

# Analysis of Academic Research Networks to Find Collaboration Partners

Kavita Asiwal<sup>(✉)</sup>, Bharath Kumar Suresh, and G. Ram Mohana Reddy

National Institute of Technology Karnataka, Surathkal, India  
kavitaswl@gmail.com, bharath30may@gmail.com,  
profgrmreddy@gmail.com

**Abstract.** Social network analysis has been used for decades to find behavioral patterns and relationships that exist between people in a network. Researchers have been collaborating for centuries with the aim of improving the quality of research, to broaden the scope of problems that they tackle, to speed up the output and to disseminate knowledge across authors. Sometimes it becomes difficult to find the right collaboration partner due to various reasons, the major one being the lack of data about individuals working in their chosen domain in geographically separated locations. In this paper, we explain how social network analysis can be used to help researchers in finding suitable collaboration partners with whom they have not worked in the past but can collaborate in the future. Further, we have considered two different analysis techniques – weighted and non-weighted graph and the results are compared based on the relevance of the outcomes.

**Keywords:** Academic research network · Social network analysis · Collaborative research

## 1 Introduction

It is observed that there has been an increase in the number of research topics that require interdisciplinary treatment, which makes it essential for departments or organizations belonging to different fields of knowledge to collaborate in the problem solving process [1–3]. To some extent, this rise can be attributed to the increasing specialization of individual academics and the broadening in scope of the problems that they tackle. Hence, it is possible to integrate scholarly communities and foster knowledge transfer between related fields through collaboration [4].

An academic research network is a network of researchers who are connected through relations like student-advisor, collaborations and citations. It can thus be used to analyze the collaboration relation among a group of authors. From the researcher's point of view, people who have worked together previously are much more likely to succeed in future collaborations, as they understand each other's areas of interest, approaches and methodologies [5]. In this paper, we propose an automated tool, which can be used by researchers to visualize their collaboration network, and thereby finding the prospective collaboration partners using a graph theory based approach.

As a case study, academic research network of National Institute of Technology Karnataka (NITK) Surathkal, Mangalore, India is considered, and the extracted results have been analyzed using social network analysis techniques. The remainder of this paper is presented as follows: Sect. 2 deals with the related work in the field of academic research network analysis. Section 3 describes the proposed methodology. Section 4 describes the implementation details. Obtained results and discussion are presented in Sect. 5. Finally, the conclusion and future work are given in Sect. 6.

## 2 Related Work

Much of the previous research in this area has focused on co-authorship analysis to analyze the collaboration networks [6, 7]. The co-authorship networks exhibit characteristics similar to the much studied citation networks [8], but co-authorship implies a much stronger link than citation, which can occur without the authors knowing each other. Scientists who have authored a paper together are considered connected and a large number of such connections constitute the research network. Co-authorship analysis has been used to assess the collaboration among academic institutions in a certain geographical environment [1].

Scientific collaboration is accepted as a positive phenomenon and is found to have a significant influence on the performance of individual researchers and institutions, in terms of effectiveness, efficiency and productivity [9, 10]. The analysis of social networks has been widely used to understand the implications of the relationship patterns between researchers in various fields [11, 12]. Further, in [12] bibliometrics information for the period between 1999 and 2005 has been used.

## 3 Proposed Methodology for Development of Automated Tool

Figure 1 shows the proposed framework for creation of the automated tool using graph theory approach. In step 1, the data about researchers of NITK, Surathkal is collected and stored in the database. The data stored includes areas of interest and publications of each researcher. In step 2, the authors are divided into groups based on their research domain. Authors may belong to more than one group depending on how varied their research domain is. The grouping is done using natural language processing (NLP) as different researchers may use different terminology to specify the same area.

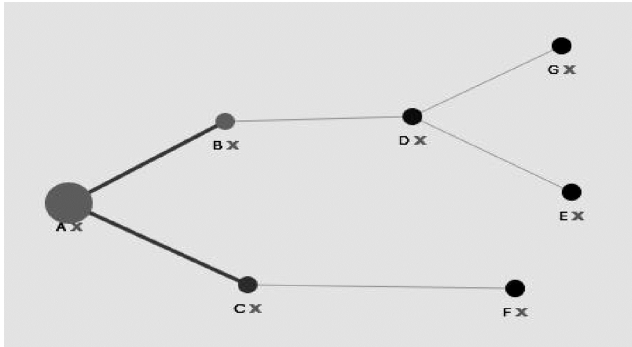


**Fig. 1.** Proposed framework

In step 3, the data stored is used to build the research network. From the data that is collected on papers authored by a researcher, his co-authors are found. The academic research network is represented in the form of an undirected Graph  $G = (V, E)$  where  $V$  denotes a finite set of nodes and  $E$  denotes a finite set of edges. Each node in  $V$  represents a researcher and each edge in  $E$  represents a co-authorship relation between a pair of researchers. In step 4, the network is analyzed using two different approaches, which are explained next.

### 3.1 Non-weighted Graph Approach

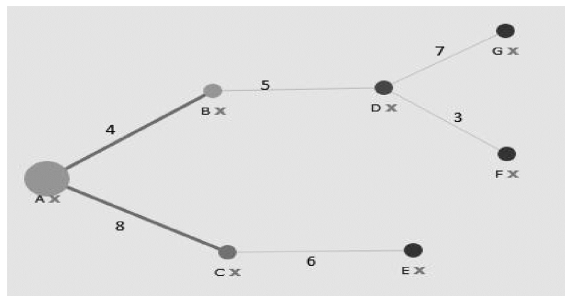
In this approach, the graph used for analysis has non-weighted edges i.e. all edges represent same type of association. A node is created for each researcher and an edge is added between a pair of nodes if the researchers representing those two nodes have co-authored a paper/journal. When researcher  $A$  tries to find his possible collaboration partners, a modification of breadth first search graph algorithm is applied which finds all nodes that can be reached indirectly starting from  $A$  i.e.  $D, E, F$  and  $G$  in Fig. 2. The nodes reached are sorted in ascending order of their distance from  $A$  i.e.  $F$  and  $D$  appear above  $G$  and  $E$ . The details of authors representing these nodes are then displayed to the researcher  $A$  followed by a list of authors from  $A$ 's area of interest group which cannot be reached starting from  $A$ . This is done as there is always a slight probability for researchers with the same interests to collaborate irrespective of whether they have any association or not.



**Fig. 2.** Non-weighted graph

### 3.2 Weighted Graph Approach

In this approach, the graph used for analysis has weighted edges i.e. each edge is different in terms of strength of association. Just like non-weighted approach, here also a node is created for each researcher and an edge is added between a pair of nodes if the researchers representing those two nodes have co-authored a paper/journal. But unlike the previous approach here each edge is assigned a weight which is equal to the number of co-authored research papers/journals. When a researcher A tries to find out his possible collaboration partners, a modified depth first search graph algorithm is applied which finds all nodes that can be reached indirectly starting from A i.e. D, E, F and G in Fig. 3. The nodes reached are sorted in descending order of the weight of paths from A i.e. G, F, E and D in this case. The details of authors representing these nodes are then displayed to the user followed by a list of authors from A's area of interest group which cannot be reached starting from A (reason specified in Non-Weighted graph approach).



**Fig. 3.** Weighted graph

## 4 Implementation

The data on papers published by researchers from NITK during 2007–2012 are considered for co-authorship analysis. The researchers are provided with the option of viewing their collaboration network. The pseudo codes for both the analysis techniques are given next.

### 4.1 Pseudo Code

Input: A Network graph  $G$  and a vertex  $v$  of  $G$  for which collaboration partners are to be suggested.

Output: A list of vertices which are the suggested collaboration partners for  $v$ .

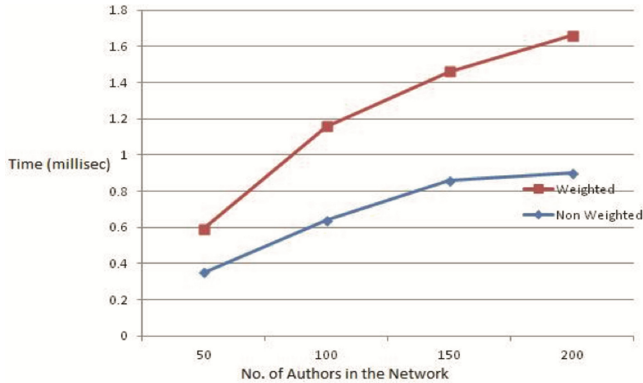
```

procedure suggestCollaborationPartners( $G, v, \text{graphType}$ ):
1:  for all vertices  $v'$  in  $G$  do
2:    if there is an edge from  $v$  to  $v'$  then
3:      for all vertices  $v''$  in  $G$  do
4:        if there is an edge from  $v'$  to  $v''$  and no edge
           from  $v$  to  $v''$  then
5:          if vertex  $v''$  is not same as vertex  $v$  then
6:            add  $v''$  to the result_list
7:            if graphType is Non-Weighted
8:              update length of path for  $v''$  in
                 path_length_list
9:            else if graphType is Weighted
10:             update weight of path for  $v''$  in
                 path_length_list
11:   sort result_list based on path_length_list
12:   return sorted result_list

```

## 5 Results and Discussion

A study is conducted on a group of 200 researchers from NITK. In this study the tool developed is used to find collaboration partners for each of the 200 researchers and the results obtained are analyzed by matching the areas of research to find similarity or some kind of interdependency in the areas. In case of weighted technique, 94 % of the time the results are found to be accurate, while in case of Non-weighted the accuracy dropped down to 83 %. A comparison is performed based on the time taken to analyze the academic research network graph by varying the number of authors in the network. The results obtained are summarized in Fig. 4 and it is observed that the time taken is slightly higher in case of weighted technique.



**Fig. 4.** Processing time versus No. of authors in the Network

## 6 Conclusion and Future Work

The tool developed helps the researchers in finding collaboration partners successfully. The time taken to analyze the network is comparatively less in case of non-weighted graph as the processing time spent for calculating the weights of edges is saved. The results obtained are more accurate in case of weighted approach as more importance is given to the strength of past successful collaborations. The limitation of this approach is that determining the accuracy of results obtained is a subjective process. Further research can be extended in this area by grouping authors based on factors other than area of research i.e. geographical location, work experience or affiliating institute etc.

## References

1. Toral, S.L., et al.: An exploratory social network analysis of academic research networks. In: 2011 Third International Conference on Intelligent Networking and Collaborative Systems (INCoS). IEEE (2011)
2. Bessis, N., Bessis, N.: Grid Technology for Maximizing Collaborative Decision Management and Support: Advancing Effective Virtual Organizations. Information Science Reference (2009)
3. Bordons, M., et al.: Measuring interdisciplinary collaboration within a university: The effects of the multidisciplinary research programme. *Scientometrics* **46**(3), 383–398 (1999)
4. Nikhil, J., Ramage, D., Jurafsky, D.: A study of academic collaboration in computational linguistics with latent mixtures of authors. In: *ACL HLT 2011*, p. 124 (2011)
5. Cummings, J.N., Kiesler, S.: Who collaborates successfully?: prior experience reduces collaboration barriers in distributed interdisciplinary research. In: *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work*. ACM (2008)
6. Rodriguez, M.A., Pepe, A.: On the relationship between the structural and socioacademic communities of a coauthorship network. *J. Informetrics* **2**(3), 195–201 (2008)
7. Farkas, I., et al.: Networks in life: Scaling properties and eigenvalue spectra. *Physica A Stat. Mech. Appl.* **314**(1), 25–34 (2002)

8. Garfield, E.: *Citation Indexing: Its Theory and Application in Science, Technology, and Humanities*, vol. 8. Wiley, New York (1979)
9. Abramo, G., D'Angelo, C.A., Di Costa, F.: Research collaboration and productivity: is there correlation? *High. Educ.* **57**(2), 155–171 (2009)
10. Lee, S., Bozeman, B.: The impact of research collaboration on scientific productivity. *Soc. Stud. Sci.* **35**(5), 673–702 (2005)
11. Monclar, R.S., et al.: Using social networks analysis for collaboration and team formation identification. In: 2011 15th International Conference on Computer Supported Cooperative Work in Design (CSCWD). IEEE (2011)
12. Benckendorff, P.: Exploring the limits of tourism research collaboration: a social network analysis of co-authorship patterns in Australian and New Zealand tourism research. In: 20th Annual Council for Australian University Tourism and Hospitality Education Conference (CAUTHE 2010). University of Tasmania (2010)
13. Sayogo, D.S., et al.: Building the academic community of e-government research on cross-boundary information integration and sharing. In: 2012 45th Hawaii International Conference on System Science (HICSS). IEEE (2012)