic distributions per document is 'α'
- The parameter of the Dirichlet prior on the word distribution per topic is' β'
- The topic distribution for Document 'di' is 'θi'
- The word distribution for the topic 'k' is 'φ_k'
- The topic for the word 'wj' in document 'di' is 'Zij' and 'wij' is the specific word.

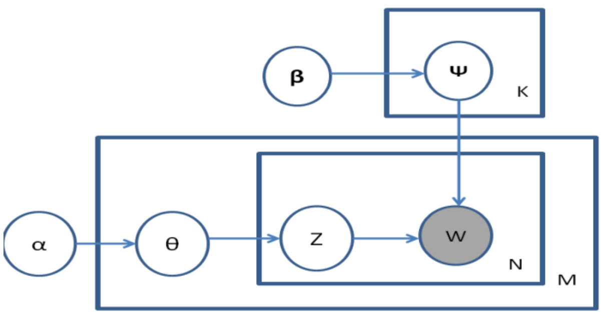


Figure 2. LDA topic model

LDA encompasses fully reliable generative semantics based on blend of topics spread as a 'k' parameter random variable [19, 28, 36]. In Fig. 2, random variables are indicated with vertices, shaded and highlighted vertices represent the identified and latent variables, respectively. Here edges indicate dependency among random variables, and directed edge indicates the dependency between variables, and plates indicate replicas. The documents are indicated by plate and topics are the inner plates representing the frequent collection of topics and keywords within a document. Working of LDA model can be described in four phases as-

Phase1:For every document, initialize the word to one of the 'K' topics randomly. Here '*k'* is prior decided value on the basis of approximately in how many research domains probably the set of documents be distributed. Further two frequencies can be computed—i) frequency distribution: the counts of topics in each document known as topic frequency and ii) frequency distribution: the counts of words in each topic known as word frequency.

Phase2: For every document 'D', and for every word say 'w' and compute: P(T |D). This is a proportion of words from D that are allocated to topic 'T'.

Phase 3: For every document 'D', and for every word say '*w'* and compute P(W |T). This is a proportion of allocated words to topic 'T' over all documents that include the word 'W'.

Phase 4: Reallocate the word 'W' with topic 'T' with the largest conditional probability P(T |D)´ P(W |T) considering all other words and their topic assignments.

We have used title, abstract and introduction section to represent the content information of the publications of expert for building the expert's profile. We have excluded the section keywords for two reasons- keywords section is often not accurately filled in and in most of the publications section 'keywords' is missing. We believe that the sections of paper- title, abstract, and introduction establish the candidate profile. The core domain topics of the expertise are extracted from the corpus of these sections. The algorithm is implemented in Python language. Following are the major steps of algorithm.

1. Extract title, abstract and introduction sections text of all publications
2. Tokenize words and Clean-up text
3. Create Bigram and Trigram Models
4. Remove stop words, Make Bigrams and Lemmatize
5. Create the Dictionary and Corpus for Topic Modelling
6. Building the Topic Model and View the topics in LDA model
7. Finding the dominant topic in each document
8. Find the most representative expert for each topic

Once the topics are extracted, for each topic the most relevant top 25 experts are extracted. We further rank these experts based on their count of publications in that domain, freshness/ recency, and h-index. This helps is assigning the most appropriate reviewer for the paper. The reviewer who has more publications in specific research domain that too in recent years and having higher h-index is the most authentic expert to review the manuscript in that domain.

6. Experimental Results

For reviewer paper assignment problem, there are few challenges. First the datasets availability- the datasets are not available as labeled data. None of the conferences provide paper-reviewer pairs. Secondly, the available manuscripts are not provided with their research domains. Thirdly, the reviewer's database too is not available with labels with their domains of expertise. We have extracted reviewers and their publications from DBLP. From publications further we extracted title, abstract and introduction sections text. All these are used to build corpus that is given as input to LDA model for topics building.

The core difficulty with expert's publications and processing to identify topics is measuring the performance as the publications are not labeled. There is no reference available to compare the output of proposed technique. We have used precision to measure performance of proposed technique. Here we use manual annotations for comparison. We asked the domain expert to mark the top @5 topic labels with grades as relevant or irrelevant [1, 2, 20, 23]. Then the precision is computed of each set of topic labels using the following formula (equation 1).

$$\text{Precision} = |P \cap Q| / |Q|(1) \quad \ldots\ldots\ldots \quad (1)$$

Here P denotes set of topics marked with 'Relevant' and Q denotes set of all extracted topics as top@5 or top@10 and so on.

Topic modeling with LDA is an investigative process as it extracts the hidden topic constructions in documents using a generative probabilistic process. These identified topics serve as inputs for further analysis. LDA uses Bayesian statistics and Dirichlet distributions through an iterative process to model topics [6, 10, 12, 22]. LDA works on principle of combined examination of topic distributions within corpus and word distributions within topics. This feature of LDA helps to identify the coherent topics with an iterative algorithm.

Table 1. Top 5 the most relevant reviewers for all 10 topics out of total 107 reviewers

Topic	**Rank**									
	1		**2**		**3**		**4**		**5**	
	Reviewer-ID and topic weight									
T-01	R104	0.552	R011	0.483	R040	0.383	R049	0.334	R068	0.280
T-02	R077	0.584	R031	0.455	R008	0.449	R033	0.434	R057	0.435
T-03	R020	0.527	R106	0.382	R010	0.313	R079	0.229	R025	0.203
T-04	R082	0.434	R004	0.388	R088	0.315	R028	0.293	R001	0.209
T-05	R055	0.524	R003	0.459	R029	0.394	R080	0.362	R026	0.357
T-06	R083	0.448	R040	0.361	R066	0.301	R044	0.284	R005	0.275
T-07	R097	0.531	R034	0.044	R037	0.366	R031	0.328	R061	0.283
T-08	R051	0.051	R101	0.416	R092	0.368	R089	0.359	R070	0.341
T-09	R056	0.682	R045	0.646	R093	0.607	R048	0.604	R038	0.592
T-10	R087	0.747	R046	0.579	R081	0.411	R062	0.354	R035	0.242

One of the most popular ranking metric is precision at k (P@k) where P@k is computed as the percentage of relevant reviewers in the top@k selections over total manuscripts [4, 18, 22, 25]. Table1 to 3 present result for topics=10.We provide test results for dataset of 107 reviewers with their 900 publications collected from DBLP. Once LDA model is build, we get 10 topics with bi-gram keywords as ‘machine learning’, ‘graph partitioning’, ‘data mining’ with weights. Further relevance of each reviewer and topics is calculated based on relevance related to experts publications. We select 25 most relevant experts per topic and rank them further based on their recent publications in that topic specific domain and h-index. For proper presentation only top five are presented in tables.

Table 1 presents top 5 relevant reviewers-reviewer ID and weight among top 25 selected reviewers. These reviewers are ranked from 1 to 5 based on weight for topic. These top 25 reviewers’ topic weights are again ranked based on total number of publications.

In table 2, these re-ranked reviewers are presented. From table 1, we notice that the reviewer R104 is at rank-1 and R040 is as at rank-3 computed based on relevance with topic 1 calculated from their published papers. When we rearrange them based on total number of publications, in table 2 the ranks are changed. We can notice in table 2 that R006 is now at rank-1 and R078 is pushed up at rank-2 as total publications are higher for reviewer-id R006. As total publications are higher, the reviewer R049 too is pushed up at rank-3.

Table 2. Top 5 reviewer for 10 topics out of total 107 reviewers selected based on highest topic relevance and maximum number of publications

Topic	Reviewer ID				
	1	2	3	4	5
T-01	R006	R078	R049	R101	R050
T-02	R078	R009	R005	R103	R062
T-03	R106	R009	R002	R004	R049
T-04	R106	R078	R002	R004	R005
T-05	R004	R094	R063	R055	R003
T-06	R106	R009	R002	R049	R005
T-07	R096	R009	R094	R003	R105
T-08	R002	R101	R089	R058	R070
T-09	R096	R094	R093	R059	R075
T-10	R006	R009	R004	R079	R104

Table 2 presents top 5 relevant reviewer ID among top 25 selected relevant reviewers again ranked based on total number of publications. But it is more important that most of the papers are in topic specific domain than higher count of publications as a whole. Further they are sorted and again ranked based on number of publications in that specific topic domain and put in table 3. Here we notice that as for reviewer R104 and R101, their recent publications are in topic-1 domain, they remain at rank-1 and rank-2 respectively.

Table 3. Top 5 reviewer for 10 topics out of total 107 reviewers selected based on highest topic relevance and maximum number of publications in topic specific domain.

Topic	Reviewer ID				
	1	2	3	4	5
T-01	R104	R101	R040	R026	R013
T-02	R008	R077	R064	R061	R021
T-03	R091	R079	R010	R106	R020
T-04	R088	R004	R028	R106	R082
T-05	R063	R055	R080	R074	R065
T-06	R083	R044	R040	R007	R105
T-07	R031	R057	R024	R018	R008
T-08	R051	R089	R054	R092	R070
T-09	R013	R084	R048	R032	R059
T-10	R062	R081	R074	R023	R092

Table 4 presents top 5 relevant reviewers ID among top 25 selected and ranked based on relevance with topic, total publications, and total publications in specific topic domain, freshness/recency and h-index. Here in table 4, we notice that reviewer R078 and R006 are at rank-1 and rank-2 respectively. We also notice that R104 and R101 are at rank-3 and rank-4 respectively.

Table 4. Top 5 reviewers for 10 topics out of total 107 reviewers selected based on highest topic relevance, maximum number of publications in topics specific domain in recent years and h-index

Topic	Reviewer ID				
	1	2	3	4	5
T-01	R078	R006	R104	R101	R050
T-02	R078	R009	R005	R021	R064
T-03	R002	R009	R104	R079	R004
T-04	R078	R002	R005	R004	R063
T-05	R094	R004	R072	R063	R091
T-06	R089	R009	R002	R005	R083
T-07	R096	R009	R094	R072	R024
T-08	R089	R002	R101	R014	R054
T-09	R096	R094	R093	R059	R012
T-10	R006	R009	R104	R079	R004

The table 5 is listing for number of topics values and corresponding LDA model perplexity and coherence. As a rule of thumb for a good LDA model, the perplexity score should be low while coherence should be high. As lowest perplexity and highest coherence is observed at topic=10, we selected LDA model with topic =10.

Table 5. Number of topics values and corresponding LDA model perplexity and coherence

Number of Topics	Perplexity	Coherence
3	-8.331189855	-1.052427613
5	-8.395786477	-5.213751281
7	-8.740134876	-4.077492887
10	-11.43009841	-1.039389374
12	-10.03012161	-4.468324876
15	-8.285181911	-1.096359289

Further experimentation is done by varying number of topics=3,5,7,10,12,15 and extracting top@k for k=2,4,5 and 7. Precision is computed by comparing resultant topic label by manual annotation by expert. And table 6 presents average precision values. It is observed that higher precision as 81.03 for topic=10 and top@5.

Table 6. Average precision@k for N topics

	Average precision @k			
Number of topics	P@2	P@4	p@5	p@7
3	63.56	65.30	69.00	64.43
5	66.79	68.78	72.67	70.56
7	69.45	71.09	79.34	74.66
10	73.60	75.00	**81.03**	78.89

12	71.34	74.88	78.99	76.00
15	70.12	73.90	77.54	76.78

Figure 3 is plot of Precision for Top@K by varying number of topics. We can notice that precision at number topics =10 and for top@5 is higher as compared to topics= 3, 5, 12 and 15 and top@2,4,and 7.

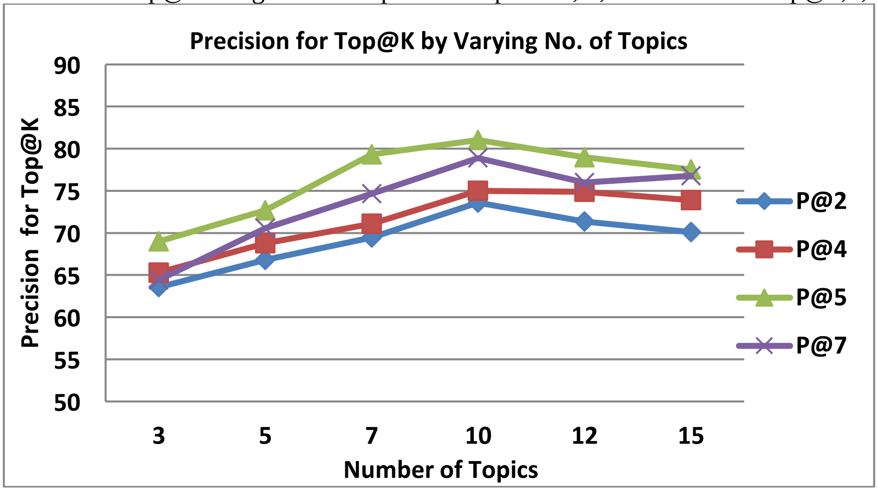


Figure 3. Plot of Precision for Top@K by Varying No. of Topics

7. Conclusion

Topic modeling is one of the most popular areas of natural language processing which processes the text when it is not labeled. Publications of experts and database of experts are not labeled with research domains. For reviewer paper assignment, the main task is to identify the expertise of researchers for matching with manuscripts. Our technique identifies multiple research domains of experts by processing corpus of their publications using LDA topic building algorithm. Corpus is built using sections of papers- title, abstract and introduction. Further the set of most relevant experts are ranked based on recency, number of publications in that domain, and h-index. This supports to assigning the most accurate reviewer to paper. Precision for top@5 topics with reference to manual annotations is 81%. The work can be further extended to build a combined corpus of submitted manuscripts and publications of experts to process it and provide multiple topic labels to experts and manuscripts.

8. Acknowledgement

We are very much thankful to research centre at department of Computer Engineering, Smt. Kashibai Navale College of Engineering, Pune for providing all resources and support to carry out this research work.

References

[1] A. H. Patil & P. N. Mahalle. (2020). Trends and Challenges in Measuring Performance of Reviewer Paper Assignmen. *Elsevier Journal, Procedia Computer Science*, 709-718.
[2] A.S. James, E. Lanham, V. Mak-Hau, L. Pan, T. Wilkin & G. Wood-Bradley. (2018). Identifying items for moderation in a peer assessment framework. *Knowledge Based System*, 162, 211–219.
[3] Ouni, R.G. Kula & K. Inoue. (2016). Search-based peer reviewer's recommendation in modern code review.*2016 IEEE International Conference on Software Maintenance and Evolution (ICSME), IEEE*,.367–377.
[4] Yesilcimen& E. A. Yíldírím. (2019) An alternative polynomial-sized formulation and an optimization based heuristic for the reviewer assignment problem", *European Journal of Operation Research*, 276 (2), 436–450.
[5] A.G. OrorbiaIi, T.Mikolov, & D. Reitter. (2017). Learning simpler language models with the differential state framework. *Neural Computing*, 9(12), 3327–3352.
[6] D. K. Tayal, P.C. Saxena, A. Sharma, G. Khanna & S. Gupta. (2014). New method for solving reviewer assignment problem using type-2 fuzzy sets and fuzzy functions. *Appl.*, 40(1), 54–73.
[7] Dey, N., Wagh, S., Mahalle, P. N., & Pathan, M. S. (Eds.),Applied Machine Learning for Smart Data Analysis. *CRC Press*.
[8] D. M. Blei, A.Y. Ng & M.I. Jordan. Latent Dirichlet Allocation, J. *Mach. Learn. Re*, 993–1022.
[9] D.M. Soares, M.L. De Lima Júnior, A. Plastino, L. &Murta. (2014). What factors influence the reviewer assignment to pull requests. *Inf. Softw. Technol*, 98, 32–43.

[10] G.S. Daş& T. Göçken (2014). A fuzzy approach for the reviewer assignment problem. *Comput. Ind. Eng*, 40, 50–57.
[11] H. Peng, H. Hu & K. Wang & X. Wang. (2017). Time-aware and topic-based reviewer assignment. *International Conference on Database Systems for Advanced Applications*, *Springer*, *Cham*,.145–157
[12] H. B. Jeon & S. Y. Lee. (2016). Language model adaptation based on topic probability of latent dirichlet allocation. *ETRI J*. 38(3), 487–493.
[13] Jaroslaw Protasiewicz.. (2014). A Support System for Selection of Reviewers. *IEEE International Conference on Systems*, *Man, and Cybernetics*, 3062-3065.
[14] JC Kim, K Chung (2020).Neural-network based adaptive context prediction model for ambient intelligence. *Journal of Ambient Intelligence and Humanized Computing*, 11, 1451–1458
[15] J. Jin, B. Niu, P. Ji & Q. Geng.(2018). An integer linear programming model of reviewer assignment with research interest considerations. *Ann. OperRes*, 1–25.
[16] J. Protasiewicz, W. Pedrycz, M. Kozłowski, S. Dadas, T. Stanisławek, A. Kopacza, M. Gałe&żewskaa. (2016). A recommender system of reviewers and experts in reviewing problems. *Knowledge -Based Systems*, 2016, 106, 164–178.
[17] K. Chen, Z. Zhang, J. Long & H. Zhang. (2016). "Turning from TF-IDF to TF-IGM for term weighting in text classification. *Expert Syst. Appl. Int. J.*, 66 (C)
[18] K. Zupanc, Z. Bosnić. (2017). Automated essay evaluation with semantic analysis*.* *Knowledge -Based Systems*, 120, pp.118–132.
[19] K.K. Janke, A.S. Bzowyckyj& A.P. Traynor. (2017). Editors' perspectives on enhancing manuscript quality and editorial decisions through peer review and reviewer development", *Am. J. Pharm. Education,* 81(4), 1–11.
[20] L. Li, Y. Wang, G. Liu & Y. Sun.(2015). Context-aware reviewer assignment for trust enhanced peer review", *PLoS One*, 10(6), 1–28.
[21] M. El-Assady, F. Sperrle, O. Deussen, D. Keim& C. Collins (2019).Visual analytics for topic model optimization based on user-steerable speculative execution", *IEEE Trans. Vis. Computer Graphics*, 25(1), 374–384.
[22] M. Karimzadehgan, C.X. Zhai& G. Belford. (2018). Multi-aspect expertise matching for review assignment. *Proceedings of the 17th ACM Conference on Information and Knowledge Management*, 1113–1122.
[23] M. Mirzaei, J. Sander & E. Stroulia. (2019). Multi-aspect review-team assignment using latent research areas. *Inf. Process. Manage*, 56(3), 858–878.
[24] M. Saleem & N.M. Minhas. (2018). Information retrieval based requirement traceability recovery approaches-a systematic literature review. *University of Sindh J. Inf. Commun. Techn*., 2(8), 180–188.
[25] M. A. Rodriguez & J. Bollen. (2008). An algorithm to determine peer-reviewers", Proceedings of the 17th ACM Conference on Information and Knowledge Management, *ACM*, pp.319–328.
[26] M.S. Wu.. (2015). Modeling query-document dependencies with topic language models for information retrieval, *Inform. Sc*, 312, 1–12.
[27] N.M. Kou, H.U. Leong & N. Mamoulis, Y. Li, Y. Li, Z. Gong.(2015). A topic-based reviewer assignment system. *Proc. VLDB Endow*, 8 (12), 1852–1855.
[28] O. Liu, J. Wang, J. Ma & Y. Sun. (2016). An intelligent decision support approach for reviewer assignment in R & D project selection, *Comput. Ind*, 76, 1–10.
[29] O. Toubia, G. Iyengar, R. Bunnell & A. Lemaire. (2019). Extracting features of entertainment products: a guided latent Dirichlet allocation approach informed by the psychology of media consumption.*J. Mark. Res*, 56(1), 18–36.
[30] P. Thongtanunam, R.G. Kula, A.E.C. Cruz, N. Yoshida & H. Iida. (2014). Improving code review effectiveness through reviewer recommendations. *Proceedings of the 7th International Workshop on Cooperative and Human Aspects of Software Engineering*, *ACM,* 119–122.
[31] R. Ramezanian& A. Ezzatpanah. (2015). Modeling and solving multi-objective mixed-model assembly line balancing and worker assignment problem. *Comput. Ind. Eng*, 87, 74–80.
[32] R. C. Chen, P. H. Shang & M.C. Che. (2012). A two-stage approach for project reviewer assignment problem. *Adv. Mater*. 452, 369–373.
[33] S. Bouajaja& N. Dridi.(2017). "A survey on human resource allocation problem and its applications",*Operations Research*, 17(2),. 339–369.
[34] S. Price & P.A. Flach (2017). Computational support for academic peer review: a perspective from artificial intelligence. *Communication ACM*, 60(3), 70–79.
[35] S. Zhao, D. Zhang, Z. Duan, J. Chen, Y. Zhang & J. Tang. (2018). A novel classification method for paper-reviewer recommendation. *Scientometrics*, 115(3), 293–1313.
[36] T. Kolasa& D. Krol (2011).A survey of algorithms for paper-reviewer assignment problem. *IETE Tech. Rev* 28(2), . 123–134.

[37] Von Bakanic, Clark Mcphail& Rita J. Simon. (1987). The Manuscript Review and Decision Making Process", *Journal of American Sociological Review*. 52(5),631-642.
[38] X. Li, T. Watanabe. (2013). Automatic paper-to-reviewer assignment, based on the matching degree of the reviewers. *Procedia Computer Science*,22, 633–642.
[39] Y. Xu, J. Ma, Y. Sun, G. Hao, W. Xu & D. Zhao.(2010). A decision support approach for assigning reviewers to proposals. *Expert Syst. Appl*, 37(10), 6948–6956.
[40] Z. Bo, H. Zhou, G. Li, Y & Huang, Zenlda. (2018). Large-scale topic model training on distributed data-parallel platform. *Big Data Min. Analysis, 1(* 1), 57–74.