ChinaRxiv

AI translation · View original & related papers at
chinarxiv.org/items/chinaxiv-201711.01942

Postprint: A Scientific Collaboration Recommendation Method Combining Link Prediction and ET Machine Learning

Authors: Lü Weimin, Wang Xiaomei, Han Tao

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

Abstract

Purpose: To propose a novel method for recommending future scientific collaborations by integrating link prediction with machine learning, thereby improving recommendation accuracy over approaches based solely on link prediction.

Method: We construct a weighted author collaboration network, employ various link prediction indices as feature inputs, utilize the Extremely Randomized Trees (ET) machine learning algorithm to train a classifier, and apply a traversal algorithm to obtain the optimal weight combination for classification results, selecting top-accuracy predictions as collaboration recommendations.

Results: Empirical analysis is conducted using SCI paper data from the nanotechnology field spanning 2008–2010. In city collaboration recommendation, the improved ET method outperforms existing approaches with favorable recommendation success rates; the prediction method is less susceptible to factors such as network structure and demonstrates broader applicability.

Limitations: Scientific collaboration is influenced by various factors including collaboration motivation, geography, and language; the weighted author collaboration network neither reflects multiple authors from the same city or institution within a single paper nor accounts for these aforementioned factors.

Conclusion: The improved algorithm can produce more accurate collaboration recommendation suggestions than individual prediction indices and provides a reference for extension to more granular application levels, such as universities and other institutions, and individuals.

Full Text

Recommending Scientific Research Collaborators with Link Prediction and Extremely Randomized Trees Algorithm

Lv Weimin[1,2], **Wang Xiaomei**[3], **Han Tao**[1]
[1](National Science Library, Chinese Academy of Sciences, Beijing 100190, China)
[2](University of Chinese Academy of Sciences, Beijing 100049, China)
[3](Institutes of Science and Development, Chinese Academy of Sciences, Beijing 100190, China)

Abstract

[**Objective**] This paper proposes a novel method for recommending future scientific research collaborations by combining link prediction with machine learning, aiming to improve recommendation accuracy beyond methods based solely on link prediction. [**Methods**] We constructed a weighted author collaboration network and employed the Extremely Randomized Trees (ET) machine learning algorithm for classification training, using various link prediction metrics as feature inputs. A traversal algorithm was then applied to obtain the optimal weight combination for classification results, with top-accuracy predictions selected as collaboration recommendations. [**Results**] Using SCI paper data from the nanotechnology field (2008–2010) for empirical validation, the improved ET method outperformed existing approaches in city-level collaboration recommendations with high success rates. The method was also less influenced by network structure and other factors, demonstrating broader applicability. [**Limitations**] Research collaboration is influenced by numerous factors including motivation, geography, and language. The weighted author collaboration network does not reflect multiple authors from the same city or institution within a single paper, nor does it account for these influencing factors. [**Conclusions**] The improved algorithm generates more accurate collaboration recommendations than individual prediction metrics and provides a reference for extending applications to more granular contexts such as universities, institutions, and individuals.

Keywords: Scientific Research Collaboration Network; Link Prediction; Machine Learning; Random Forest; Extremely Randomized Trees; Recommendation

---

1. Introduction

In the era of the knowledge economy, collaborative relationships facilitate the exchange, transfer, and sharing of knowledge within certain social relationships [1]. As a crucial form of scientific production, research collaboration has become a powerful driver of scientific output growth and innovation. From the perspective of research outputs, co-authorship represents one of the most explicit manifestations of research collaboration, with complex relationships among co-authors forming scientific collaboration networks. These networks evolve over time, and scholars have investigated them from various angles including network structure [2-4], evolutionary mechanisms [5], and network growth [6-7]. Kretschmer [8] focused on individual collaborative behaviors.

In recent years, link prediction methods in complex networks have attracted increasing attention, with important applications in network reconstruction, evaluation of network evolution models, and recommendation systems [9-10]. Due to their significant impact across numerous fields and superior predictive performance, these methods have gradually been introduced to library and information science. Within this domain, link prediction methods for scientific collaboration networks have achieved considerable progress [11-14]. However, the prediction accuracy of link prediction itself heavily depends on network topology, resulting in limited applicability. Ensemble learning in machine learning addresses this limitation by integrating multiple different link prediction algorithms, which not only significantly expands the applicability of link prediction methods but also further improves their recommendation accuracy.

This paper summarizes research progress on link prediction methods in scientific collaboration networks. Using nanotechnology research collaborations as an example, we explore methods for recommending future research partners based on link prediction and Extremely Randomized Trees. By combining link prediction with machine learning and comparing Random Forest and Extremely Randomized Trees algorithms, we find that the Extremely Randomized Trees algorithm achieves higher prediction accuracy. We further improve the method by using enumeration to obtain optimal combination weights for recommendation ranking, yielding more precise results. This provides an effective approach for generating accurate collaboration recommendations for both collaborators themselves and policy makers.

2. Link Prediction and Machine Learning

Link prediction in networks involves predicting the existence of edges between two nodes based on node characteristics or existing structural features. Getoor et al. [15] early proposed that the link prediction problem in networks refers to how to predict the likelihood of connection between two nodes that have not yet been linked, using known network nodes and structural information.

Link prediction has been extensively studied in complex networks. Liben-Nowell et al. [16] provided pioneering work on link prediction methodologies, and Lü Linyuan et al. [9, 17] introduced these ideas to the domestic literature, summarizing three research approaches based on network topology: node similarity-based link prediction, maximum likelihood estimation-based link prediction, and probabilistic model-based link prediction. Each approach has distinct advantages and limitations. Node similarity-based methods only involve network structural information with relatively simple similarity index calculations, but their predictive power varies across different networks, with accuracy depending on whether the similarity measure effectively captures the target network' s structural characteristics. Maximum likelihood estimation-based methods target the entire network structure, resulting in high computational complexity that makes them unsuitable for large-scale networks. Probabilistic model-based methods offer higher prediction accuracy by utilizing both network structure and node attribute information, but suffer from high computational complexity and non-universal parameters that limit their application scope.

Due to their strong performance, link prediction methods have attracted widespread attention from scientists across different fields and backgrounds. For collaboration networks constructed among knowledge creation entities, with co-authored papers as the basic manifestation, link prediction has been integrated into research collaboration recommendation methods. Related research can be broadly categorized into four approaches (see).

The first approach involves index weighting, where collaboration frequency is added as a weight to metrics, which can improve link prediction accuracy in certain contexts. The second is time-series analysis, which improves prediction metrics by incorporating temporal factors to simulate evolution processes. The third involves comparing networks at different levels, applying link prediction methods at national, institutional, and author levels, finding that prediction accuracy increases at more macroscopic levels. The fourth uses weighted networks, constructing networks from different perspectives, applying link prediction methods to each network separately, and then weighting the similarity scores before ranking.

These four categories primarily employ node similarity-based link prediction metrics. However, they either analyze single link prediction metrics directly or perform simple linear weighting of similarity scores from several metrics, which fails to achieve optimal prediction effectiveness. To date, more than 30 metrics have been used for link prediction, but single metrics consider relatively limited information and their recommendation success rates depend on the network' s topological structure, resulting in limited method applicability. Therefore, finding appropriate ways to integrate these metrics or incorporate more attributes into a single metric to achieve broader applicability and higher recommendation success rates has become a research direction.

Mitchell [22] proposed that ensemble learning in machine learning importantly combines various features, leveraging their respective strengths through diverse integrated learning systems. Ensemble learning has been attempted in research collaboration domains. Guns et al. [12] proposed combining Random Forest (RF) methods to study inter-city research collaboration in Africa, the Middle East, and South Asia, constructing weighted collaboration networks for malaria and tuberculosis research across three consecutive periods (1997–2001, 2002–2006, 2007–2011). Using the Random Forest algorithm from ensemble learning to build classifiers, they recommended top-ranked predictions as collaboration suggestions, achieving better recommendation accuracy than individual link prediction metrics.

In recent years, machine learning has been increasingly applied across academic fields, and the idea of combining it with link prediction has begun to attract attention [23]. This paper improves upon Guns et al.' s [12] RF method (hereinafter referred to as the "RF method"), with the improved method called the "improved ET method." By combining link prediction with machine learning, we aim to improve the accuracy of research collaboration recommendations and enhance practicality.

3. Data Description and Network Construction

This study conducts research collaboration recommendations in the nanotechnology field. Using search strategies constructed by Arora et al. [24], we retrieved data from the Web of Science (WOS) Core Collection (the retrieval formula includes various nanotechnology keywords with some stop words removed; omitted here due to space limitations). We selected three time periods—2008, 2009, and 2010—with all nanotechnology Article-type papers from each year forming a research collaboration network, creating three continuously evolving networks as shown in .

When constructing networks, we extracted cities to build weighted collaboration networks, with each city representing a node in the network and all cities forming the node set. Specifically, if a paper had two authors from different cities, those two cities were recorded as collaborating once. If a paper had multiple authors from different cities, those cities were connected by an edge in the network, with the edge weight incremented by 1. The weight between cities A and B equals the number of articles co-authored by authors from cities A and B. Due to significant variations in address formats and data inconsistencies, all results were manually checked and corrected as necessary.

4. Research Collaboration Recommendation Based on Link Prediction and ET

The approach involves first extracting appropriate link prediction metric features, then building machine learning models based on these features, and finally integrating different features within the ET algorithm to achieve better results.

4.1 Feature Selection for Link Prediction Metrics

We focus on cities within target countries that have not yet collaborated and are interested in whether they will form collaborations. For each link prediction metric, we determine the correlation score S between each pair of nodes based on the existing network, select pairs that are not yet connected, and rank them in descending order according to S to generate a list of cities most likely to collaborate in the future.

Among the three link prediction approaches, node similarity-based methods are computationally simple, applicable to large-scale networks, and perform well in networks with appropriate topological structures, making them widely used in knowledge network research. This paper primarily focuses on node similarity-based link prediction methods. Considering both algorithm efficiency and prediction performance [25], we selected six metrics as machine learning input features: metrics considering neighbor information including Common Neighbours (CN), Adamic/Adar [26] (AA), and Resource Allocation (RA); and metrics considering overall network topology including Katz [27], Graph Distance (GD), and SimRank. Guns' [18] research indicates that appropriately incorporating weights into metrics can improve prediction accuracy, so our experiments only include weighted versions.

Feature extraction using link prediction methods mainly involves two steps [28]: applying link prediction metrics to a training network to predict potential new links, and evaluating the metrics by comparing them with a test network.

4.2 Integrating the ET Method

Both Random Forest (RF) and ET belong to ensemble learning in machine learning. Random Forest selects the most effective feature from the feature set for classification, potentially yielding slightly better results, but multiple runs may produce unstable outcomes. Subsequently, ET was developed, which selects features completely randomly, producing results with smaller variance and greater stability. After comparing the two, this paper primarily selects ET as the training method, with specific steps as follows:

1. Divide the dataset into three periods: 2008, 2009, and 2010. Use 2008 data to construct the early network A1, 2009 data to construct the later network A2, and 2010 data as the validation set.
2. Select appropriate link prediction metrics and calculate correlation scores for each node pair in A1.
3. Use correlation scores as features. ET learns from features in A1 and corresponding classification data in A2 (whether connected), using the known existence of edges in A2 (represented by 0 or 1) as classification results to build the model. By learning the correlation strength of each prediction metric in the current time slice and whether connections occur in the next time slice, the method constructs a relatively accurate classifier.
4. For each possible connection in A2, match the correlation scores obtained from link prediction metrics in A1.
5. Use features from A2 as the test set and apply the previously constructed classifier for classification, providing predicted classification results and selecting potentially connectable node pairs for re-evaluation.
6. Assign weights to the re-selected potentially connectable node pairs from

step 5, with each weight combination corresponding to a set of recommendation accuracies.
7. Enumerate all possible weight combinations, select the weight combination corresponding to the highest accuracy recommendation result, and recommend the top n predicted results as collaboration pairs.

We used the Python package LinkPred provided by Guns to calculate link prediction metrics and Scikit-learn [29] for machine learning training. The number of trees in the RF/ET algorithm forest was set to 400, and recommendation accuracies represent averages of 10 computational runs.

4.3 Evaluation Metrics

Commonly used evaluation metrics for link prediction include AUC, Precision, and Ranking Score, each measuring prediction accuracy from different perspectives. AUC measures overall algorithm accuracy [30], Precision only considers whether edges in the top L positions are predicted correctly [31], and Ranking Score focuses more on the ranking of predicted edges [32].

In practical applications, decision-makers and researchers are interested in collaborations with high potential and typically only focus on the top few most likely collaboration groups rather than those ranked beyond several dozen. Therefore, this paper adopts Precision to evaluate recommendation results, selecting the top 10 predictions for recommendation. Using networks from 2008 and 2009 as the training set to generate recommendations for the post-2009 collaboration network, we define s as the number of successful recommendations and n as the total number of recommendations. Recommendation accuracy is defined as s/n, serving as a measure of recommendation quality.

5. Results and Analysis

5.1 Accuracy Comparison of Individual Metrics, RF, and ET

Predicting collaborations between research institutes in different cities based on correlation scores, we applied six weighted metrics—AA, CN, GD, Katz, RA, and SimRank—individually. The prediction accuracy results are shown in and [Figure 1: see original paper]. For the weighted versions, parameters were: Weighted Katz: α=1; Weighted Graph Distance: parameter not specified; Weighted SimRank: C not specified.

Using the same approach but employing the ET algorithm for training, the resulting recommendation accuracies are also shown in . Comparing the six link prediction metrics, the RF algorithm, and ET reveals that the six link prediction metrics maintain prediction effectiveness around 80%, except for SimRank, which performs poorly. Typically, prediction accuracy decreases as the number of recommendations increases, though some metrics exhibit different patterns. Surprisingly, the existing integrated RF method achieves only about 60% accuracy, which is lower than some individual link prediction metrics. The integrated

ET method improves recommendation accuracy to some extent, but overall performance remains inferior to some individual metrics. Therefore, we improved Guns' algorithm by adopting the ET algorithm for model construction and implementing a traversal algorithm to enumerate all possible weight combinations, selecting the combination with the best predictive performance as the optimal weight combination.

5.2 Obtaining Optimal Weights

In the further improved method, we do not directly use Gini Importance as weights. Instead, we use a traversal algorithm to enumerate all possible weight combinations, with each combination corresponding to a set of recommendation accuracy results. To prevent excessive randomness in ET prediction results, accuracies represent averages of 10 computational runs. We selected the top 5 weight combinations with the best accuracy, as shown in .

5.3 Analysis of Results

The improved ET method shows decreasing prediction accuracy as the number of recommendations increases, but among the top 10 recommendations, accuracy equals or approaches 100%, outperforming both individual link prediction metrics and RF and ET methods. The reasons are as follows:

1. Intuitively, if five metrics all suggest that two cities may collaborate, the integrated learning result is more likely to predict collaboration. If only two metrics predict potential collaboration while the other three predict no collaboration, the integrated learning result may not strongly favor collaboration between these two cities. Therefore, integrated learning achieves higher recommendation accuracy than individual metrics.
2. The RF method integrates six link prediction metrics with varying individual accuracies, and less accurate metrics affect final prediction results. The improved ET method extracts five metrics with prediction effectiveness above 70%, representing a better feature selection process. As shown in , SimRank performs poorly on our dataset, so we excluded it from the ET integration algorithm, using only the top five metrics.
3. The RF method' s prediction accuracy is not optimal because Gini Importance does not always accurately reflect weights in some cases—important metrics do not necessarily have larger weights—and RF does not perform well on our dataset. The improved ET method, by traversing all possible weight combinations, can achieve optimal accuracy. Moreover, the improved ET algorithm integrates different link prediction metrics with distinct features, enabling effective training across networks with different topological structures and broader applicability.

The improved ET algorithm demonstrates higher stability, with minimal variance for each metric across the five effective weight combinations. For example, GD' s weight interval across the five combinations is [0.85,1.0], and RA' s weight interval is [0.0,0.1]. Therefore, we can confidently adopt the weight combination ([AA, CN, GD, Katz, RA]:[0.0, 0.0, 1.0, 0.0, 0.0]) as optimal.

In the improved ET algorithm, link prediction metrics with higher individual prediction accuracy receive relatively higher corresponding weights, indicating their greater role in predictions. In practical applications, one can first select numerous link prediction metrics for prediction, then extract the better-performing metrics as feature inputs for the ET algorithm to ensure better machine learning effectiveness.

Compared with traditional library and information science approaches that rely solely on citation frequency and limited factors for network construction and link prediction, combining link prediction methods with ET and referencing more factors to predict future collaborations achieves better method applicability and accuracy. Additionally, this machine learning approach offers a prominent advantage: it can effectively reduce time complexity and improve efficiency when dealing with large-scale collaboration network problems.

6. Conclusion

This paper introduces an improved research collaboration recommendation method based on link prediction and ET, proposing the ET algorithm with better predictive performance. By adding a step to enumerate all possible weights to solve for the optimal weight combination and using this optimal combination for ranking, recommendation accuracy is greatly improved, with success rates exceeding those of all individual metrics. The ET algorithm integrates different link prediction metrics with distinct features, enabling effective training across networks with different topological structures, making the results more stable and broadly applicable than individual link prediction methods. This represents an excellent approach for recommending research partners. Compared with machine learning methods such as support vector machines, it offers better time complexity, providing prominent advantages when processing large-scale research collaboration networks and enabling application at more granular levels for research collaboration recommendations.

When constructing our dataset, if multiple authors from the same city contributed to one paper, they were counted as a single author. For example, a paper with five authors from city A and three from city B was treated as having one author from each city. Future research will attempt to consider multiple authors from the same city separately from a heterogeneous network perspective.

At the city level, cities worldwide have distinct research characteristics, and finding research partners can maximize one' s research advantages. Collaboration between less research-intensive cities and more developed ones is also attractive [34], such as establishing local elite centers and gaining more comprehensive understanding of shared needs and problems in developing countries [35]. Additionally, our method can be extended to more granular contexts such as research institutions and individual researchers, providing more practical value. However, networks constructed at micro-levels become excessively large, making link prediction analysis inefficient with existing tools. If existing link prediction methods or datasets can be improved to obtain prediction results from various metrics, our method can be applied for collaboration recommendation, which is also a research objective of this paper.

Overall, the application of link prediction in research collaboration recommendation remains in an exploratory and empirical research stage. Different prediction metrics and methods have their own advantages, disadvantages, and application scopes. Therefore, finding appropriate ways to integrate different prediction results or incorporate as much information as possible (such as network node attributes and network topology) into metrics represents a future research direction. Meanwhile, leveraging the theoretical framework and evaluation methods of link prediction can provide recommendations for interdisciplinary collaboration trends and research groups, thereby solving the challenge of finding collaborators [36]. Further conclusions require more experimental support, which constitutes future work.

References

[1] Zhang Bin, Ma Feicheng. A Review on Link Prediction of Scientific Knowledge Network[J]. Journal of Library Science in China, 2015, 41(3): 99-113.

[2] Newman M E J. Scientific Collaboration Networks. I. Network Construction and Fundamental Results[J]. Physical Review E, 2001, 64(1): 016131.

[3] Newman M E J. Scientific Collaboration Networks. II. Shortest Paths, Weighted Networks, and Centrality[J]. Physical Review E, 2001, 64(1): 016132.

[4] Newman M E J. The Structure of Scientific Collaboration Networks[J]. Proceedings of the National Academy of Sciences, 2001, 98(2): 404-409.

[5] Barabási A L, Jeong H, Néda Z, et al. Evolution of the Social Network of Scientific Collaborations[J]. Physica A: Statistical Mechanics and Its Applications, 2002, 311(3-4): 590-614.

[6] De Solla Price D J. Little Science, Big Science ···and Beyond[M]. New York: Columbia University Press, 1986.

[7] Zuckerman H A. Patterns of Name Ordering Among Authors of Scientific Papers: A Study of Social Symbolism and Its Ambiguity[J]. American Journal of Sociology, 1968, 74(3): 276-291.

[8] Kretschmer H. Author Productivity and Geodesic Distance in Bibliographic Co-authorship Networks, and Visibility on the Web[J]. Scientometrics, 2004, 60(3): 409-420.

[9] Lü L, Zhou T. Link Prediction in Complex Networks: A Survey[J]. Physica A: Statistical Mechanics and Its Applications, 2011, 390(6): 1150-1170.

[10] Zhu B, Xia Y. An Information-theoretic Model for Link Prediction in Complex Networks[J]. Scientific Reports, 2015, 5: Article No. 13707.

[11] Guns R, Rousseau R. Predicting and Recommending Potential Research Collaborations[C]//Proceedings of ISSI. 2013: 193-203.

[12] Guns R, Rousseau R. Recommending Research Collaborations Using Link Prediction and Random Forest Classifiers[J]. Scientometrics, 2014, 101(2): 1461-1473.

[13] Yan E, Guns R. Predicting and Recommending Collaborations: An Author-, Institution-, and Country-level Analysis[J]. Journal of Informetrics, 2014, 8(2): 295-309.

[14] Zhang Bin, Li Yating. A Review of the Evolution Model of Scientific Knowledge Network[J]. Journal of Library Science in China, 2016, 42(5): 85-101.

[15] Getoor L, Diehl C P. Link Mining: A Survey[J]. ACM SIGKDD Explorations Newsletter, 2005, 7(2): 3-12.

[16] Liben-Nowell D, Kleinberg J. The Link Prediction Problem for Social Networks[J]. Journal of the Association for Information Science and Technology, 2007, 58(7): 1019-1031.

[17] Lv Linyuan. Link Prediction on Complex Networks[J]. Journal of University of Electronic Science and Technology of China, 2010, 39(5): 651-661.

[18] Guns R. Missing Links: Predicting Interactions Based on a Multi-relational Network Structure with Applications in Informetrics[A]. // Missing Links: Predicting Interactions Based on a Multi-relational Network Structure with Applications in Informetrics[M]. Universiteit Antwerpen (Belgium). 2012.

[19] Tylenda T, Angelova R, Bedathur S. Towards Time-aware Link Prediction in Evolving Social Networks[C]//Proceedings of the 3rd Workshop on Social Network Mining and Analysis. ACM, 2009: 1-10.

[20] Wang C, Satuluri V, Parthasarathy S. Local Probabilistic Models for Link Prediction[C]//Proceedings of the 7th IEEE International Conference on Data Mining. IEEE, 2007: 322-331.

[21] Guns R. Generalizing Link Prediction: Collaboration at the University of Antwerp as a Case Study[J]. Proceedings of the American Society for Information Science and Technology, 2009, 46(1): 1-15.

[22] Mitchell T M. Machine Learning. 1997[J]. Burr Ridge, IL: McGraw Hill, 1997, 45(37): 870-877.

[23] Backstrom L, Leskovec J. Supervised Random Walks: Predicting and Recommending Links in Social Networks[C]//Proceedings of the 4th ACM International Conference on Web Search and Data Mining. ACM, 2011: 635-644.

[24] Arora S K, Porter A L, Youtie J, et al. Capturing New Developments in an Emerging Technology: An Updated Search Strategy for Identifying Nanotechnology Research Outputs[J]. Scientometrics, 2013, 95(1): 351-370.

[25] Guns R. Bipartite Networks for Link Prediction: Can They Improve Prediction Performance[C]//Proceedings of ISSI. 2011: 249-260.

[26] Adamic L A, Adar E. Friends and Neighbors on the Web[J]. Social Networks, 2003, 25(3): 211-230.

[27] Katz L. A New Status Index Derived from Sociometric Analysis[J]. Psychometrika, 1953, 18(1): 39-43.

[28] Guns R. Link Prediction[A]//Measuring Scholarly Impact[M]. Springer International Publishing, 2014: 35-55.

[29] Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: Machine Learning in Python[J]. Journal of Machine Learning Research, 2013, 12(10): 2825-2830.

[30] Hanley J A, McNeil B J. The Meaning and Use of the Area Under a Receiver Operating Characteristic (ROC) Curve[J]. Radiology, 1982, 143(1): 29-36.

[31] Herlocker J L, Konstan J A, Terveen L G, et al. Evaluating Collaborative Filtering Recommender Systems[J]. ACM Transactions on Information Systems (TOIS), 2004, 22(1): 5-53.

[32] Zhou T, Ren J, Medo M, et al. Bipartite Network Projection and Personal Recommendation[J]. Physical Review E, 2007, 76(2): 046115.

[33] Breiman L, Friedman J, Stone C J, et al. Classification and Regression Trees[M]. CRC Press, 1984.

[34] Schubert T, Sooryamoorthy R. Can the Centre–periphery Model Explain Patterns of International Scientific Collaboration Among Threshold Industrialised Countries? The Case of South Africa and Germany[J]. Scientometrics, 2010, 83(1): 181-203.

[35] Boshoff N. South–South Research Collaboration of Countries in the Southern African Development Community (SADC)[J]. Scientometrics, 2010, 84(2): 481-503.

[36] Pavlov M, Ichise R. Finding Experts by Link Prediction in Co-authorship Networks[C]//Proceedings of the International Conference on Finding Experts on the Web with Semantics. 2007.

Author Contributions

Lv Weimin: Conceived research ideas, designed research methodology, conducted experiments, analyzed data, and wrote the manuscript.
Wang Xiaomei: Revised the final manuscript version.
Han Tao: Performed data analysis and methodology analysis.

Conflict of Interest Statement

All authors declare no conflict of interest.

Supporting Data

Supporting data is self-archived by the authors, E-mail: lvweimin@mail.las.ac.cn.

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

Source: ChinaXiv —Machine translation. Verify with original.