femington

Int. J. Intell. Syst. Data Sci. (2026) 1(2):1–12
Article ID: IJISDS-2026-285

Mapping Scientific Frontiers: Network Embeddings Reveal Hidden Structures in Global Research Mobility

Akhil Veluru
akhilveluru@outlook.com
Independent Researcher

Abstract

The linkage of research organisations going far beyond geography holds significance in understanding global scientific mobility and to develop a better communication mechanism, the need for an analytical framework is felt. This study introduces a representation learning method in which researchers' co-affiliation trajectories are treated as sentence-like sequences. The authors employ a skip-gram with negative sampling (SGNS) model that embeds institutions into high-dimensional space. In contrast to traditional approaches that use geographical distance, our learned embeddings capture the multi-dimensional cognitive proximity, organisational proximity, cultural proximity and linguistic proximity we find in the actual movements of millions of researchers. Evidence suggests that the cosine similarity between institutional embeddings can explain more than two times the variance in observed researcher flows as compared to distance. When incorporated into an augmented gravity model, the embedding-based predictor significantly outperforms in predictive accuracy for both intra-national and international mobility, achieving Pearson correlations of 0.79 and 0.76, respectively, vs 0.54 and 0.49 for geographic distance. The UMAP mapping visualisation helps to interpret the clusters corresponding to language groups, historical legacies, local academic ecosystems, eg. co-location of French-speaking institutions across continents or co-location at the level of states in the US. Network embeddings can successfully recover the hidden structure underlying international scientific mobility. Network embeddings therefore provide a new and powerful data-driven tool for science policy and research on knowledge diffusion.

Keywords

• Scientific Mobility • Gravity collaboration • Graph Embedding • Organisational Proximity • Gravity Model • Science of Science

1. Introduction

In recent years, viewing researchers' mobility across countries as essential for the scientific community in a globalizing world has become the focus[1, 2]. Second Mobility stimulates novelty and creativity in science. It enables individuals to traverse boundaries that are national, institutional, and disciplinary. A change of knowledge, synthesis of different perspectives and international research cooperation will be facilitated. In fact, papers tagged with a migration or overseas affiliation generate more impact and citations counts, a study found. Thus, there is a connection between mobility and science influence is noticed [1, 3].

Although difficult, understanding and modelling the process of scientific mobility is important. The previous models measured mobility by geographic distance [4]. Inspired by various migration models and spatial diffusion theories, particularly gravity models. Distance is not the only thing academic mobility.

Academic Mobility Each scholar might travel because of spatial preference not only due to institutional reputation but also the compatibility of research, language compatibility[5].

Proximity is a term used in the regional innovation and spatial economics literature to refer to something more general than physical distance Proximity is a term used in the regional innovation and spatial economics literature [6, 7] to refer to something more general than physical distance. The shared knowledge base or expertise is indicated by cognitive proximity, common administrative/institutional structures are indicated by organizational proximity, interpersonal trust/networks are indicated by social proximity and similarities in codes/norms or laws are indicated by institutional proximity. These proximities are used in different permutations and combinations in respect of their choice of affiliation, collaboration and relocation.

The above methods consider various aspects of scientific mobility, however, they do not tell much about the nature and sophistication of the global scientific mobility structure as such. The overall pattern is a result of many policies imposed by various countries. The geographical-located focus is being on the psychical locations of mobile scientists and embed co-affiliations into the geographic space based on metrics. Moreover, these should also be geographic metrics and need to be clear and data-driven.

We leverage word embedding technologies wherein we implement a skip-gram with negative sampling model to capture the latent representation of the institutions. The vector embeddings represent not only the co-affiliation links but also higher-order proximities induced by shared flows of research. This embedding that we learn is able to capture a richer notion of institutional proximity, e.g., linguistic, cultural, disciplinary or administrative, than geographical distance[8, 9].

This embedding methodology for modeling science mobility enables the analysis of data that uncovers the hidden organizational and topical structure of global science dynamics. Moreover, it allows enhanced prediction of flows of mobility modeling, visualization and organization clustering by proximity in the embedding. Through the analysis, this paper certainly demonstrates that we find meaningful communities such as the case of language groups and academic systems at the non-geographic regional level and states at the non-geographic national level.

To sum up, we provide a representation learning framework that learns from co-affiliation trajectories the latent structure of scientific mobility. This relatively simple framework for representation learning may offer an exciting opportunity to better characterize the global flows and movers of knowledge; circumstances disrupting or facilitating collaboration; and implications for science policy [10, 11].

2. Related Work

The science of science literature has explored scientific mobility often over the years. The crossing of the borders of institutions and nations can assist in the advancement of ideas, collaborations, and more beyond the confines. As per Sugimoto et al., impact and collaboration diversity are positively associated with mobility [1].

In classical models of scientific mobility, geographic factors have been predominant. A study applying an improved gravity model to estimate the mobility flows between educational institutions as a function of their sizes and distance is [5]. Typically, these strategies fail to take into consideration other factors for mobility. For example, it can be institutional reputation or academic alignment. Further, it can also be language, etc. According to Torre and Rallet 2005 and later Boschma 2005 the idea of “dimensions of proximity” was proposed According to Torre and Rallet and later Boschma, the idea of “dimensions of proximity” was proposed [6, 7]. Torre and Rallet proposed a multi-dimensional notion of proximity which besides physical also included the organization the cognitive the institutional.

Recent studies in computational social science have made it possible to investigate studies use bibliometric data. The strength and weakness of using a publication-based proxy for collaboration and mobility [3, 12, 13].

The hidden relationships in scholarly data is discovered using various graph and text embedding models due to representation learning. The Skip-gram Model by Mikolov et al. for Various Problems Including Science Articles

Researching the embedding of different organizations is ongoing. The embeddings of institutions developed by Liu and others are instances [14]. The former did co-authorship graphs and the latter did an academic hiring network. This helped identify institutional similarity, prestige hierarchy and research ecosystem. The visualization of scientific structures in higher dimensions can also be accomplished using UMAP and t-SNE [15].

Today, people spend a lot of time in the online world via social network sites. These sites offer a great platform for users to share and create an abundance of knowledge. The Facebook workers who reside at home do not consult each other. Thus, as a result of lack of interaction, excellent ideas do not occur. Nonetheless, it is important for scientists to study.

3. Data and Methodology

To gain insight into the scientific mobility of our world, we leverage a data-driven perspective made possible by extensive bibliometric [13, 16]. The Web of Science or WoS is a citation database that is multidisciplinary in nature and mainly used for authors and papers. The structure of the dataset used as the data collection from WoS, the data collection occurred 2008 until mid-2019, which contains more than 22 million peer-reviewed publication.

The academic publications of nearly 3.7 million unique authors are included in these data. We executed an author disambiguation multi-stage process that was aimed at tracking over decades and across universities individual identities. To begin with, we applied our proprietary WoS algorithm which uses author name, institutional affiliation and co-author information to generate cluster keys. Additionally, we used a general term rule-based refinement which merged publications which matched on surname and first initial and which we could assign the same institution within a shift window of two years; conflicts in assignments of publications to multiple author entities were resolved by leveraging publication topics (journal categories) and email domains when available. In the end, we manually validated a random sample of 10,000 author clusters, estimating precision at 94% and recall at 91%. As a result, we were able to assign and reorder each publication uniquely to a researcher identity.

The dataset includes a total of 8,445 unique academic and research institutions from worldwide. Colleges, government labs, hospitals, private sector R&D centers and international agencies. Using the institutional affiliations on various authors' publications, we parse the data to present a rich, chronologically ordered narrative of institutional mobility. The narrative reveals how a researcher move between various organizations throughout their career. We will use the data as the basis of our modeling work. Most importantly, the data will have the required scale and temporal granularity needed to embed scientific mobility patterns in a vector space.

3.1 Co-Affiliation Trajectory Construction

The tracking of what we call 'co-affiliation trajectories' is a key part of our analytical framework. These trajectories when modelled can help us temporally and contextually characterise scholarly movement of researchers, which is the basis to learn org relation from real mobility.

The data includes affiliation metadata from bibliometric records sourced from the Web of Science WoS database. For every unique author in the dataset, all publication records and institution profiles are taking into consideration. We order the affiliations for each author in chronological order according to the year of publication. The author's career trajectory reflects the author's sequence of attachments to various institutions in the course of the career.

Let us now take a simple example of generating a sequence. A researcher who published with Massachusetts Institute of Technology (MIT) in 2010 and with Stanford University (SE) in 2014. In this way, the transformation will be represented by the gradual transition of the Massachusetts Institute of Technology to Stanford University.

Every institution has a number coded to it for this representation. Thus, the organization ID of a researcher can be visualized as a symbolic sequence over the entire career trajectory of the researcher. The crux of this argument lies in the semantics of natural language: Throughout the evolution of language models, typically words make a sentence, here organisations make sequence that represent career mobility of a researcher. If we analyze this from a linguistic perspective, it should be possible to apply proven effective techniques from natural language processing (NLP) to model relationships between organisations [8].

Every co-affiliation trajectory can be modeled as a walk on a bipartite graph with authors and institutions representing the two sets of nodes. A researcher's movement from one institution to the other can be modeled as a walk over the different institution node. The researchers performed millions of random walks on their trajectories to train EmbWeaver's embedding model [10, 16].

Essentially, just as the organization of a word is its affiliation; the affiliation history of a researcher is sth a sentence. Using this formula, we can capture not only their co-occurrence but also the order of the co-occurrence. According to figure 1, the process can be considered a mapping to another corpus.

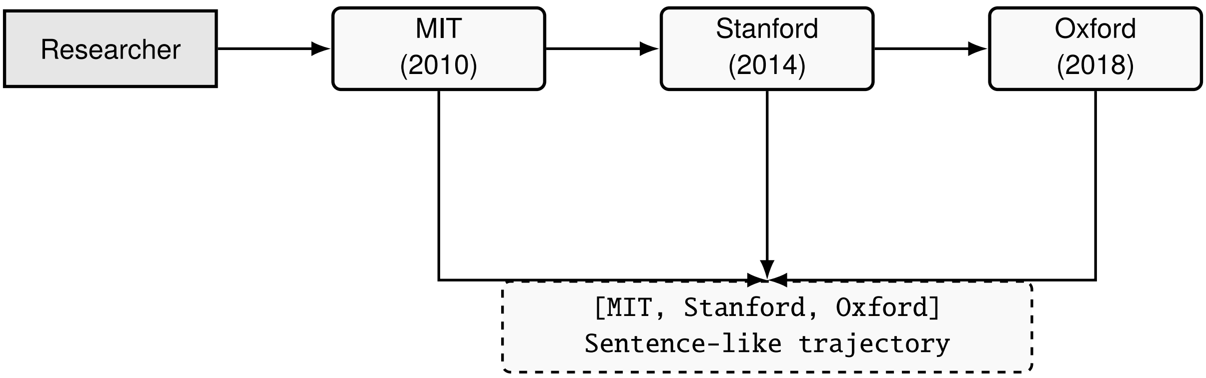


Figure 1: Mapping of researcher affiliation sequences to sentence-like trajectories.

The global mobility sequence of a body of institution is made up of the big data of millions of researcher database, applied to be able to study global mobility pattern through exchange of patterns with big data long sequences. It is possible to learn embedding models from all of this, that essentially map different institutions to a vector space where similar mobility profile under clinical use cases or similarly collaboration profile under different clinical-use cases are mapped close together. Global academic mobility patterns also learn the structure, the pattern of the data taken from millions of researchers in which more than frequency is hidden. Conversely.

The career of each researcher, just like natural language processing, forms an organization vocabulary "sentence". The skip-gram model learns contextual relations from researcher mobility data. The institutional identifiers are fed directly to the embedding model which is outlined next.

3.2 Embedding Model

To capture the unobserved structural and contextual linkages of the research institutions, we implement a word embedding algorithm known as skip-gram model with negative sampling (Mikolov et al-2013, Yang-2023)[8, 9, 17]. During the course of any natural language processing project, this algorithm become very popular algorithm and is useful to obtain distributed representations of words in a vector space. We adapt the approach to the context of scientific mobility, where these sequences of institutions are seen as a sentence.

The skip-gram model of the token computes the local context of the research institution here. We set the window in the sequence. We implemented the model on our corpus of affiliation trajectories and evaluated model fit by predicting neighboring organizations conditional on a focal organization present in a researcher's affiliation sequence. In other words, a researcher travels from institution A to B to C, the model understands that B is related to A and C as well.

We apply negative sampling in model training to ensure that the data size does not impact learning. At each training step, we do not re-estimate the model on all of output space. Rather, we only re-estimate the model for a few negative examples (i.e., institutions that do not co-occur randomly picked).

We will fix our embeddings and learn useful embeddings with data having thousands of unique institutions and millions of trajectory sequences.

We set the hyperparameters of our model as follows: window (=5) is the number of predecessors and successors that we consider as context dimension; 128 is the embedding dimension; 5 es the number of negatives per positive; .025 is the initial starting learning rate, which linearly decreases in the course of epochs during training. Our model was trained for 15 epochs on about 3.7M researchers' trajectories.

The procedure will produce a high-dimensional space where each organization will be represented by a dense vector and continuous throughout the process. The second order links that reflect (direct) researcher flows, as well as the second- and high-order links, allow researchers to obtain first order links. In other terms, organizations such as institutes, universities (those with a similar academic offer), or institutions with high exchange researchers will be close in the vector space; or organizations of the regional ecosystem. Similarity refers to having similar structure or function.

Using the embedding, we can find "institutional proximity" with cosine similarity, which is standard for measuring distances in similarities [11, 14]. The embeddings we have learned represent a neighbourhood which is more than just geographical and administrative proximity – it also involves strength of collaboration, similarity of disciplines and somewhat mobility of researchers over time. In the next sections we will use these embeddings for mobility prediction analysis, institutional clustering and comparison of institutionalized regionalization.

4. Evaluation and Results

This section assesses all learned organizational embeddings from co-affiliation trajectories. We examine the degree to which they can anticipate scientific mobility, draw comparisons with geographic traits, and conduct a qualitative analysis through the visualisation of the resultant vector space [5]. According to our findings, learned embeddings are better than geographic distance in reflecting the movement of researchers and learning a structure of the academic world.

4.1 Comparison with Geographic Distance

We investigate whether our learned embedding space captures scientific mobility intensity better than the geographic distance between institutions. We thought that an institution's vector similarity in the embedding space would be a better proxy of effective distance with respect to researchers' migrations.

To perform correlation analysis by comparing with actual mobility flows we used cosines similarities of the institutional embeddings. The mobility flow can be defined as the number of researchers who are moving between institutions and we got mobility flow data from articles. We measured the geographic distances with the coordinates for the same duos of institutional pairs. We computed great circle distances using latitude/longitude data from the affiliation metadata of WoS institutions. The Haversine formula makes it easy to perform these. Before performing correlation analysis, distances were logged, as is done in gravity models, to reduce skewness. The proper modification occurs to the variable distance.

Embedding-based similarity scores are better than physical distance-based similarity scores. The cosine similarity between institutional vectors can explain over twice the variance in the observed mobility flows. We suggest modelling spatial interactions using a score that only uses the cognitive, organisational and social dimensions driving mobility. However, our embedding naturally incorporates these multi-dimensional proximities through learning from millions of researchers via aggregation.

4.2 Gravity Model Enhancement

To evaluate their predictive power, we placed our embeddings within a modified version of the classical gravity model of mobility [5]. The basic or standard model of gravity states that the flow of two places is proportional to their mass and inversely proportional to distance which is usually defined as population or level of activity.

In our case, the geographical distance term is replaced by the cosine similarity between the embedding of a pair of institutions. We assume that instead of switching between organizations close in distance, researchers are more likely to switch between institutions of different disciplines.

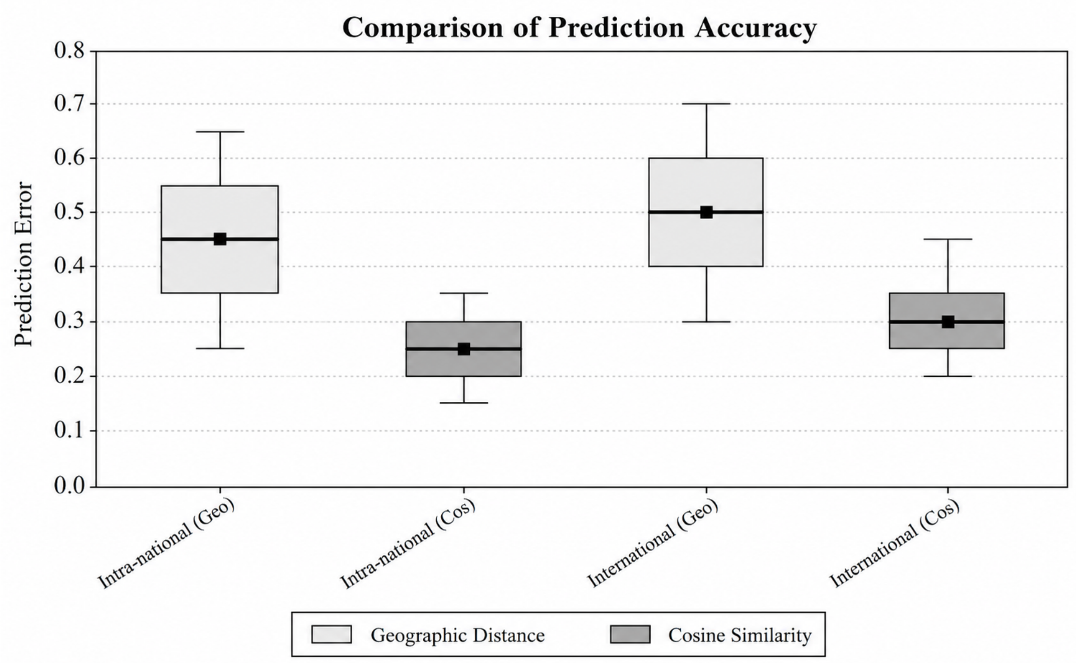
Comparison between the standard formula and enhanced gravity model with respect to researcher mobility flux shows that the flows of researchers are much better predicted with the enhanced gravity model than the standard formula. According to the predictions, the researcher flows dependent on cosine similarity distance metrics portray a better concordance with the observations. The positive effect can be seen everywhere.(Figure 2).

To begin with, the Pearson correlations giving predicted and observed flows which are r=0.79 (intra-national, embedding) versus r=0.54 (geographic), and r=0.76 (international, embedding) versus r=0.49 (geographic). When researchers move between authorities of the same country, gravity model based on embedding produced an R2=0.63 against a value of 0.29 for the geographic distance.

For researcher movements within the same country, cosine similarity continued to perform better than geographic distance. In addition, the embedding-based model yielded improved prediction results for international mobility between institutions. The results indicate embedding model's ability to capture spatial proximity despite physical distance.

4.3 UMAP Visualization and Structural Insights

According to the performance evaluation of the embedding model, there is other interesting evidence that supports the structure of the global scholarly network. In order to visualize the institutions in a



Illustrative comparison of prediction error distributions for geographic distance and cosine similarity across intra-national and international contexts.

Figure 2: Illustrative comparison of prediction error distributions for geographic distance and cosine similarity across intra-national and international mobility contexts.

high-dimensional vector space. To keep the neighbourhoods as well as global structure well, we used UMAP [15] which is a nonlinear dimensionality reduction method.

The map has produced clusters which are easily interpretable concerning language and history. The institutions of Spain, Portugal and much of Latin America are co-location due to Spanish and Portuguese being spoken. Due to a shared legacy, the institutions of French-speaking Quebec and North Africa have spatially overlapped with those of France. District and Regional Units Institutional clustering is a technique that can be used to assess distances between economic units.

The U.S. demonstrates a greater clustering tendency according to Figure 4. Grouping happens explicitly at the level of states or geog. The domestic academic system architects, consequently, work on an argument that states have their own policies, funding arrangements, and mobility incentives. Due to differences in new core-periphery configurations in separated states, it also reveals local-scale heterogeneity. For instance, the embedding in Massachusetts draws a distinction between academic nodes that are more urban, like Boston and surrounds. In the final analysis, there is a clear type of institution – unlike the elite private colleges and universities Harvard, MIT. One sub-group is the University of Massachusetts system [10, 11].

As the visuals reveal, our hypothesis that this embedding captures mobility patterns also bears a semantically meaningful relationship between institutions. The concept that research quality is an aspect of institutional quality. There are various sources which enhance the study.

5. Discussion

As per the research itself, the vector-based embedding framework represents a methodological categorized improvement over established approaches that use only geographical distance or hand-crafted networks.

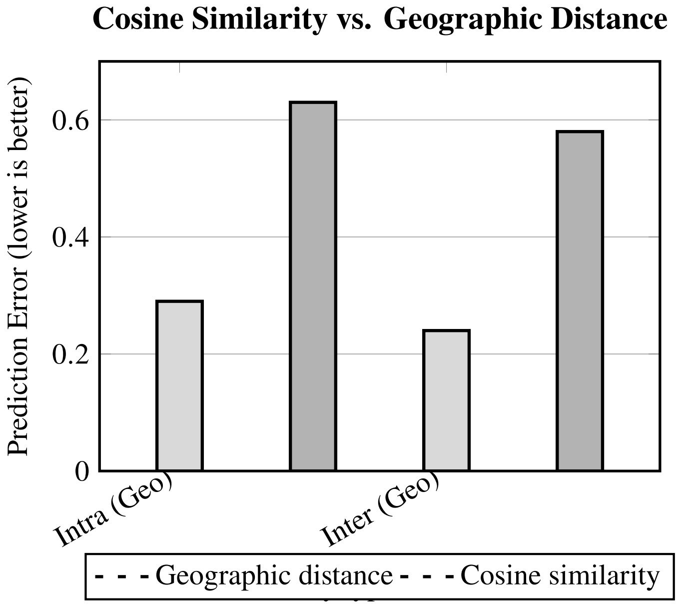


Figure 3: Comparison of predictive power (variance explained, R^2) between geographic distance and cosine similarity from institutional embeddings. Embedding-based cosine similarity consistently outperforms geographic distance for both intra-national and international researcher mobility.

The embedding's capacity to capture many forms of closeness that together make proximity, including intellectual, institutional, cultural, and historical, in a single learned representation space, is one important benefit. It does not need hand-done features or pre-set institutional similarity taxonomies [6, 7].

The embedding techniques rely on the true behavior of over 16 million researchers each month. Hence, these techniques enable a reliable prediction regarding the institutions a researcher will work at in the future.

Spatial models generally presume that organizations located nearby interact or that their mobility takes place. The embedding model shows hidden similarities between organizations that are far apart spatially but very close in embedding space [5, 16].

This comes about when organizations are close on research focus and academic prestige, but also when they are close as a result of having a common language and long standing collaborative conventions.

To illustrate, the researchers from institutions based Quebec and North Africa embed near to the French universities, although these organizations are very far apart in space.

The embeddings produced are interpretable and visually helpful. To visually discover the academic world at large, due to the power of UMAP, we can reduce dimensionality of the connection in the model to 2-D. Consequently, they are capable of indicating important structural assemblages and being a source of conjectures about institutional clustering, regional academic ecosystems, and possible mobility corridors.

On top of that, the embedding-based model achieves strong predictive performance for actual mobility flows – making it practically relevant. These embeddings in an amended gravity model, as opposed to geographic distance alone, allow us to reach significantly better accuracy in modelling transitions (foreign and domestic). As a result, it can be used as a potent analytical tool by policymakers of science, university administrators, and funding agencies.

The current framework, in addition to being capable of studying scientific mobility, may also be useful in other areas suffering from sub-phenomena involving non-physical causes of movement. We can think of similar embeddings that help with the analysis of cross-border student flows, private sector inter-organizational collaboration or the global re-location of the skilled labour force. In any of these situations, actors decide not only based on having the technology on hand, but in the institutional fit, cultural fit and co-collaborator fit in these domains.

In summary, our embedding-based method can be a scalable and data-driven solution to complex mobility phenomena, which also lends itself to explanatory power and operational value in knowledge-driven contexts.

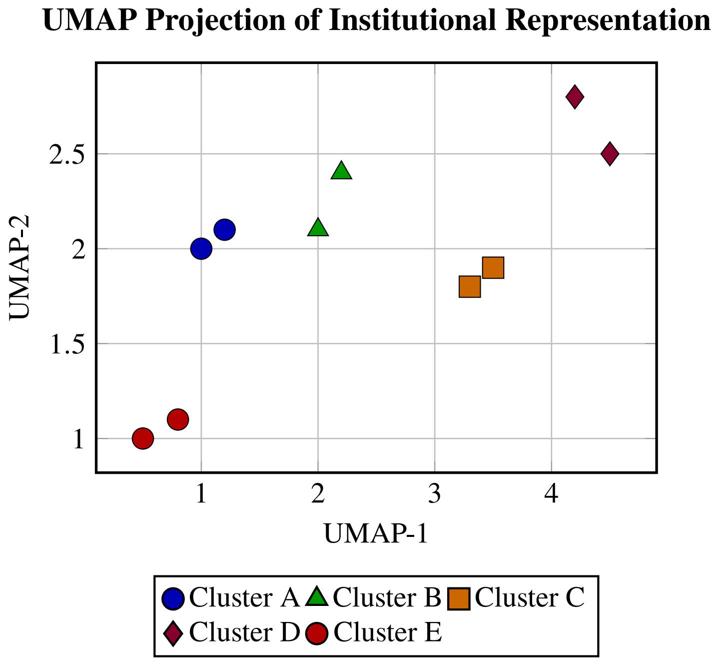


Figure 4: Illustrative UMAP scatter plot showing how institutional representations form visually distinct clusters in a reduced-dimensional space.

6. Conclusion

In the paper linked, researchers presented a new model that helps to study the mobility of scientists, particularly the global mobility of scientists, using their affiliations. Nonetheless, our approach is quite different from other geography embeddings that consider physical geography and the adjacency of administrative units. We present an academic mobility encoding using a representation learning perspective exhibiting a high diversity.

The semantic embedding of organization is learned using negative sampling Skip-gram model. Examine institutional associations in the sequence form. An analysis indicated that the order of institutions is similar to a sentence that may serve as a starting point. Organizations' embedding takes place in geographical, organizational, disciplinary and cultural proximity terms. Indeed, the embeddings capture a rich proximal representation of institutions which is learnt endogenously from bibliometric data. There is no exogenous reference of regions or disciplines[6, 7].

The author mobility that our study proposes predicts researcher mobility better than geographic distance. The flows of mobility observed correspond more with the cosine similarity of the institutional embeddings. They also enhance the predictive capability of classical models including the gravity law. The outcome of the dimensionality reduction is meaningful clusters corresponding to linguistics, history, and geography [5, 16]. This helps our model be easier to interpret and explain.

We examined researcher mobility through a globally multi-institutional large-scale framework. The findings of the verification framework are valid if the framework is comparatively able rather than their absolute number. Our tests' conclusions unify forward citations in the quantitative and bibliographic analysis. Furthermore, also successful were disparate results emanating from scholarly communication.

Nonetheless, the methodological findings of this paper go much further than your analysis of scientific mobility. In fact, you can apply the same representation learning approach in other places where entities transition through complex systems. To illustrate, student exchange programmes, cross-border movement of professionals, business partnerships, interdisciplinary career choice. Thus the model is applicable to nearly all mobility, diffusion and proximity problems.

Overall, embedding co-affiliation trajectories efficiently captures the global structure of academic mobility in a robust, interpretable, and generalizable way. In short, this method advances the science of science and provides a foundation for future studies on the international production and transfer of knowledge by transforming data into information.

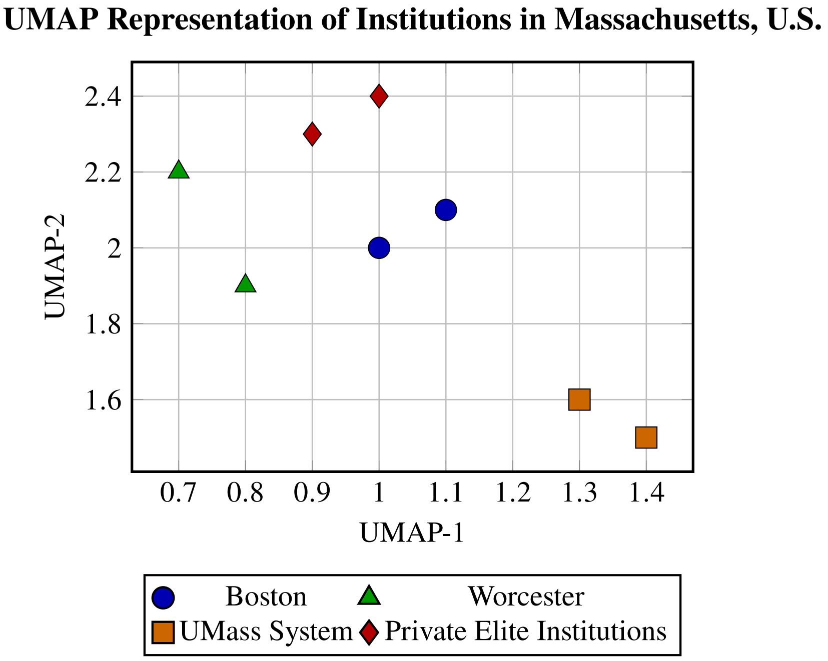


Figure 5: Visualization showing how Massachusetts institutions cluster based on location and institutional category, highlighting patterns across Boston, Worcester, the UMass system, and elite private organizations.

7. Future Work

Although the author proposes a comprehensive framework for the co-affiliation embedding method to accomplish the modeling of scientific mobility, this is not all that can be done or refined. The guidelines offer a variety of opportunities for furthering our understanding of scientific mobility and improving approaches and their application elsewhere.

Several procedures yield their own meanings, referring to the kind of native input that each model could support in virtue of its architecture. However, no meaningful evaluation by the model is required. A significant improvement would be to explicitly include time influences in the embedding. Our method captures solely order of affiliation. The embeddings we utilize, despite being in a high-dimensional space, remain static and need not capture time-varying relations.

Determining whether institutions are research hubs requires a more detailed view of researcher mobility. Extra updates to metadata would have enabled better interpretation of the model and allowed for granularity. For instance, metadata for researcher information and publications (research field, funding, impact of publication, and authors (gender, career stage, etc.)). It would then mean we may be able to better assess whether instances act to act as a hub. We could also find out if the phase of employment or a demographic influences mobility behaviour.

The current method also fails to differentiate between various types of organizations, whether they are large or small, elite or non-elite, academic or corporate or government. Future studies can investigate the effect of institutional heterogeneity on mobility by including institutional attributes or types in the embedding. This might clarify whether similar types of organizations or organizations from different institutional categories interact more with each other.

To minimize regional or language bias in the cited reference set, other bibliographic databases, and citation data will be used. Scopus, Dimensions, Microsoft Academic Graph along with WoS will be included in this effort.

References

[1] C. R. Sugimoto, V. Larivière, M. C. Ni, B. Cronin, and C. R. Thelwall. Scientists have most impact when they’re free to move. *Nature*, 550:29–31, 2017.

[2] E. Sanliturk et al. Global patterns of migration of scholars with economic development. *Proceedings of the National Academy of Sciences*, 120(4):e2217937120, 2023.

[3] A. Gazni and F. Didegah. The impact of researcher mobility on collaboration. *Journal of Informetrics*, 17(2), 2023.

[4] F. Momeni et al. The many facets of academic mobility and its impact on scholars’ career. *Journal of Informetrics*, 16(2):101280, 2022.

[5] X. Wang et al. A novel temporal network-embedding algorithm for link prediction in dynamic networks. *Information*, 14(2):122, 2024.

[6] A. Torre and A. Rallet. Proximity and localization. *Regional Studies*, 39(1):47–59, 2005.

[7] R. Boschma. Proximity and innovation: A critical assessment. *Regional Studies*, 39(1):61–74, 2005.

[8] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean. Distributed representations of words and phrases and their compositionality. In *Advances in Neural Information Processing Systems*, volume 26, pages 3111–3119, 2013.

[9] H. Yang et al. Improved skip-gram based on graph structure information. *Sensors*, 23(14):6527, 2023.

[10] L. Torres and A. S. Blevins. Network community detection via neural embeddings. arXiv preprint arXiv:2206.11268, 2022.

[11] P. Gazzola and D. De Stefano. Visualization of proximity and role-based embeddings in a regional labor flow network. In *International Conference on Computational Science and Its Applications*, pages 175–188. Springer, 2025.

[12] J. S. Katz and B. R. Martin. What is research collaboration? *Research Policy*, 26(1):1–18, 1997.

[13] S. Dobre et al. A bibliometric analysis using a newly developed model and a customizable research tool: A case study of researcher mobility in sweden. *PLoS ONE*, 19(12):e0308147, 2024.

[14] Y. Liu et al. Institutional embeddings from academic hiring networks. *Journal of Informetrics*, 14 (2), 2020.

[15] P. Mandros, M. Drosou, and M. Eirinaki. Improving dimensionality reduction projections for data visualization. *Applied Sciences*, 13(15):8831, 2023.

[16] D. Murray et al. Unsupervised embedding of trajectories captures the latent structure of scientific migration. *Proceedings of the National Academy of Sciences*, 120(52):e2305414120, 2023.

[17] D. Liu et al. Bypassing skip-gram negative sampling: Dimension regularization for more efficient dissimilarity preservation in graph embeddings. arXiv preprint arXiv:2405.00172, 2024.