Poster

STI 2022 Conference Proceedings

Proceedings of the 26th International Conference on Science and Technology Indicators

All papers published in this conference proceedings have been peer reviewed through a peer review process administered by the proceedings Editors. Reviews were conducted by expert referees to the professional and scientific standards expected of a conference proceedings.

Proceeding Editors

Nicolas Robinson-Garcia
Daniel Torres-Salinas
Wenceslao Arroyo-Machado

Citation: Dudek, J., Van Hoed, M., Núñez, López, L., & Costas, R.(2022). Mobility footprints: A notation method for studying researcher moves. In N. Robinson-Garcia, D. Torres-Salinas, & W. Arroyo-Machado (Eds.), *26th International Conference on Science and Technology Indicators*, STI 2022 (sti22229). https://doi.org/10.5281/zenodo.6957327

Collection: https://zenodo.org/communities/sti2022grx/

Mobility footprints: A notation method for studying researcher moves

Jonathan Dudek[*], Miriam Van Hoed[**], Lidia Núñez López[**] and Rodrigo Costas[*,***]

[*] *j.dudek@cwts.leidenuniv.nl; rcostas@cwts.leidenuniv.nl*
Centre for Science and Technology Studies (CWTS), Leiden University (Netherlands)

[**] *miriam.vanhoed@ideaconsult.be; lidia.nunez@ideaconsult.be*
IDEA Consult (Belgium)

[***]
Centre for Research on Evaluation, Science and Technology (CREST), Stellenbosch University (South Africa)

Introduction

Scientists move across institutions and countries. This has implications for the researchers, and also for the institutions or countries they move to or from. For example, some countries may experience increases or decreases in their workforce. One way of capturing researcher mobility is by using publication metadata,[1] which allows to track changes in the affiliations of researchers and their successive career steps over their entire publication activity.

It is only relatively recently that bibliometric approaches have begun to leverage the potential of affiliation data for mobility studies (Moed, Aisati, & Plume, 2013; Robinson-García et al., 2019). However, there are still important nuances to mobility (e.g., the return of researchers, sporadic relationships between a researcher and affiliations) that await incorporation (El-Ouahi, Robinson-García, & Costas, 2021). This necessitates more fine-grained and flexible frameworks able to capture the diverse types of academic mobility.

Sociological studies have modelled careers as sequences of abstract events (Abbott & Tsay, 2000) – e.g., the careers of musicians (Abbott, 1990). Inspired by that, we propose an approach for academic mobility in which affiliation changes identified in researchers' publications will be framed as mobility sequences (or 'mobility footprints'). Accordingly, this paper has two main objectives: 1) to propose a terminology that covers a larger variety of 'mobility events' (Robinson-García et al., 2019) in affiliation data; and 2) to translate this terminology into sequences that can be further used for the calculation of mobility indicators.

[1] Publication metadata contains information such as publication year, author name, and author address(es), including the country and organization of affiliation. A single publication's metadata does not yet provide a lot of information about an author's career. Combining the publication years and affiliation addresses in an author's records across a publication database results in a more comprehensive picture of the author's movements.

Methodology

From affiliation data to mobility events
In this study we conceptualize a *mobile researcher* as any researcher with more than one affiliation in the period of analysis. This turns into international mobility when the researcher is affiliated with research institutions in more than one country.

Affiliation data from multiple records provides various combinations of e.g., a country and a publication year and thus allows for looking forward – to what happens in the next year – and backwards, to what has happened in the previous year. When there is no other previous country anymore, we have arrived at the 'country of origin' or the country of the first publication(s) of a researcher (Robinson-García et al., 2019). There are different ways of determining a researcher's career start, but here we will take a researcher's first country as the country in which he or she started their publishing activity. A (internationally) mobile researcher will then have one or more countries of destination, which are the countries of affiliation that are different from the country of origin.

A researcher can be affiliated to multiple countries of origin or destination. When these *multiple* affiliations happen at the same time, we call them *simultaneous*, and when they follow on each other without overlap, they are *sequential*. Thus, a researcher relationship with a country can be characterized by different mobility events determined by the presence of publication activity at given points in time with other countries.

From mobility events to career footprints
The terminology outlined above can be operationalized by accounting for all possible combinations of mobility events. To achieve this, we translate the mobility events into 'codes' that can be accumulated in 'footprints' of the relationship of an author with her countries of affiliation. This notation combines information on the state of affiliation of a researcher with a given country and with respect to other countries. The codes are the following:

- A '1' indicates that a researcher is fully (exclusively) publishing in a country.
- A '0' indicates that a formerly active researcher is not active with a country, but active in another country. This is a break in the relationship with the country.
- A 'b' indicates a year of activity of a researcher 'before' starting to publish with a country. It also denotes a sequential move of a researcher, marking the period before the researcher becomes a 'newcomer' in a country.
- An 'f' indicates activity in a country, but with simultaneous activity in another country (a 'fractional' relationship with the country).

The single-digit/letter codes are concatenated in chronological order to form strings. Every such string is what we call a *mobility footprint* that represents the entire activity of a researcher in relation with a country. This footprint is sufficient for characterizing the researcher-country relationship, including whether a researcher is a newcomer or affiliated with a country of origin.

Illustration of application

An example may clarify what the footprints look like and how they allow for the description of different mobility situations. Table 1 shows a researcher who is affiliated with three countries in total and has registered changes of affiliation at four points in time.

Table 1. Exemplary mobility footprints for a researcher affiliated with three countries.

Researcher	Country	t1	t2	t3	t4	Footprint
Researcher 1	C1	1	0	0	1	1001
Researcher 1	C2	b	1	f	0	b1f0
Researcher 1	C3	b	b	f	0	bbf0

Country C1 is considered the country of origin of Researcher 1 because it is his country of affiliation at t1. At t2, he has completely stopped publishing with C1 and becomes a 'newcomer' for C2. This is denoted with a '0' code in the relationship of Researcher 1 with C1 at t2, and with a '1' in the relationship with C2 at t2. At t3, Researcher 1 establishes a simultaneous affiliation with C3, fractionalizing his activity between C2 and C3. This is marked with code 'f'. Finally, at t4, the researcher returns to C1, while leaving C2 and C3 again. Also, the years of activity before being active in C2 and C3 are denoted by code 'b'.

Based on these conventions, mobility footprints can be created for all the researchers in a bibliometric database and may then be used for the calculation of mobility indicators at e.g., the country level. For example, an indicator that captures 'brain drain' could be the ratio of researchers having left their country of origin and (non-origin) researchers having entered a country. To identify them, we simply need to search for all the researchers of a country of origin with a '0' at the t that we are interested in, and all the researchers with a '1' or 'f' (at the t of interest) for whom the country is not their country of origin. Figure 1 shows what this could look like when calculated for European countries based on Web of Science data.

Figure 1. Exemplary indicator of 'brain drain' per country at the end of the period 2009-2019.

Discussion

In this paper we propose a framework for notating and organizing the various mobility events that can be deduced from bibliometric affiliation data. This approach allows for the large-scale analysis of researcher mobility at the level of researcher addresses and is more versatile than previous attempts. It is an important step towards indicators of researcher in- and outflows and of residing researchers, with researchers including any kind of publishing author.

Limitations of this approach can be the completeness of the database used and the quality of the disambiguation of authors. Also, mobility can only be captured for those researchers who have actually been active as authors.

Next, we will apply the method to Microsoft Academic Graph data, convert affiliation information into footprints, and create and calculate indicators to map global researcher mobility. Alongside this, we plan to share the SQL-code for the creation of the footprints and indicators, as well as the in- and output data, to allow the application of this method for anyone. Future studies may go beyond the researcher-country combination and include other entities: research organizations, geographical regions, or even journals and fields of science.

Acknowledgements

The method presented was developed in the context of the project 'Knowledge ecosystems in the new ERA', which was funded by the European Commission.

References

Abbott, A. (1990). A Primer on Sequence Methods. *Organization Science, 1*(4). 375-392.

Abbott, A., & Tsay, A. (2000). Sequence Analysis and Optimal Matching Methods in Sociology. Review and Prospect. *Sociological Methods & Research, 29*(1). 3-33.

El-Ouahi, J., Robinson-García, N., & Costas, R. (2021). Analyzing scientific mobility and collaboration in the Middle East and North Africa. *Quantitative Science Studies, 2*(3). 1023-1047.

Moed, H.F., Aisati, M., & Plume, A. (2013). Studying scientific migration in Scopus. *Scientometrics, 94.* 929–942.

Robinson-García, N., Sugimoto, C. R., Murray, D., Yegros-Yegros, A., Larivière, V., & Costas, R. (2019). The many faces of mobility: Using bibliometric data to measure the movement of scientists. *Journal of Informetrics, 13*(1). 50-63.