Untangling the complex pattern of international talent mobility in science

Meijun Liu[*], Alex J. Yang[**], Yi Bu[***], Yifan Tian[***] and Liying Guo[****]

[*] *meijunliu@fudan.edu.cn*
0000-0002-2800-5511
Institute for Global Public Policy, Fudan University, Shanghai, China

[**] *alexjieyang@outlook.com*
0000-0002-1276-0483
School of Information Management, Nanjing University, Nanjing, China

[***] *buyi@pku.edu.cn; tianyifan@stu.pku.edu.cn*
0000-0003-2549-4580
Department of Information Management, Peking University, Beijing, China

[****] *lyguo24@m.fudan.edu.cn*
0009-0007-2903-9565
Institute for Global Public Policy, Fudan University, Shanghai, China

This paper presents a comprehensive analysis of global talent mobility networks spanning from 1950 to 2020, utilizing data encompassing country-level, city-level, and institutional-level networks. Leveraging network science methodologies, we investigate both Small-world and Scale-Free attributes within these networks, shedding light on their structural complexities and dynamic patterns over time. Our findings reveal increasing interconnectedness and efficiency in talent mobility, as evidenced by rising average clustering coefficients and decreasing average path lengths. Additionally, we identify Scale-Free patterns characterized by heterogeneous distributions of node degrees, highlighting the presence of influential hubs within the mobility networks. These insights offer valuable perspectives on the evolving landscape of global talent exchange and collaboration within the scientific community.

1. Introduction

Understanding the dynamics of talent mobility is paramount for comprehending collaboration patterns and knowledge dissemination within the global scientific community (Azoulay et al., 2017; Pappalardo et al., 2023; Shi et al., 2023; Sugimoto et al., 2017). Over the past decades, advances in network science have facilitated the exploration of complex network structures underlying international talent mobility (Alessandretti et al., 2020; Schläpfer et al., 2021). In this study, we undertake a rigorous analysis of global talent mobility networks from 1950 to 2020, examining country, city-, and institutional-level networks. Through the lens of network science methodologies, we delve into both small-world and scale-free attributes within these networks, aiming to elucidate their structural characteristics and dynamic patterns over time.

By uncovering the structural complexities and dynamic patterns within global talent mobility networks, our study contributes to a deeper understanding of the mechanisms driving collaboration and innovation within the scientific community (Azoulay et al., 2018; Fortunato et al., 2018; Liu et al., 2023). These insights hold significant implications for policymakers, researchers, and stakeholders seeking to foster international collaboration and harness the full potential of global talent networks.

2. Identifying mobility of scientists

This study undertakes a meticulous examination of academic literature spanning the period from 1950 to 2020, drawing from the MAG database. Central to its investigation is the nuanced

exploration of talent mobility, with a keen focus on tracking shifts in authors' affiliations over time. Through the systematic analysis of each author's publication corpus using AuthorID, coupled with the corresponding affiliations for each paper, this research unveils insightful patterns. Authors' papers are meticulously arranged chronologically, facilitating the identification of mobility instances based on changes in affiliations between successive publications. If the affiliation of the author in the i^{th} paper is different from that in the $(i-1)^{th}$ paper, there is mobility, i.e., a change of affiliations, in the year of t where the author published the i^{th} paper. If the country of the origin affiliation is different from that of the destination affiliation, the change of affiliations is considered an international mobility.

Significantly, a total of 27,747,156 instances of talent mobility are unearthed, implicating 3,621,647 individuals within the scientific community. This extensive analysis encompasses a diverse array of mobility events spanning 198 countries and involving 17,183 institutions. Noteworthy emphasis is placed on international talent mobility, which reveals 9,265,316 instances of such events, implicating 1,617,001 scientific professionals. This international dimension underscores the study's focus on the global dynamics of talent mobility within the scientific domain.

Figure 1a illustrates the exponential growth of international talent mobility events across the fields of Art and Humanities, Science and Engineering, and Social Sciences from 1950 to 2020. It vividly portrays the escalating trend in mobility occurrences within each domain over the specified period. In Figure 1b, the proportion of international talent mobility relative to total mobility events is depicted. Notably, North America exhibits the smallest proportion of international talent mobility (28%).

Figure 1c provides insight into the ranking of the top 50 countries by their international talent mobility count from 1990 to 2020. Noteworthy trends emerge, including the consistent dominance of the United States as the leading country throughout the period under consideration. The United Kingdom initially holds the second position, subsequently yielding to China after 2006. Germany experiences a decline in its ranking from the top 3 to the top 4, while Japan witnesses a notable decrease in rank. Conversely, China and Singapore demonstrate significant ascents in their rankings, indicative of their growing prominence in international talent mobility during the specified timeframe. These observations underscore the evolving dynamics and shifting global landscape of talent mobility across nations.

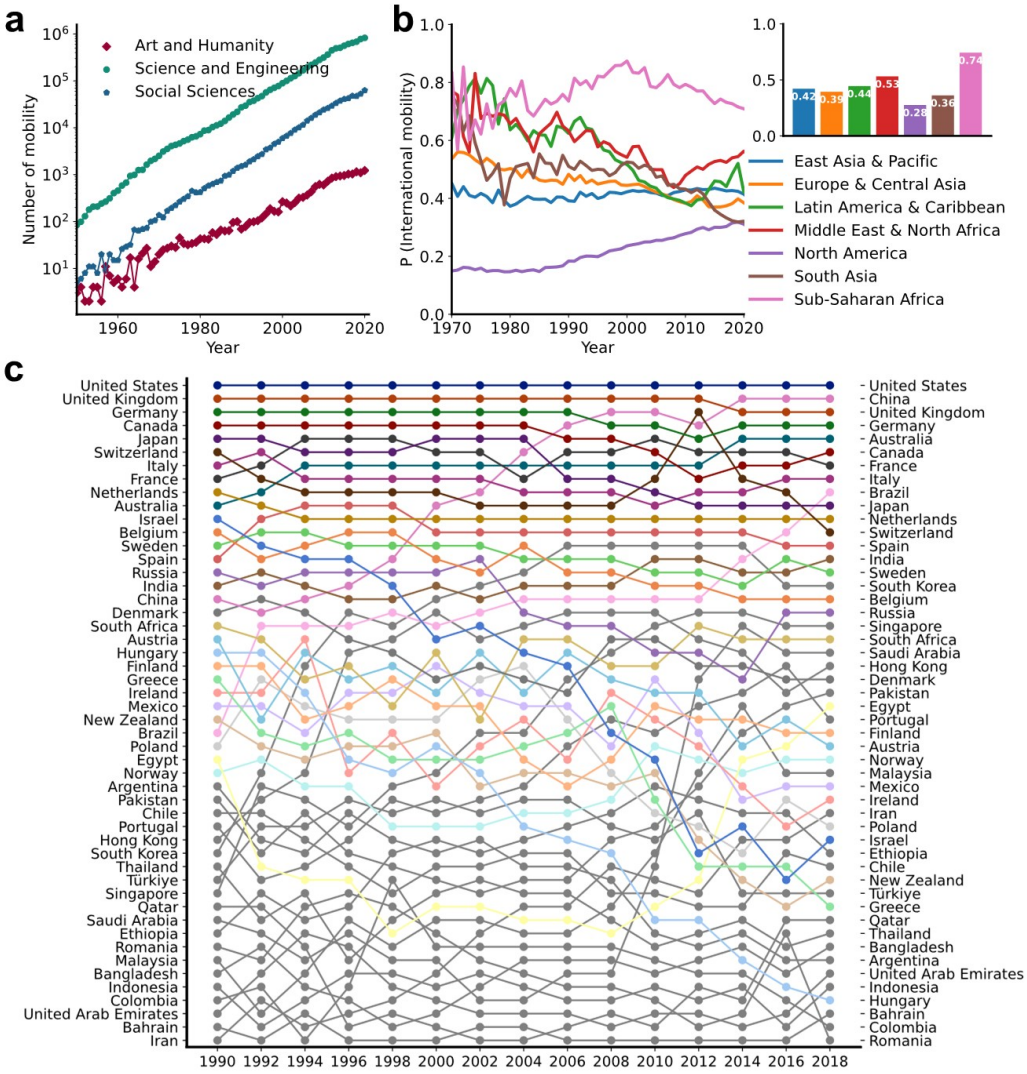


Figure 1: Number, proportion and country/region-level rank of international talent mobility.

3. Methods

3.1 Scale and intensity

Comprehending the flux of scientific talent across geographical boundaries is paramount for analysing collaboration patterns and knowledge dissemination. Utilizing network science methodologies, this analysis delves into the structure and dynamics of global scientific talent mobility networks. Key network metrics are herein explored to illuminate the scale and intensity of these networks. Within the context of talent mobility networks, these metrics provide invaluable insights into the structure, dynamics, and cohesion of the global scientific community.

(1) **The number of edges** denotes the aggregate frequency of talent movements originating from all countries, cities, or institutions within the network.

(2) **The average degree** ($\hat{k}$) signifies the average number of talent movements involving each country, city, or institution, calculated as:

$$\hat{k} = \frac{\sum_{i=1}^{N} k_i}{N} \quad (1)$$

(3) **The diameter** represents the maximum distance between any two nodes within the talent mobility network:

$$Diameter = Max\left(d_{i,j}\right) \quad (2)$$

(4) **The density** of a network is defined as the ratio of the number of edges to the maximum possible number of edges (M):

$$Density = \frac{M}{N\left(N\text{-}1\right)/2} \quad (3)$$

(5) **The number of connected components** denotes the count of distinct subnetworks within the talent mobility network, highlighting its structural segmentation. A connected component is a subset of nodes in a graph such that every pair of nodes in the subset is connected by a path.

(6) **The maximum connected component size** of a network refers to the size of the largest connected component within that network. This metric illustrates the extent of cohesive connectivity within the talent mobility network, emphasizing the prominence of major hubs or clusters of talent exchange.

3.2 Small-world attributes

Beyond mere assessment of network scale and intensity, delving into the "small-world" properties of talent mobility networks offers additional enriching insights. These networks often manifest characteristics that bridge the divide between highly localized and globally connected structures. Small-world networks typically demonstrate a low average path length coupled with a high clustering coefficient. This suggests that researchers can capitalize on the dense local connections within their immediate network (high clustering coefficient) while still having the capacity to access distant collaborators or knowledge reservoirs through a relatively small number of intermediary institutions (low average path length).

(1) **The average path length ($\hat{d}$)** characterizes the connectivity and accessibility within a network, specifically within the context of a global talent mobility network. It provides insight into the average number of steps or hops required for information or individuals to travel between any two nodes (e.g., countries, cities, or institutions) within the network. In the realm of global talent mobility, this metric helps to quantify the overall interconnectedness and ease of movement across geographical and institutional boundaries. In the context of a global talent mobility network, the average path length is calculated as the mean of the shortest paths between all pairs of nodes within the network. Formally, it is defined as the sum of the shortest path lengths divided by the total number of pairs of nodes:

$$\hat{d} = \frac{\sum_{i \geq j} d_{ij}}{N\left(N\text{-}1\right)/2} \quad (4)$$

where d_{ij} represents the shortest path length between node i and node j.

(2) **The average clustering coefficient** provides a measure of the degree to which nodes in the network tend to cluster together. It quantifies the extent to which nodes with common connections tend to form clusters or communities within the network. A high average clustering coefficient indicates that nodes in the network are densely interconnected, forming tight-knit communities, while a low coefficient suggests a sparser and more decentralized network structure. The average clustering coefficient of a global talent mobility network is calculated by averaging the clustering coefficients of all nodes in the network. The clustering coefficient of a node measures the proportion of connections between its neighbours relative to the total

possible connections among them. Mathematically, the average clustering coefficient is given by the formula:

$$\widehat{C} = \frac{\sum_{i=1}^{N} \frac{E_i}{(k_i(k_i - 1))/2}}{N} \quad (5)$$

where E_i is the number of edges between the neighbours of node i , k_i of degree of node i , and $C_i = \frac{E_i}{(k_i(k_i - 1))/2}$ denotes the clustering coefficient of node i .

3.3 Scale-Free attributes

In the context of a global talent mobility network, scale-free network attributes and power law fitting provide insights into the distribution of connections among nodes, highlighting the presence of highly connected hubs. Scale-free networks exhibit a heterogeneous distribution of node degrees, where a few nodes (hubs) possess a disproportionately large number of connections, while the majority of nodes have relatively few connections. Power law fitting is employed to characterize this distribution, revealing the relationship between the frequency of nodes with a given degree and the degree itself. Understanding scale-free network attributes aids in identifying influential nodes and understanding the resilience and robustness of the mobility network.

In a global talent mobility network, scale-free attributes refer to the presence of a power-law distribution of node degrees. This distribution implies that the probability $P(k)$ of encountering a node with degree k follows a power law:

$$P(k) = k^{-\gamma} \quad (6)$$

where γ is the power-law exponent, typically ranging between 2 and 3 in scale-free networks.

4. Global mobility networks

Figure 2 visually depicts the evolution of global talent mobility by illustrating all mobility lines in the years 1990, 2000, 2010, and 2020. A notable observation is the discernible increase in both the quantity and complexity of mobility patterns over time. This growth signifies not only a numerical expansion but also a heightened diversity and intricacy in the mobility dynamics observed. Furthermore, it underscores the increasing involvement of a greater number of institutions and countries in global talent mobility endeavours. It also highlights the evolving nature of talent mobility on a global scale, indicating a trend towards greater interconnectedness and collaboration across diverse academic and research entities.

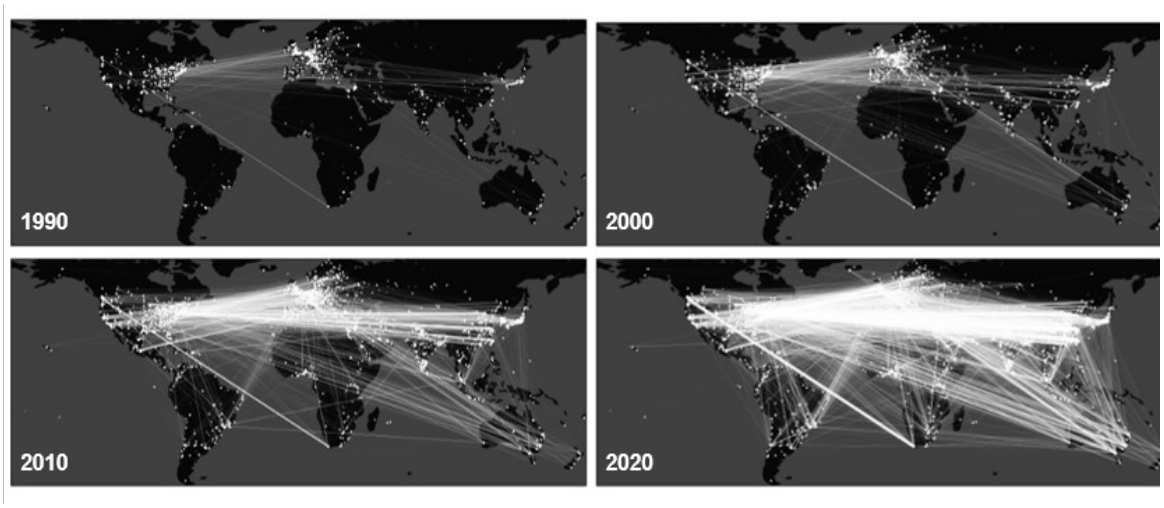


Figure 2: Figure captions should be centred and placed above the figure.

Utilizing complex network analysis tools, a country-level global talent mobility network is constructed based on international talent mobility data. Figure 3 presents the evolution of this network, showcasing significant trends over time. The number of nodes representing countries within the network demonstrates a consistent linear increase, expanding from approximately 20 in 1950 to nearly 200 in 2020. This growth reflects the expanding participation of nations in global talent mobility endeavours.

The structural evolution of the network reveals noteworthy dynamics. In its nascent stages, high-income countries prominently occupy leading positions within the network, indicative of their early engagement and dominance in global talent mobility initiatives. However, a notable shift occurs around the year 2000, where middle-income countries progressively enhance their involvement within the network. This trend underscores the evolving landscape of talent mobility, with middle-income nations increasingly assuming a more active and integral role in global talent exchange and collaboration.

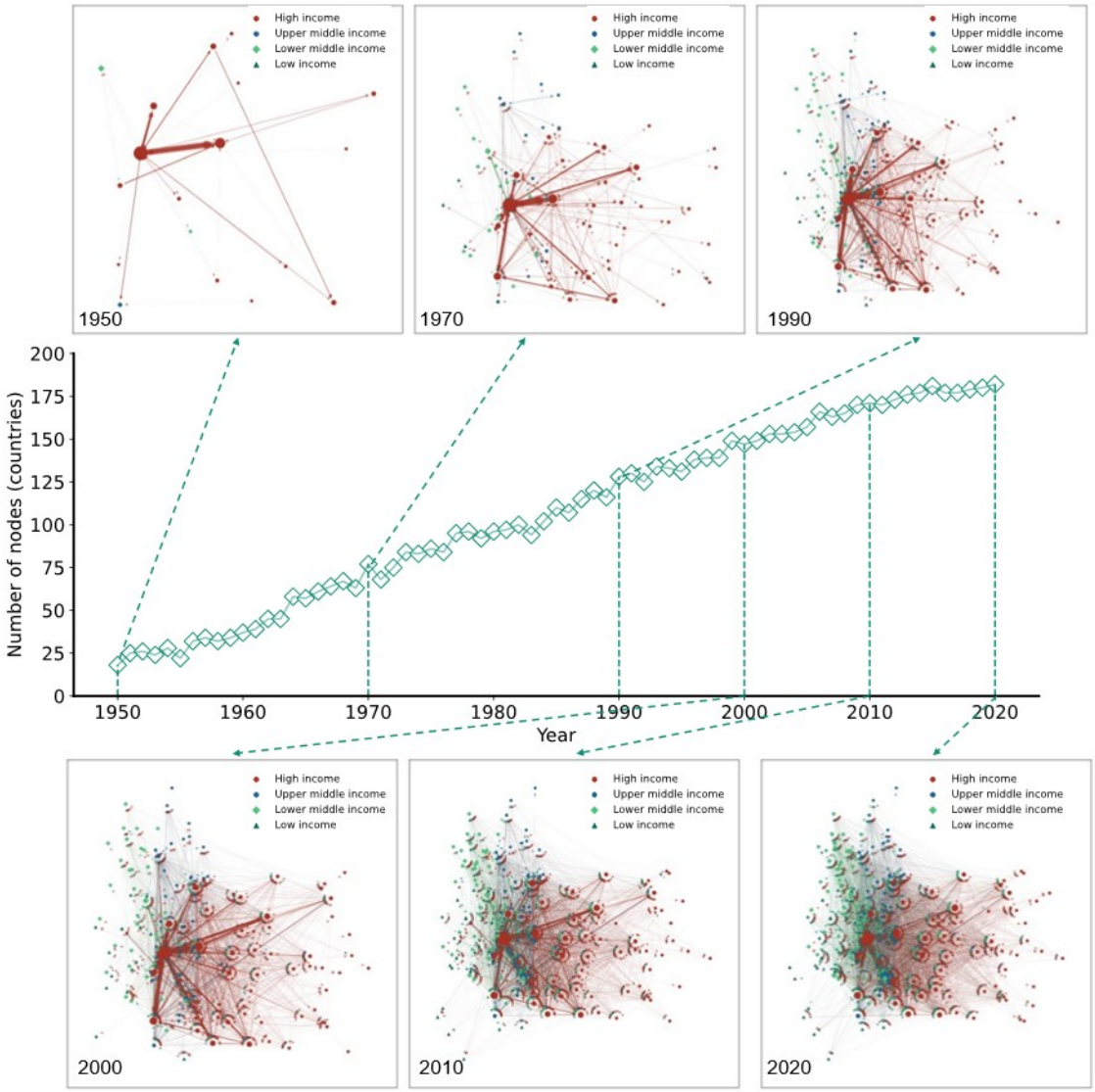


Figure 3: Evolution of country-level global talent mobility network

5. Complex pattern of international talent mobility

5.1 Scale and intensity

Utilizing international talent mobility data spanning from 1950 to 2020, multi-level mobility networks are constructed, including country-l, city-, and institutional-level networks. The dynamic patterns of complexity within these networks are analysed, focusing on key metrics such as the number of edges, average degree, diameter, density, number of connected components, and maximum connected component size.

Figure 4 illustrates the dynamic patterns of scale and intensity across these networks. Across all levels, the number of edges and average degree exhibit exponential growth, indicating increasing engagement of countries, cities, and institutions in global talent mobility. Concurrently, the diameter of city-level and institutional-level networks decreases from 1960 onwards, suggesting a closer proximity of nodes within these networks.

The density metric serves as a quantitative measure of network sparsity or density, reflecting the intensity of connections among nodes. Over time since 1960, the density of country-level and city-level networks demonstrates an increasing trend, signifying tighter interactions among network members in terms of talent exchange. Conversely, the density of institutional-level networks experiences a decrease.

Moreover, both the number of connected components and the maximum connected component size for all networks witness a steady increase over time. This trend indicates a growing complexity in international talent mobility, with networks becoming more fragmented yet also featuring larger cohesive components. These findings highlight the evolving dynamics of international talent mobility across different hierarchical levels, providing valuable insights into the complexities underlying global talent exchange processes.

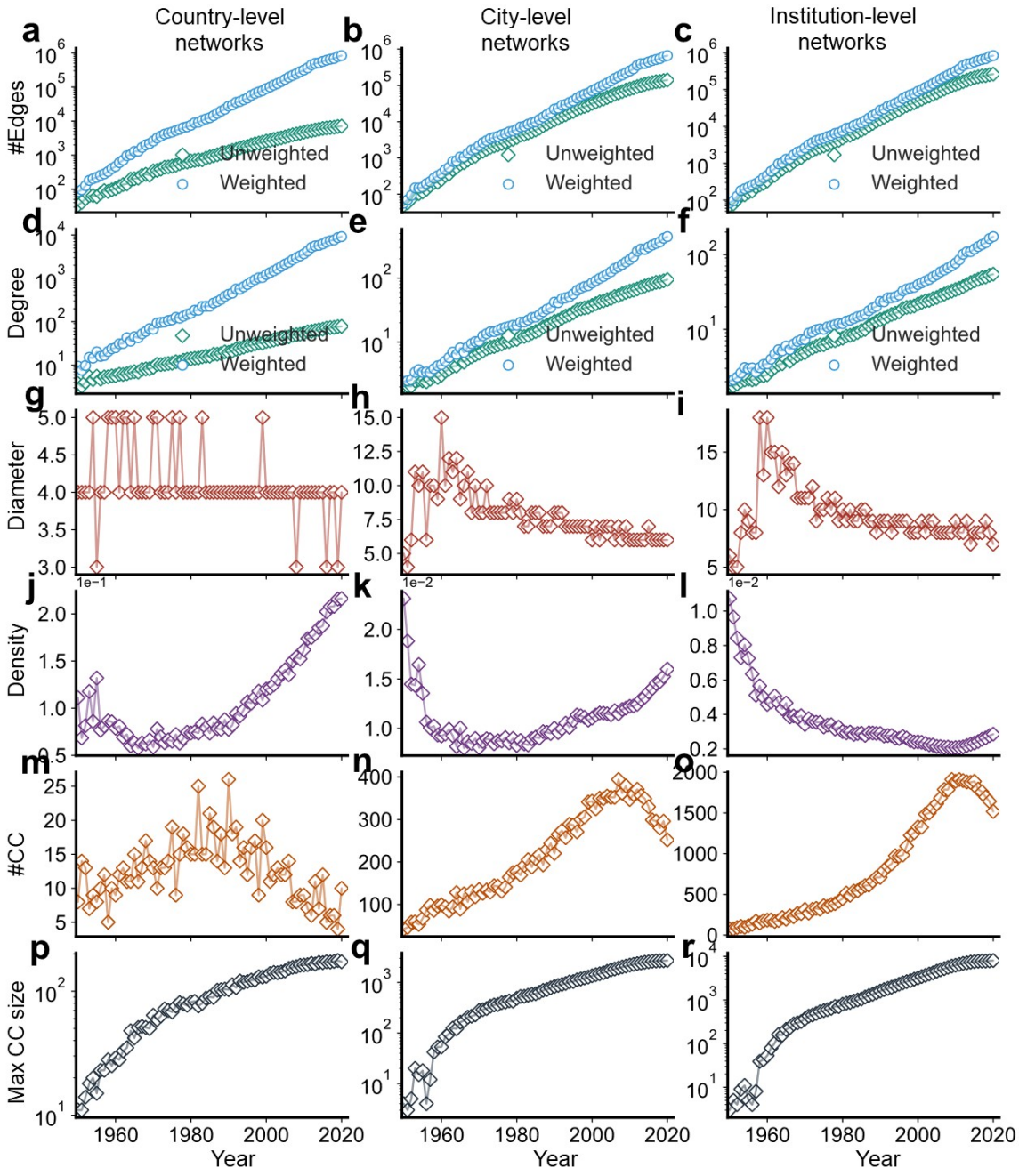


Figure 4: Dynamic patterns of the scale and intensity for multi-level mobility networks.

5.2 Small-world and Scale-Free patterns

The Small-world patterns of global talent mobility are scrutinized, with results presented in Figure 5a-f. The average clustering coefficient of country-level, city-level, and institutional-level networks exhibits a steady increase over time. This trend indicates denser interconnections among nodes within these networks, fostering cohesive clusters and communities.

Furthermore, the average path length serves as a quantitative measure of the network's efficiency in transmitting information or facilitating individual mobility. A lower average path length signifies a more efficient network, characterized by closely interconnected nodes facilitating rapid communication or mobility. Conversely, a higher average path length suggests greater separation between nodes, potentially impeding the flow of information or talent. Our analysis also reveals a decrease in the average path length of country-level, city-level, and

institutional-level networks over time since 1960. This decline suggests enhanced efficiency in talent mobility and information dissemination across these networks, with nodes becoming closer and communication pathways more direct. Collectively, these two findings signify an increasing manifestation of Small-world patterns over time within global talent mobility networks. The observed trends underscore the evolving efficiency and interconnectedness of these networks, facilitating more effective collaboration and knowledge exchange among diverse nodes.

To examine the Scale-Free pattern of global talent mobility, we utilize mobility data spanning from 1950 to 2020 to construct country-level, city-level, and institutional-level networks. We employ log-log distribution with power law fitting, as depicted in Figure 5g-i.

The power law fitting alpha values provide insights into the distribution of node degrees within each network. For the country-level network, the power law fitting alpha is 2.0, indicating a Scale-Free pattern characterized by a heterogeneous distribution of node degrees. Similarly, for the city-level network, the power law fitting alpha is 2.2, suggesting a similar pattern with slightly higher heterogeneity. Notably, the institutional-level network exhibits a higher power law fitting alpha of 2.9, indicating even greater heterogeneity in the distribution of node degrees within this network. These findings underscore the presence of highly connected hubs within the global talent mobility networks, where a few nodes possess a disproportionately large number of connections while the majority of nodes have relatively few connections. The Scale-Free attributes observed in these networks highlight the importance of influential nodes and contribute to understanding the resilience and robustness of the mobility network structure.

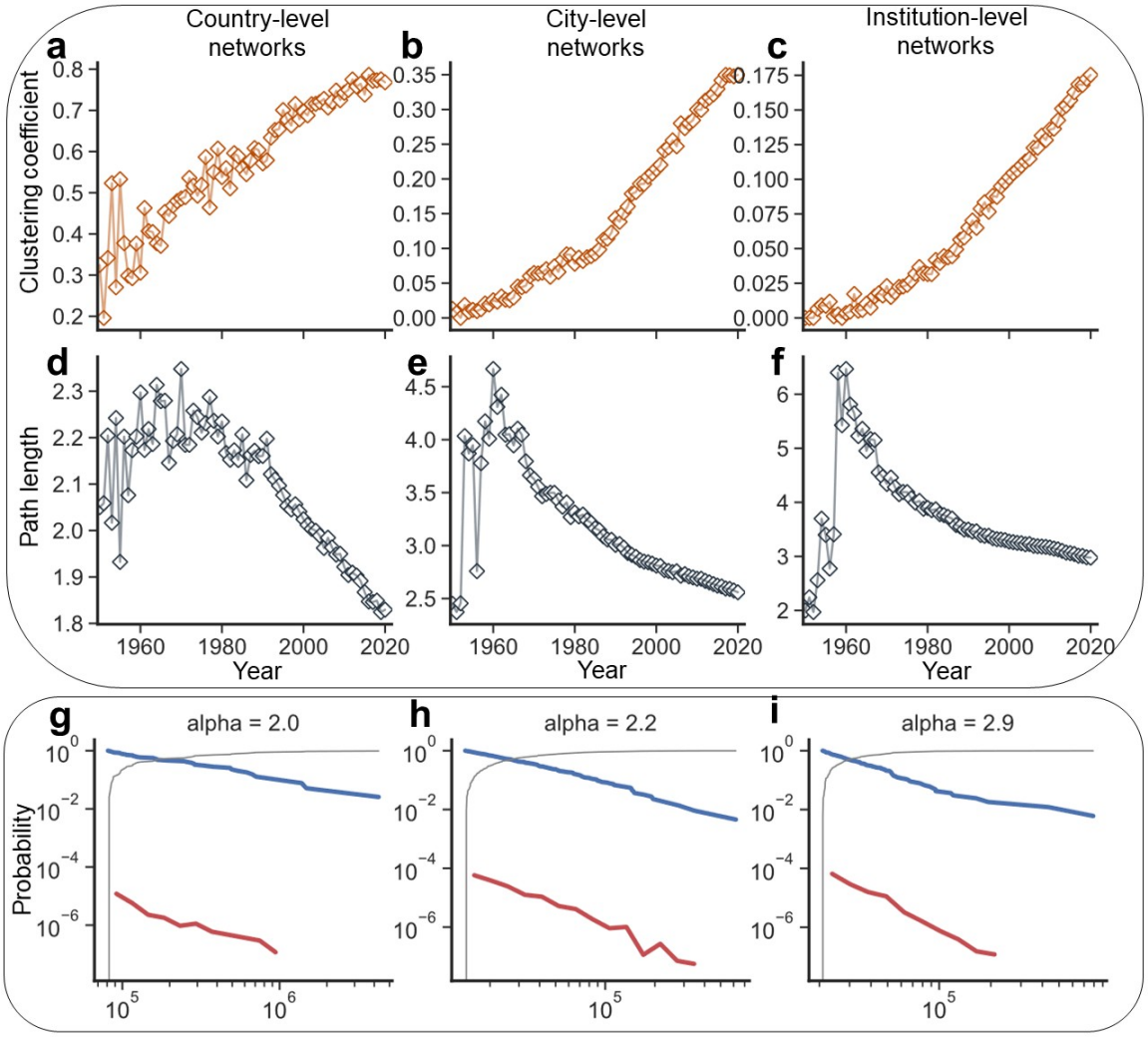


Figure 5: Small-world and Scale-Free patterns of global talent mobility.

6. Conclusion

The analysis of global talent mobility networks provides valuable insights into the structure, dynamics, and patterns underlying international collaboration and knowledge dissemination within the scientific community. By examining country-level, city-level, and institutional-level networks, this study elucidates the complexities inherent in talent mobility processes.

It is important to acknowledge certain limitations of this study. Firstly, the analysis relies on data from the MAG database, which may not capture all instances of talent mobility. Additionally, the study focuses on mobility within the scientific community and may not fully represent talent mobility in other sectors. Furthermore, the analysis is limited by the availability and quality of data, which may vary across different regions and time periods.

Future research could explore additional factors influencing talent mobility, such as cultural, economic, and policy-related factors. Moreover, incorporating qualitative methods could provide deeper insights into the motivations and experiences of individuals engaged in talent mobility. Additionally, longitudinal studies could track the long-term impact of talent mobility on scientific collaboration, innovation, and knowledge creation. Finally, efforts to enhance data collection and analysis methods could further advance our understanding of global talent mobility dynamics.

Open science practices

This study draws upon the rich repository of the Microsoft Academic Graph (MAG) (Wang et al., 2020), recognized for its unparalleled comprehensiveness in scholarly research. Specifically, we utilized the iteration of the MAG dataset released on December 6th, 2021. Spanning from 1800 to 2021, this dataset comprises over 20 million documents, including journal articles, conference proceedings, preprints, and various other research publications.

To provide a comprehensive understanding of the global talent mobility patterns, we focused our analysis on papers published between 1950 and 2020, a period marked by significant transformations in the scientific landscape. This timeframe is aligned with established research paradigms (Lin et al., 2023).

Acknowledgments

This study is sponsored by the National Natural Science Foundation of China (72104054, 72104007), the Shanghai Pujiang Talent program (21PJC026), and the Key Project of the National Natural Science Foundation of China (72234001).

Author contributions

Meijun Liu: Writing-original draft, Conceptualization, Methodology. **Alex J. Yang**: Writing-original draft, Data collection, Methodology, Validation. **Yi Bu:** Conceptualization, Writing-review & editing. **Yifan Tian**: Data collection, Writing-review & editing. **Liying Guo**: Data collection, Writing-review & editing.

Competing interests

Authors are expected to declare their competing interests in a short competing interests section at the end of their paper. Authors may also use this section to declare that they have no competing interests.

References

Alessandretti, L., Aslak, U., & Lehmann, S. (2020). The scales of human mobility. *Nature*, *587*(7834), 402-407. https://doi.org/10.1038/s41586-020-2909-1

Azoulay, P., Ganguli, I., & Graff Zivin, J. (2017). The mobility of elite life scientists: Professional and personal determinants. *Research Policy*, *46*(3), 573-590. https://doi.org/https://doi.org/10.1016/j.respol.2017.01.002

Azoulay, P., Graff-Zivin, J., Uzzi, B., Wang, D., Williams, H., Evans, J. A., Jin, G. Z., Lu, S. F., Jones, B. F., Börner, K., Lakhani, K. R., Boudreau, K. J., & Guinan, E. C. (2018). Toward a more scientific science. *Science*, *361*(6408), 1194-1197. https://doi.org/10.1126/science.aav2484

Fortunato, S., Bergstrom, C. T., Boerner, K., Evans, J. A., Helbing, D., Milojevic, S., Petersen, A. M., Radicchi, F., Sinatra, R., Uzzi, B., Vespignani, A., Waltman, L., Wang, D., & Barabasi, A.-L. (2018). Science of science. *Science*, *359*(6379), eaao0185. https://doi.org/10.1126/science.aao0185

Lin, Y., Frey, C. B., & Wu, L. (2023). Remote collaboration fuses fewer breakthrough ideas. *Nature*, *623*(7989), 987-991. https://doi.org/10.1038/s41586-023-06767-1

Liu, L., Jones, B. F., Uzzi, B., & Wang, D. (2023). Data, measurement and empirical methods in the science of science. *Nature Human Behaviour*. https://doi.org/10.1038/s41562-023-01562-4

Pappalardo, L., Manley, E., Sekara, V., & Alessandretti, L. (2023). Future directions in human mobility science. *Nature Computational Science*, *3*(7), 588-600. https://doi.org/10.1038/s43588-023-00469-4

Schläpfer, M., Dong, L., O'Keeffe, K., Santi, P., Szell, M., Salat, H., Anklesaria, S., Vazifeh, M., Ratti, C., & West, G. B. (2021). The universal visitation law of human mobility. *Nature*, *593*(7860), 522-527. https://doi.org/10.1038/s41586-021-03480-9

Shi, D., Liu, W., & Wang, Y. (2023). Has China's Young Thousand Talents program been successful in recruiting and nurturing top-caliber scientists? *Science*, *379*(6627), 62-65. https://doi.org/10.1126/science.abq1218

Sugimoto, C. R., Robinson-Garcia, N., Murray, D. S., Yegros-Yegros, A., Costas, R., & Larivière, V. (2017). Scientists have most impact when they're free to move. *Nature*, *550*(7674), 29-31. https://doi.org/10.1038/550029a

Wang, K., Shen, Z., Huang, C., Wu, C.-H., Dong, Y., & Kanakia, A. (2020). Microsoft Academic Graph: When experts are not enough. *Quantitative Science Studies*, *1*(1), 396-413. https://doi.org/10.1162/qss_a_00021