Pattern Recognition of Scholars' Dynamic Academic Impact Based on Collaboration Networks (Postprint)

Authors: Fan Ruxia, Zeng Jianxun, Gao Yaruixi

Date: 2017-11-08T00:00:00+00:00

Abstract

Purpose: To identify high-collaboration scholars within a team and their dynamic academic influence patterns using a high-collaboration scholar identification algorithm and a scholar influence pattern recognition algorithm, thereby providing references for talent development within the team. Method: High-collaboration scholars within the team were distinguished based on their collaboration counts; the publication output and degree centrality metrics of high-collaboration scholars were utilized to measure scholars' individual influence and influence within the team, thereby identifying their dynamic academic influence patterns. Results: The number of high-collaboration scholars varies across different teams, ranging from zero to multiple. The dynamic academic influence patterns of high-collaboration scholars differ and are identified as either steady growth or mature fluctuation patterns. Limitations: Only two metrics were employed to measure scholar influence; for scholars with more complex circumstances, additional metrics need to be incorporated to identify their dynamic academic influence patterns. Conclusion: The high-collaboration scholar identification algorithm and scholar influence pattern recognition algorithm can reasonably identify high-collaboration scholars within a team and their dynamic academic influence patterns.

Full Text

Preamble

Recognizing Dynamic Academic Impacts of Scholars Based on Cooperative Network

Fan Ruxia[1], Zeng Jianxun[1], Gao Yaruixi[2]
[1](Institute of Scientific and Technical Information of China, Beijing 100038, China)

[2](China Defense Science and Technology Information Center, Beijing 100142, China)

Abstract

[**Objective**] This study aims to identify highly cooperative scholars within research teams and characterize their dynamic academic impact patterns, providing valuable references for talent development. [**Methods**] We first distinguished highly cooperative scholars based on their collaboration patterns. Then, we measured individual and team-level influence using publication output and degree centrality metrics to identify dynamic academic impact patterns. [**Results**] The number of highly cooperative scholars varied across teams, ranging from zero to multiple individuals. Their dynamic academic impact patterns fell into two categories: steady growth or mature fluctuation. [**Limitations**] Only two indicators were employed to measure scholarly impact; more complex cases would require additional metrics for comprehensive pattern identification. [**Conclusions**] The proposed algorithms for identifying highly cooperative scholars and their influence patterns can effectively recognize key collaborative researchers and track their evolving academic impacts within teams.

Keywords: Academic Impact; Social Network Analysis; Dynamic Pattern

Introduction

Research on the evolution and evaluation of scholars' academic influence holds significant importance for research management and talent development. As research publications represent a primary output of scholarly activity, monitoring and evaluating academic impact through publication records has become a standard approach. Traditional evaluation methods primarily include peer review, publication counts (total and average), citation metrics (total citations, average citations, impact factor), and combined indicators such as the H-index and G-index. In recent years, social network analysis has gained traction in scholarly evaluation by assessing collaboration patterns and measuring researchers' positions within co-authorship networks.

Academic influence is inherently dynamic, and only through longitudinal analysis can we fully understand talent development trajectories, predict future directions, and inform talent cultivation strategies. However, existing research predominantly adopts static perspectives, with few studies examining dynamic patterns. Those that do incorporate temporal dimensions typically focus on citation trends over time. Moreover, research on scholars' evolving influence within teams remains limited, particularly in teams with multiple high-impact scholars. Tracking their dynamic performance can provide deeper insights into their development, identify rapidly rising collaborators, and guide targeted support for talent cultivation.

This study addresses these gaps by leveraging author collaboration networks.

We propose algorithms to identify highly cooperative scholars and track their influence patterns, enabling the discovery of rapidly growing collaborators for talent selection purposes.

Literature Review

Co-authorship networks have proven valuable for assessing scholarly influence. Thomson Reuters, for instance, identifies research teams through co-authorship relationships to award "Research Team Awards" to Chinese scholars. The critical step in using collaboration networks for influence measurement is team identification, typically achieved through clique analysis or n-clique analysis of co-authorship networks. While these methods effectively identify tightly-knit groups, some researchers have sought more stable and complete teams by first identifying small groups using 2-clique analysis, then applying a snowball sampling approach to expand these groups within the broader network. Although this method yields larger teams encompassing various cliques (2-clique, 3-clique, 4-clique, etc.), the resulting collaboration relationships tend to be sparse and unstable. Our approach first identifies stable, complete teams before discovering highly cooperative scholars, addressing limitations of methods that pre-select scholars before defining teams.

Regarding methods for identifying high-impact scholars, Yan et al. demonstrated that centrality measures in collaboration networks can indicate a scholar' s influence within their field. Abbasi et al. examined relationships between network metrics (degree centrality, closeness centrality, betweenness centrality, eigenvector centrality, average connection strength, and efficiency) and citation-based impact indicators (H-index, G-index), showing that network characteristics can predict future academic performance. Wang Feifei identified high-impact authors in knowledge exchange networks using centrality measures, while Zhu Tian et al. ranked important authors based on collaboration quantity and scope. However, these studies primarily provide "snapshot" evaluations of overall collaboration states, lacking temporal analysis of influence evolution and comprehensive understanding of talent development trajectories.

Some scholars have introduced temporal dimensions for dynamic impact assessment using personal impact factors, hr(y) index, and other methods. Personal impact factors inherit limitations from journal impact factors, including the problematic practice of using short-term average citations. The hr(y) index, derived from the H-index, suffers from the H-index' s inherent drawback of poor discrimination among average performers. Currently, dynamic impact indicators remain underexplored. Many universities use publication counts for promotion evaluation, such as Zhejiang University' s professor appointment criteria, which involve statistical assessment of research output. Degree centrality is frequently employed to measure scholars' network positions and academic relationship influence. This study combines publication output and degree centrality to measure individual and team-level influence across time, supplementing dynamic academic impact assessment.

Methodology

This research adopts a team-level perspective, identifying highly cooperative scholars based on collaboration network size within sub-networks. We then analyze their dynamic academic influence to discover rapidly growing collaborators. Degree centrality represents team influence, while publication output represents individual influence. By tracking temporal changes in these metrics, we classify dynamic patterns as either steady growth or mature fluctuation.

Key Definitions:

Cooperation Sub-network: Within the overall collaboration network, a stable sub-network where scholars have collaborated more than twice.

Scholar Cooperation Count: In a co-authorship network where nodes represent authors and edges represent collaborations, the number of edges directly connected to a node defines that scholar' s cooperation count.

Normalized Degree Centrality: Degree centrality normalized using min-max standardization to enable comparison across different network sizes.

Algorithm 1: High-Cooperation Scholar Identification

This algorithm identifies highly cooperative scholars and characterizes their quantity features within teams. The steps are as follows:

1. Partition the overall collaboration network into cooperation sub-networks using n-clique analysis. To analyze stable and complete teams, we employ a trial method: start with n=2. If m subgroups are identified at n-clique analysis and w subgroups at (n+1)-clique analysis, and at least 50% of the m subgroups maintain stable membership with corresponding subgroups in w, then n-clique analysis is adopted. Otherwise, increment n by 1 and repeat.

2. Construct a non-0-1 collaboration matrix M(ai,j) for each sub-network, where ai,j represents the number of collaborations between authors i and j (greater than 2).

3. For each node i in M, calculate the cooperation count vector C, where element Ci is computed by initializing Ci to 0 and incrementing by 1 for each collaboration partner j.

4. If there exists a unique scholar i where Ci/Z $\geq$ 50% (Z = total scholars in the network), then only one highly cooperative scholar exists–scholar i who collaborates with over half the team.

5. If multiple scholars i, j, k (within 20% of team size) satisfy (Ci + Cj + Ck)/Z $\geq$ 80% (with scholars counted only once even if they collaborate with multiple high-cooperation scholars), then multiple highly cooperative scholars exist, following the Pareto principle (80/20 rule). Scholars are

ranked by cooperation count; if counts are equal, frequency determines priority.

6. If neither condition is met, collaborations are balanced and no highly cooperative scholars exist.

[Figure 1: see original paper]

Algorithm 2: Scholar Influence Pattern Recognition

After identifying highly cooperative scholars, this algorithm detects their dynamic academic influence patterns:

1. For each identified scholar, construct annual collaboration networks across the study period. For example, to analyze dynamic influence from year m to n, build m-n+1 separate networks. Author order weights are considered (first author weight w1, second author w2, etc.).
2. Calculate the average normalized degree centrality across all m-n+1 years and compare it with the average of the most recent (m-n)/2 years (or (m-n+1)/2 if odd). Degree centrality incorporates author order weights: if scholar i collaborates with scholar A as first author A1 times, second author A2 times, etc., the weighted collaboration count is A1×w1 + A2×w2 + A3×w3 + ···The total degree centrality Ci is the sum of weighted collaborations with all partners.
3. Calculate average publication output across m-n+1 years and the most recent (m-n)/2 years, with publications weighted by author order (F1×w1 + F2×w2 + F3×w3 + ···).
4. Pattern classification: Let Di = (Ci - Cmin)/(Cmax - Cmin) be normalized degree centrality. If both average Fi and average Di in the recent period exceed the long-term averages, the pattern is **steady growth**. If both are lower, the pattern is **mature fluctuation**.

[Figure 2: see original paper]

Data Source

For feasibility and data accessibility, we constructed a co-authorship network in the library and information science field to identify highly cooperative scholars and their dynamic impacts. The dataset comprised 24,711 papers published between 2011-2015 in 18 CSSCI-indexed journals, involving 17,486 scholars. The collaboration network was built by connecting all pairs of co-authors in each paper, yielding 9,859 authors and 26,710 co-authorship ties. To focus on stable collaborations, we restricted the analysis to scholars with more than two collaborations, resulting in 1,454 scholars and 1,485 ties. While author contribution weights could be considered, existing indices like Nature Index and ESI treat all authors equally; we therefore assigned equal weights (wi = 1) to avoid subjectivity and discipline-specific complications.

Results and Analysis

High-Cooperation Scholar Identification

Applying n-clique analysis to the stable collaboration network, we determined n=4 through trial-and-error, yielding 227 sub-networks. When n=5, 126 sub-networks maintained identical membership, confirming that 4-clique analysis produced stable, complete teams. Running the identification algorithm across all 227 sub-networks revealed varying numbers of highly cooperative scholars per team (see).

Single High-Cooperation Scholar Example: In Zhu Qinghua's sub-network ([Figure 3: see original paper]), Zhu had 15 collaborators while Zhao Yuxiang had 7 in a 19-member network. Since Zhu' s collaboration count exceeded 50% of the team, he was identified as the sole highly cooperative scholar.

Multiple High-Cooperation Scholars Example: In Huang Shuiqing' s 15-member sub-network ([Figure 4: see original paper]), no single scholar exceeded 50% collaboration. However, the top three scholars (Huang Shuiqing, He Lin, Wang Dongbo) within 20% of team size had combined unique collaborations of 14/15 (>80%), satisfying the 80/20 rule. Thus, all three were identified as highly cooperative scholars.

Scholar Influence Pattern Recognition

Analyzing the three highly cooperative scholars in Huang Shuiqing' s team, we constructed 5-year (2011-2015) and 2-year (2014-2015) average networks ([Figure 5: see original paper] and [Figure 6: see original paper]). The 5-year average reflects long-term performance, while the 2-year average captures recent activity. Publication output represents research productivity, and normalized degree centrality indicates network position importance across differently-sized teams.

Steady Growth Pattern: He Lin' s average publication output and normalized degree centrality in 2014-2015 exceeded the 5-year averages (), indicating enhanced recent productivity and more prominent network position. This constitutes a steady growth pattern.

Mature Fluctuation Pattern: Huang Shuiqing and Wang Dongbo showed decreased average publication output in 2014-2015 compared to the 5-year average (). While Wang Dongbo' s average collaboration frequency increased, his normalized degree centrality declined, indicating reduced relative importance in the recent network. Huang Shuiqing experienced slight decreases in both metrics. Both scholars exhibit mature fluctuation patterns.

Validation using total citation frequencies across years () supports these findings. Despite citation lag (2015 papers show zero citations), He Lin' s recent papers received substantially more citations than earlier work, while Huang Shuiqing and Wang Dongbo's recent citations declined, confirming the identified patterns.

ChinaRxiv 中X

Discussion

This study explores dynamic academic impact patterns using publication output and degree centrality, offering a novel perspective on highly cooperative scholars' evolution to inform talent development. Future research should incorporate additional metrics for more comprehensive pattern identification. By measuring current and future academic impacts of highly cooperative scholars, this approach better reveals talent trajectories and can incentivize growth by directing resources toward scholars with steady growth patterns.

References

[1] Abramo G, D' Angelo C A, Viel F. Peer Review Research Assessment: A Sensitivity Analysis of Performance Ranks to the Share of Research Product Evaluated[J]. Scientometrics, 2010, 85(3): 705-720.

[2] Svider P F, Choudhry Z A, Choudhry O J, et al. The Use of the H-index in Academic Otolaryngology[J]. The Laryngoscope, 2013, 123(1): 103-106.

[3] Waltman L, Van Eck N J. The Inconsistency of H-index[J]. Journal of the American Society for Information Science and Technology, 2012, 63(2): 406-415.

[4] Liu Xuan, Duan Yufeng, Zhu Qinghua, et al. Study on Scholars Evaluation Method Based on Co-author Network[J]. Journal of Intelligence, 2014, 33(12): 77-82.

[5] Lin Runhui, Fan Jianhong. A Study of Characteristics of Co-authorship Network of Chinese Scholars in Management and Their Impacts on Cooperative Performance[J]. R & D Management, 2012, 24(4): 81-92.

[6] Tang Qiang, Wang Yamin. Study High Collaboration Scholars Evaluation Value and Cooperation Network Based on the Contribution in the Field[J]. Information Studies: Theory & Application, 2015, 38(1): 85-89.

[7] Song Xiaohong. Research on Chinese Scholars' International Coauthor Network' Evolution in Management and Academic Contribution[D]. Harbin: Harbin Institute of Technology, 2011.

[8] Yang Siluo, Qiu Junping, Ding Jingda, et al. Differences of Citing Behavior over Time and Across Fields in China: A Diachronous Analysis[J]. Journal of Library Science in China, 2016, 42(2): 18-31.

[9] Thomson Reuters. Thomson Reuters China Citation Laureates 2014[EB/OL]. [2016-12-09]. http://www.thomsonscientific.com.cn/chinacitationlaureates/2014/.

[10] Pang Hongshen, Fang Shu, Yang Bo, et al. Research on the Close Degree of Cooperation in Research Team——An Empirical Study in WISE Laboratory in Dalian University of Technology[J]. Library and Information Service, 2011, 55(4): 28-32, 99.

[11] Gao Yang. Research on the Virtual Team Knowledge Sharing Promotion Strategy Base on the Social Networks[D]. Qinhuangdao: Yanshan University, 2012.

[12] Liu Xuan, Zhu Qinghua, Duan Yufeng, et al. An Empirical Study on the Application of Social Network Analysis to the Discovery and Evaluation of Scientific Research Team[J]. Journal of Information Resources Management, 2011, 1(3): 32-37, 52.

[13] Shang Shan, Meng Qi. Aggregation of Expert Team Based on Cohesive Subgroup[J]. Library Theory and Practice, 2014(8): 12-16.

[14] Yan E J, Ding Y. Applying Centrality Measures to Impact Analysis: A Coauthorship Network Analysis[J]. Journal of the American Society for Information Science and Technology, 2009, 60(10): 2107-2118.

[15] Abbasi A, Altmann J, Hossain L. Identifying the Effects of Co-authorship Networks on the Performance of Scholars: A Correlation and Regression Analysis of Performance Measures and Social Network Analysis Measures[J]. Journal of Informetrics, 2011, 5(4): 594-607.

[16] Abbasi A, Chung K S K, Hossain L. Egocentric Analysis of Co-authorship Network Structure, Position and Performance[J]. Information Processing & Management, 2012, 48(4): 671-679.

[17] Wang Feifei. The Influence of the Core Authors in the Field of Metrology from the Perspective of the Integration of Document and Citation[J]. Science of Science and Management of S. & T., 2014(12): 45-55.

[18] Zhu Tian, Wu Bin, Wang Bai, et al. Core Author Discovery in Science Collaboration Network[J]. Digital Library Forum, 2010(8): 29-35.

[19] Gao Zhi, Zhang Zhiqiang. A Review of Dynamic Quantitative Evaluation Methods for Individual Academic Impact[J]. Journal of Intelligence, 2015, 34(11): 40-43, 78.

[20] Gupta H M, Campanha J R, Pesce R A G. Power-Law Distributions for the Citation Index of Scientific Publications and Scientists[J]. Brazilian Journal of Physics, 2005, 35(4a): 981-986.

[21] Wang Meiying, Liu Xueli. Research Advances on h Index and Its Expanded Variants[J]. Chinese Journal of Scientific and Technical Periodicals, 2011, 22(2): 184-189.

[22] Department of Chemistry, Zhejiang University. 2015 Basic Requirements for Position of Chemistry Department Professor[EB/OL]. [2017-02-18]. http://www.chem.zju.edu.cn/chinese/redir.php?catalog_{id}=473&page=0.

[23] Zuo Meiyun, Wen Xiaowei, Hua Xiaoqing. Analysis on Co-authorship Network of Scholars in Knowledge Management[J]. Journal of Information Resources Management, 2012, 2(4): 4-15.

[24] Zhang Shanshan. An Exploration Study of Contribution Factors and Weight Algorithm of Author Contribution Statements in China and Abroad[J]. Library and Information Service, 2016, 60(1): 125-134.

Author Contributions

Fan Ruxia: Conceptualized the study, collected and analyzed data, wrote and finalized the manuscript.
Zeng Jianxun: Refined the research framework and methodology, revised and improved the manuscript content.
Gao Yaruixi: Developed algorithm programs to identify highly cooperative scholars across different teams.

Conflict of Interest Statement

All authors declare no conflict of interest.

Supporting Data

Supporting data are self-archived by the authors and available upon request at: fanrx2015@istic.ac.cn.
[1] Fan Ruxia. author.xlsx. Paper author statistics table.
[2] Fan Ruxia. team.xlsx. Team pattern statistics table.

Received: December 29, 2016
Revised: March 16, 2017

Note: Figure translations are in progress. See original paper for figures.