PROPOSED FRAMEWORK FOR A PAPER-REVIEWER ASSIGNMENT SYSTEM USING WORD2VEC

OGUNLEYE Olawole, IFEBANJO Temitope, ABIODUN Theresa, ADEBIYI Ayodele
Department of Computer and Information Science, Covenant University, Nigeria

Abstract
The task of assigning papers to reviewers is crucial to the realization of a peer-to-peer review process of academic conferences. The manual process of ensuring submissions assigned to reviewers is related to their knowledge domain can be very cumbersome. Besides, poor quality reviews results from an ineffective assignment of papers. From extant literature, automated reviewer assignment systems based on distributional semantic models have been used to capture semantics with the shortcoming of limited in the bag of words models. Neural Network Language models have been used to eliminate the limitations of bag of words of models in expertise finding and product recommendation. Thus this paper proposes a framework based on neural network language models to derive suitability scores based on the semantic relatedness between a paper meant for review and a reviewer's representation papers. The present performance of the neural network language model compared to distributional semantic models used in solving reviewer-assignment. This ensures the semantic relatedness of paper and reviewer knowledge representation in allocating a paper, which improves the overall success of the peer-to-peer review process.
Keywords: Paper-reviewer assignment, Word Vectors, Neural Network Language Model, conference, Assignment Optimization

INTRODUCTION

Yearly, academic conferences receive a lot of paper submissions for publications, which must go through a peer review exercise before they can be accepted. The review process serves as a filter to sieve out quality papers from unsuitable papers for the conference. The crucial task of assigning submitted papers to reviewers is the starting point of the peer-review process, which is also known as paper-reviewer assignment (PRA). For the paper review to be fair and thorough the paper must be assigned to reviewers based on their knowledge domain, and are to some degree competent to review the paper.

The manual process of assigning large papers in an academic conferences whilst taking into cognizance the suitability of the reviewer would require a lot of time and laborious effort from the organizers. Therefore, this necessitates the automation of paper-reviewer assignment, which promises to save time and effort of the organizers for other tasks that will contribute to the overall quality of the conference. The approach to automating PRA can be undertaken as an information retrieval problem where a submitted paper is used as the query and each reviewer's set of published works typifies the document representations to be matched (Long*et al.,* 2013). The information retrieval model seeks to compute a matching score, which represents the relevance between the paper and the reviewer, which in turn

e-Governance Conference
Covenant University Conference on e-Governance in Nigeria - CUCEN2017

determines the suitability of the reviewer to review the queried paper (Li & Hou, 2016).

Drawing from literature, the state of the art information retrieval models used in PRA includes: Term Frequency-Inverse Document Frequency (TF-IDF) (Hettich & Pazzani, 2006), Latent Semantic Indexing (LSI) (Zablocki & Lee, 2012) and Latent Dirichlet Analysis (LDA) (Charlin & Zemel, 2013).A common fact about the mentioned approaches is that they are innately Bag-of-Words (BOW) models. Bag-of-Words view a document, such as the paper and reviewer's expertise representation as a set of terms where the frequency of each term is significant but the ordering of each of the terms is ignored (Manning *et al.*, 2008).But the shortcoming of the bag-of-words approaches is that they are likely to suffer from the query document mismatch. According to Li and Xu (2014), query document mismatch happens when both the searcher and author are using different expressions to describe the same meaning or vice versa. Moreover, bag-of-words approaches on large data sets often produces vectors with large sparseness, where the n-dimension of the vector representation of each document is equal to the n-size of the document vocabulary. This dimensionality problem is also known as the curse of dimensionality.

Word Vectors also known as Word Embedding are continuous-valued distributed representations for words with reduced dimensional space. They have recently being used by Neural Network Language Models in learning word sequence distributions to eliminate the curse of dimensionality plagued by bag-of-words models (Bengio *et al.*, 2003). Also, Word Vectors can handle semantic parsing of a corpus by extracting meaning from text to generate a language. Distributed word vectors are powerful and can be used for many applications, particularly word prediction, translation, and automated speech recognition (Mikolov *et al.,* 2010). Word2vec (Mikolov *et al.,* 2013) is a neural network implementation that learns distributed representations for words; it can learn the meaning of a word from its context and encoded into a vector of values. A document representation can be expressed as a matrix of the encoded word vectors. Cosine distance can be used to determine the semantic similarity between document vectors.

The aim of this paper is thus to implement a paper-reviewer assignment system using a Neural Network Language Model, Word2Vec and comparing it with the state of the art models used in information retrieval. The rest of this paper is structured as follows: Section 2 presents the review of related works. In Section 3, the methodology is presented. Section 4, discusses the implementation of the system as an application. Section 5 discusses the results obtained and Section 6 concludes the paper.

RELATED WORKS

The literature is toward real-world implementations and deployment of paper-reviewer assignments for academic conferences.

The study in Zablocki & Lee (2012), improves an existing automated assignment task of START, a web-based conference management system for managing every process of the

Governance and Public Service Delivery in Nigeria: The Role of Information and Communication technologies **CUCEN2017**

Ogunleye *et al.*

conference. The system has modules for collecting and storing the reviewers' information and papers in a database. Also the conference organizers can add reviewers to a conference and committees. Automatically matching paper submissions to reviewers uses a modified latent semantic indexing to achieve topic-topic comparisons. The system computes the total expertise score for a paper, which is the summation of the cosine similarities between the paper's topic and each primary expertise of the reviewer, multiplied by the weight given to primary expertise matches over secondary.

The Toronto Paper Matching System (TPMS), a framework presented in Charlin & Zemel (2013) hasbeen used in machine learning and computer vision conferences since 2010, such as Neural Information Processing Systems (NIPS), International Conference on Machine Learning (ICML), and International Conference on Computer Vision (ICCV) and being integrated into the Microsoft Research Conference Management Toolkit to handle paper submissions and review. The suitability score is computed from the unsupervised learning of the encoded vector representations of the reviewer's archives and the paper submissions using LDA model. Specifically, the suitability scores are computed from the probability of the normalized distribution on the word count for each word appearing in the reviewer's published work.

In addition, *Erie* (Li & Hou, 2016) is a system that supports the conference management for IEEE INFOCOM. The system was adapted from and integrated to an existing conference management software application, *EDAS*, which was lacking in fully automated assignments of paper to reviewer capability. The system integrates with the EDAS's remote database system, which has been used to collate the reviewer's information. The system uses LSI model to compute the suitability scores between the encoded vector representation of a reviewer's corpus and a paper. Particularly, LSI uses singular value decomposition (SVD) technique to reduce the sparseness of the vector representation. Then, the cosine similarity between the paper-reviewer vector matrices is computed to deliver the suitability scores. In summary, the uniqueness of our proposed work is the integration of Word2Vec to the system that will provide better semantic retrieval and assignment of papers to the reviewers.

METHODOLOGY

This section describes the proposed approach for paper-reviewer assignment using Word2Vec model. The following will show the different phases consisted in our methodology in determining the semantic similarity. Finally, we intend to compare the scores from Word2vec with other bag-of-words models such as TF-IDF, LSI, and LDA.

A)   Data Preparation

There is no publicly available dataset for paper-reviewer assignment that will sufficiently meet the set out objectives of this paper so there was a need to curate datasets from a typical large conference such as NIPS to test our model. NIPS is a machine learning conference that holds yearly since 1985. We were able to access the papers from 1985-2016 from their website. The reviewers' list will be 2016 reviewers list on the NIPS, which includes about

Ogunleye *et al.*

100 Area chairs. The 2016 set of papers from 1985 – 2016 dataset will serve as the query papers for assignment are up to 562.

The Dataset gotten are in raw form and would require text processing, which includes removal unwanted characters or non-printable, tokenization, elimination of numbers and punctuation, removal of stop words of the datasets and lemmatization. In modeling the reviewer's knowledge domain, we intend to concatenate each of the publication of a reviewer to form a corpus that would represent the reviewer. We would first download the paper in pdf format from their Google Scholar profile page and then we would use native tools such as *pdf to text* to extract the text from the downloaded files and concatenate them to form a single corpus.

B)   Overview of our Methodology

Central to the quest of AI is the ability of machines to estimate some likeness between things in the way a human is able to compare stimuli. In this regard semantic measures are being developed to determine a similarity between concepts according to the information supporting their meaning (Harispe *et al.,* 2016). Therefore, determining semantic similarity is an important technique in artificial intelligence, which is used to assist humans in determining how similar or related concepts.

Figure 1 shows the process flow of our methodology, showing the different phases towards determining the semantic similarity, which is later used for assignment.
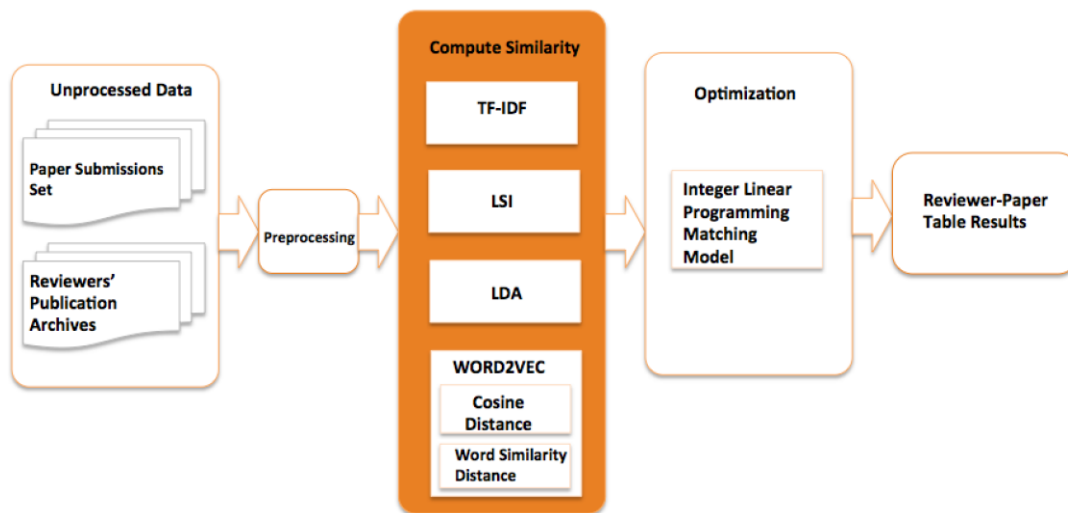

Fig. 1. Overview of the methodology

After building the reviewer's corpus, the 2016 paper set and the domain papers needed for training to develop the domain space, we would use the Word2Vec model in determining the vector encoded word embedding for each of the reviewers and for the papers. Then we find the cosine distance, which would use as the semantic similarity between the reviewer and the paper. This will be done for the 89 reviewers to the 562 papers. We would compare our

Governance and Public Service Delivery in Nigeria: The Role of Information and Communication technologies | **CUCEN2017**

Ogunleye *et al.*

results with the bag-of-word models such as TF-IDF, LSI and LDA.

C)   Assignment Optimization

In optimizing the assignment submitted papers to the most suitable reviewers, we plan to use the Integer Linear Programming formulation as presented by Taylor (2008) with an objective of maximizing the overall sum of suitability scores globally. This is subject to constraints that each submitted paper should be assigned to no more than a certain number of reviewers, and no reviewer should be assigned more than a maximum workload of submitted papers.

$$max \sum_r \sum_p S_{rp} \alpha_{rp}$$

s.t

$$\sum_p \alpha_{rp} \leq W_r^{Max}, \forall r$$

$$\sum_p \alpha_{rp} \geq W_r^{Min}, \forall r$$

$$\sum_p \alpha_{rp} = R_p^{Min}, \forall p$$

$$\sum_p \alpha_{rp} \in \{0,1\}. \forall r \quad \forall p$$

SYSTEM IMPLEMENTATION

The system will be implemented as a web application with a set of modules written in Python. The built-in data structures, and ready-to-use and available third-party tools made Python a first choice for implementing the six modules of the system. From Figure 2, which shows the proposed system architecture of the developed system consists of Paper Collation, Preprocessing & Repository Database, Similarity Computation Module, Assignment Optimization Module and Matching Results View modules.

Governance and Public Service Delivery in Nigeria: The Role of Information and Communication technologies | **CUCEN2017**
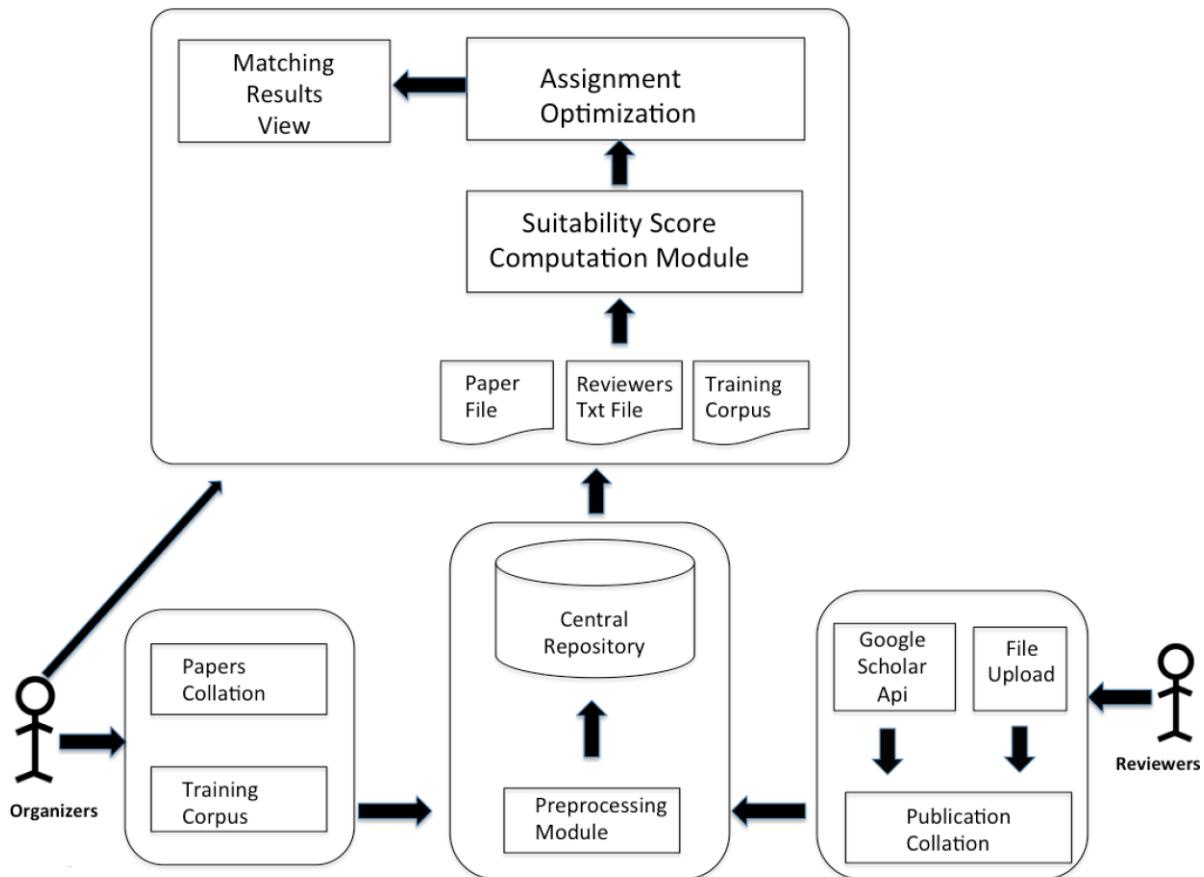
Ogunleye *et al.*

Fig. 2. Overview of the Propose System Architecture

To initially populate the reviewers' repository we plan to obtain the Google Scholar profile ID for each reviewer to download the reviewer's publication papers to achieve the reviewers' corpus. The prepossessing module would use the Glyph & Cog *pdf to text* for extracting the text while natural language toolkit (NLTK 3) would be used to perform natural language processing (NLP) text processing tasks on the extracted data.

The Computing Similarity Scores Module uses Python third-party tools, SKLearn and Gensim, for implementing the methods for computing the similarity scores between reviewers archive and papers submission using TF-IDF, LSI, LDA and Word2Vec models. The Assignment Module uses the optimization integer linear programming model as presented by Taylor (2008) in maximizing the assignments based on the produced suitability scores.

DISCUSSION
The proposed approach has the tendency to provide a fast and more robust platform for paper-reviewer assignment for large conferences. By considering semantic similarity using neural network models such as Word2Vec, assigning a paper to the most suitable reviewer

Ogunleye *et al.*

will be a goal achieved. The Word2Vec model adopted for this work will provides a mechanism for intelligent semantic assignment that will ensure the decision of assigning papers that are relevant to the domain expertise of the reviewer. Also, by integrating an Integer Linear Programing model to the system will ensure that the assignments are maximized so the load balance is achieved across the reviewers. This will ensure the most optimized assignment is delivered to the system. This paper only presents the workflow and the system architecture of the work, which would serve as a guide in implementing the model as discussed in this paper.

CONCLUSION

The automated paper-reviewer assignment system for conference management would be of tremendous benefit to conference managers and chairpersons. It would allow for reviewers management, besides the reviewers would be able to upload their publications to their personalized repository on the platform or use the capability of the system to crawl the Internet to download publication files as provide on their Google Scholar Profile, so as to create the reviewer's representation. The system users would be able to compute suitability scores of the paper-reviewer assignments using pre-trained models using Word2Vec or with other Bag of Words models. The suitability scores would be used to deliver an optimized assignment of papers to reviewers.

REFERENCES

Bengio, Y., Ducharme, R., Vincent, P., & Jauvin, C. (2003). A neural probabilistic language model. Journal of machine learning research, 1137-1155.

Charlin, L., & Zemel, R. S. (2013). The Toronto paper matching system: an automated paper-reviewer assignment system. In International Conference on Machine Learning (ICML) (Vol. 10).

Harispe, S., Ranwez, S., Janaqi, S., & Montmain, J. (2015). Semantic similarity from natural language and ontology analysis. Synthesis Lectures on Human Language Technologies, 8(1), 1-254.

Hettich, S., & Pazzani, M. J. (2006). Mining for proposal reviewers: lessons learned at the national science foundation. In Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 862-871). ACM.

Li, B., & Hou, Y. T. (2016). The new automated IEEE INFOCOM review assignment system. IEEE Network, 30(5), 18-24.

Li, H., & Xu, J. (2014). Semantic matching in search. Foundations and Trends® in Information Retrieval, 7(5), 343-469.

Long, C., Wong, R. C. W., Peng, Y., & Ye, L. (2013). On good and fair paper-reviewer assignment. In Data Mining (ICDM), 2013 IEEE 13th International Conference on (pp. 1145-1150). IEEE.

Manning, C. D., Raghavan, P., & Schütze, H. (2008). Introduction to information retrieval, 1(1), (p. 496). Cambridge: Cambridge university press.

Mikolov, T., Karafiát, M., Burget, L., Cernocký, J., & Khudanpur, S. (2010). Recurrent

Governance and Public Service Delivery in Nigeria: The Role of Information and Communication technologies | **CUCEN2017**

Ogunleye *et al.*

neural network based language model. In Interspeech (Vol. 2, p. 3).

Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. In Advances in neural information processing systems (pp. 3111-3119).

Taylor, C. J. (2008). On the optimal assignment of conference papers to reviewers.

Zablocki, J., & Lee, R. Auto-Assign: An Implementation to Assign Reviewers using Topic Comparison in START.