Tracking the Evolution of Technology and Science using Patent and Publication Data

Gariman Singh[1], Dr Sarabjot Singh Anand[2], Dr Amit Kulshrestha[1]

[1]Indian Institute of Science Education and Research (IISER), Mohali, India

[2]Sabudh Foundation Mohali, India

Abstract

Today, much effort is being put into tracking the changes in Technology and Science. This is because the results produced are not only informative but also useful, as is in the case with figuring out “Emerging Technologies”. In this work, we approach the problem from the perspective of Graph Theory or Network Science. We design a way to construct keyword networks and identifying groups of communities within them using the techniques of Graph Convolutional Neural Network and Random Walks. Each major community identified represents a hot topic of research in the year of a field for which data has been taken.

Data

Where?

The data for this project has been obtained from Scopus – a bibliographic database of scientific publications.

What?

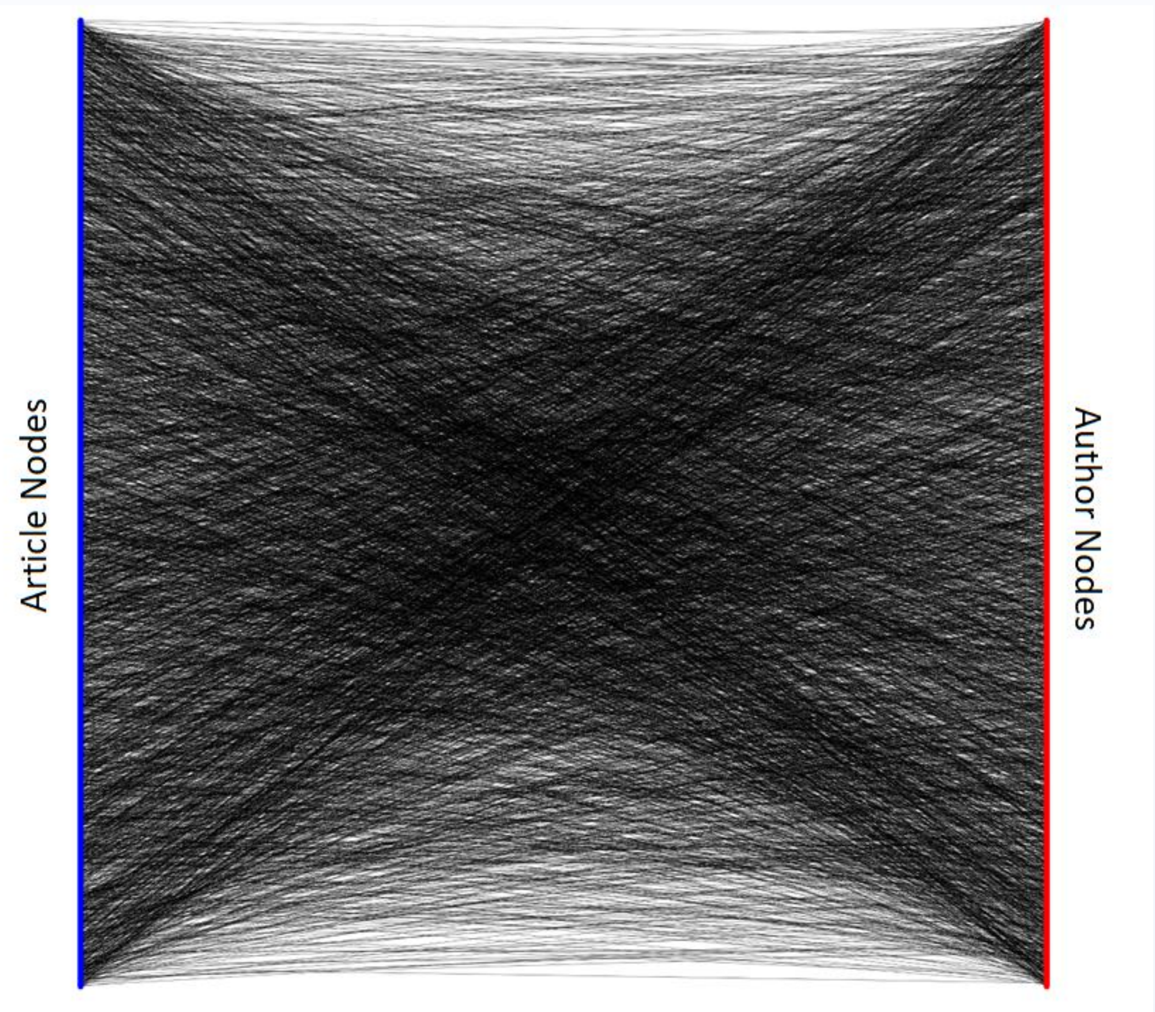
For illustrative purposes, articles from the field of Computer Science published in the year 2015 have been considered.

Author Collaboration Network

To establish an Author Collaboration Network, we constructed a bipartite graph of articles and authors.

Bipartite Graphs

The nodes of such a graph can be divided into two disjoint sets such that there are no edges between any two nodes of the same set.



The bipartite graph was then converted to an author-author unipartite graph by taking the inner product for each pair of authors.

$$\text{Inner Product} = \sum_{k=1}^{n} C_{ik} C_{jk}$$

where C_{ik} is 1 if the i^{th} author is a co-author to the k^{th} article, and 0 otherwise.

Author Collaboration Network

Using the Girvan-Newman Algorithm for community detection, we identified communities of scientific collaborators from the Author Collaboration Network:

```
Wu Z.-G. Shi P. Zhu Y. Zhang L. Su H. Lu R. Zhao Y.-X. Ahn C.K.
Chen M. Alamri A. Xiang Y. Hu L. Alelaiwi A. Zhang Y. Li Y. Hassan M.M.
Mei J. Zhang F. Khan S.U. Li K. Li K.
Rodrigues J.J.P.C. Javaid N. Shu L. Jiang J. Imran M. Khan Z.A. Guizani M. Han G.
Hu X. Wu X. Wang H. Li P. Li L.
Sciuto E.L. Conoci S. Libertino S. Petralia S. Castagna M.E.
Huang S.Y. Chen J.V. Chiu A.-A. Widjaja A.E. Yen D.C.
Romagnoli C. Abolmaesumi P. Fels S. Fenster A. Chang S. Ward A.
Heo S. Kim M. Suk J.-H. Yang Y.S. Kim J. Park K.-T. Oh J.
Yan S. Hong R. Qian X. Wang M. Liu L. Wang S. Tian Q. Huang Q.
Wan J. Tao D. Xu C. Li X. Yu J. Tao D. Gao X.
```

Keyword Network

In a process similar to the one adopted to construct Author Collaboration Network, we constructed a Keyword Network from an Article-Keyword bipartite graph. However, unlike the Author Network, Cosine Similarity was applied to calculate keyword-keyword weights.

$$\text{Cosine Similarity} = \frac{\sum_{k=1}^{n} C_{ik} C_{jk}}{\sqrt{\sum_{k=1}^{n} C_{ik}^2}\sqrt{\sum_{k=1}^{n} C_{jk}^2}}$$

where C_{ik} is 1 if i^{th} keyword belongs to the k^{th} article, and 0 otherwise.

On the Keyword Network obtained, communities of keywords were identified using Graph Convolutional Neural Network and Random Walks.

Graph Convolutional Neural Network and Random Walks

Graph Convolutional Neural Networks extend the idea of Convolutional Neural Networks to graphs. Since graphs are an example of non-Euclidean data, convolution is performed on them in the Fourier basis.

The Fourier Transform of graph is taken by projecting it on to the orthonormal basis of eigenfunctions of the Laplacian L.

$$L = D - A$$

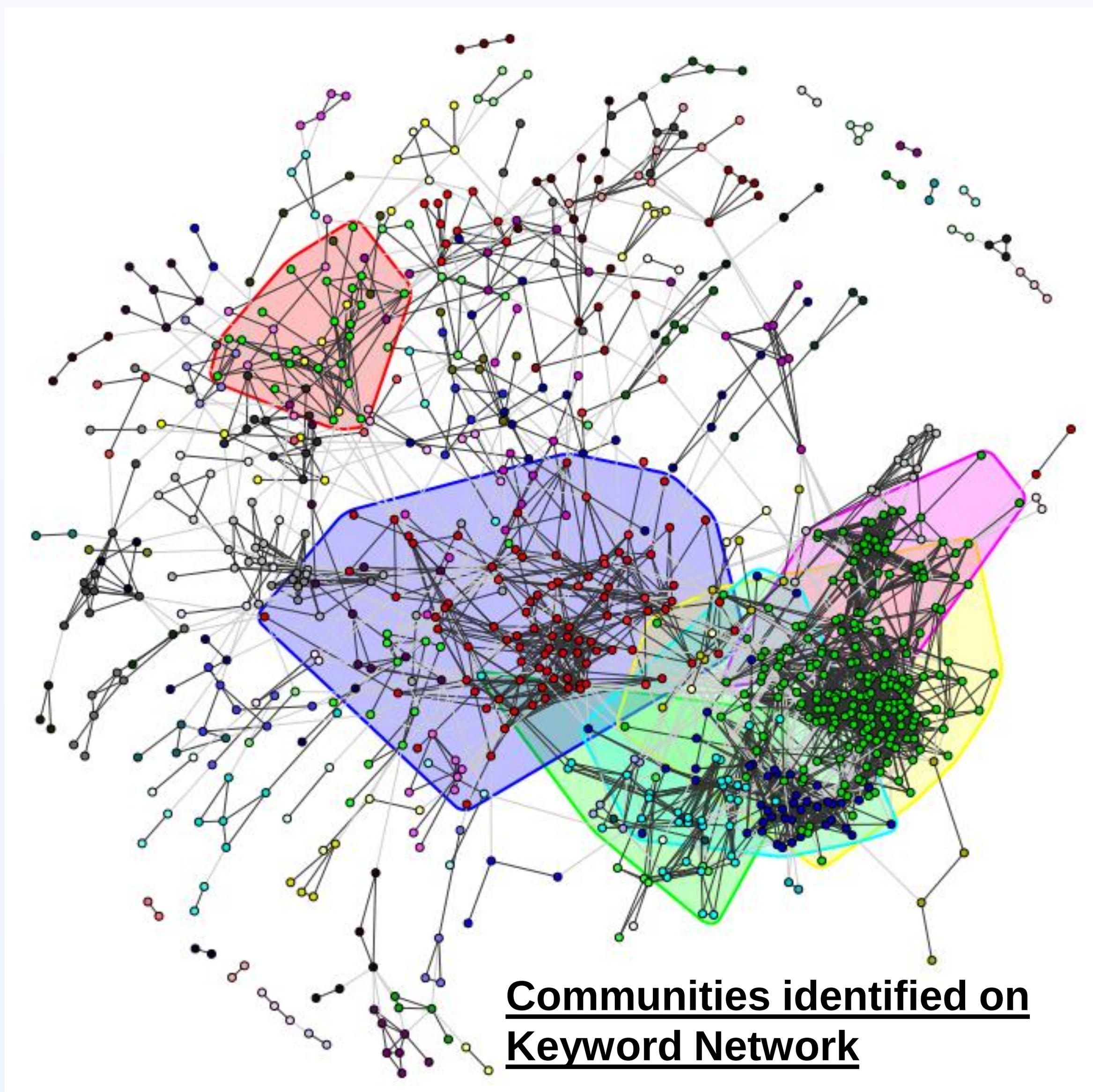
where D is the diagonal matrix and A is the adjacency matrix of the graph.

After making some approximations, the convolved signal is obtained as:

$$\boxed{x * g = D^{-\frac{1}{2}}(A + I)D^{-\frac{1}{2}}XW}$$

Communities of keywords were identified using Walktrap – an approach based on Random Walks.

Results



Communities identified on Keyword Network

Target tracking
Vehicle Routing Problems
Optimization
Conditional random field
Clustering algorithms
Financial markets
Backpropagation
Reliability
Linear transformations
Graph algorithms
Air navigation
Data packet dropout
Convergence
Control engineering
Job shop scheduling

Keywords with highest Degree Centrality in their respective communities

References

[1] K. Boyack, K. Börner, and R. Klavans, “Mapping the structure and evolution of chemistry research,” Scientometrics, vol. 79, no. 1, pp. 45–60, 2009.
[2] J. Bruna, W. Zaremba, A. Szlam, and Y. LeCun, “Spectral networks and locally connected networks on graphs,” arXiv preprint arXiv:1312.6203, 2013.
[3] M. Defferrard, X. Bresson, and P. Vandergheynst, “Convolutional neural networks on graphs with fast localized spectral filtering,” Advances in neural information processing systems, vol. 29, 2016.
[4] T. N. Kipf and M. Welling, “Semi-supervised classification with graph convolutional networks,” arXiv preprint arXiv:1609.02907, 2016.
[5] A.-L. Barabási, “Network science,” Network Science, 2016.

Acknowledgement

We would like to express our sincere gratitude to the institute library of IISER Mohali for providing access to the Scopus database for our publication data requirements.