



Anomalous citations detection in academic networks

Jiaying Liu¹ · Xiaomei Bai² · Mengying Wang³ · Suppawong Tuarob⁴ · Feng Xia⁵

Accepted: 19 December 2023 / Published online: 29 March 2024
© The Author(s) 2024

Abstract

Citation network analysis attracts increasing attention from disciplines of complex network analysis and science of science. One big challenge in this regard is that there are unreasonable citations in citation networks, i.e., cited papers are not relevant to the citing paper. Existing research on citation analysis has primarily concentrated on the contents and ignored the complex relations between academic entities. In this paper, we propose a novel research topic, that is, how to detect anomalous citations. To be specific, we first define anomalous citations and propose a unified framework, named ACTION, to detect anomalous citations in a heterogeneous academic network. ACTION is established based on non-negative matrix factorization and network representation learning, which considers not only the relevance of citation contents but also the relationships among academic entities including journals, papers, and authors. To evaluate the performance of ACTION, we construct three anomalous citation datasets. Experimental results demonstrate the effectiveness of the proposed method. Detecting anomalous citations carry profound significance for academic fairness.

Keywords Anomalous citation · Scholarly big data · Network representation · Non-negative matrix factorization

1 Introduction

Citations could be regarded as currencies for assessing the scholarly impact for both scholars and institutions in academia. At the same time, citation analysis can help scholars understand the development trend of specific disciplines and the frontier of science (Liu et al. 2021a; Fortunato et al. 2018). The development of information technology and the

✉ Xiaomei Bai
xiaomeibai@outlook.com

¹ School of Economics and Management, Dalian University of Technology, Dalian 116024, China

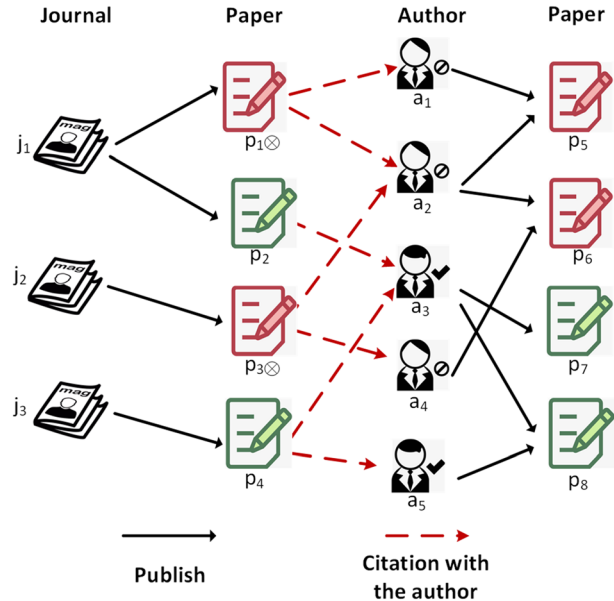
² Computing Center, Anshan Normal University, Anshan 114007, China

³ Department of Computer and Information Science, University of Konstanz, Konstanz 78464, Germany

⁴ Faculty of Information and Communication Technology, Mahidol University, Salaya 73170, Thailand

⁵ School of Computing Technologies, RMIT University, Melbourne VIC 3000, Australia

Fig. 1 The example of a heterogeneous academic network which contains three types of academic entities, i.e., publications, journal, and author



expansion of large databases provide the opportunity to achieve the multi-angle and systematic analysis of academic networks (Xia et al. 2017). For citation bibliographic information analysis, researchers have proposed a large number of citation-based evaluation indicators such as h -index and g -index. In these indicators, citations are considered equal and valid. However, different types of citations should be treated differently (Zhu et al. 2015).

Although scholars have distinguished different types of citations in citation network analysis, most research has focused on the bibliographic information (Liu and Fang 2020; Cai et al. 2019; Siudem et al. 2020), while few of them concentrate on the citation contents. At the same time, the number of suspected citations increases,¹ which are established to enhance the impact of publications or authors intentionally rather than disseminate priors scientific advances contributing to the publication.² In this paper, we regard the above-mentioned citations as anomalous citations. Anomalous citations are academic misconducts that are currently highly concerned by the academic community (Franck 1999; Bai et al. 2016; Chorus and Waltman 2016; Mimouni et al. 2016; Bai et al. 2017, 2020; Liu et al. 2022). It is essential to detect anomalous citations because they will bring a lot of negative effects. For example, they will affect researchers' judgment on the real quality and value of papers. Moreover, they could result in academic false prosperity and even undermine academic fairness. Anomalous citations also affect citation-based measurement indices, which are associated with scholars' promotion and awards.

In this work, we aim to give an effective solution that addresses the problem of anomalous citations identification in the heterogeneous academic network. To this end, we propose ACTION (Anomalous CItations detECTION) to simultaneously exploit complex

¹ <http://help.prod-incites.com/incitesLive/JCR/JCRGroup/titleSuppressions>.

² <https://scholarlykitchen.sspnet.org/2012/04/10/emergence-of-a-citation-cartel/>.

relations among multiple academic entities. Figure 1 gives the illustration of a heterogeneous academic network and the relations among the academic entities including journals, authors, and publications. In Fig. 1, papers p_1 and p_2 are both published on journal j_1 . p_3 and p_4 are published on j_2 and j_3 , respectively. a_1, a_2, \dots, a_5 are authors who cite these papers. In addition, there may be collaboration relationships existing between these authors. For example, a_1 and a_2 have published p_5 together. a_2 and a_4 have published p_6 together. In addition, we use different colours and symbols to distinguish whether a paper or author has anomalous citation behaviour in Fig. 1. For example, papers p_1, p_3, p_5 , and p_6 have anomalous references. Authors a_1 and a_2 have published p_5 , so they have anomalous citation behaviour.

More specifically, in the proposed framework, the first step is to represent the relationships in the academic heterogeneous network by constructing relational matrices. Then ACTION models paper contents, journal-paper relationship, and author-paper relationship based on the constructed matrices. Finally, the proposed framework integrates these three parts and combines the semi-supervised method to identify anomalous citations. Furthermore, we also identify the papers with anomalous citations and observe the effectiveness of ACTION when anomalous boundary changes. The key contributions of our work include:

- We formally define a novel topic, i.e., anomalous citations detection, which widely covers a series of real-world issues.
- We propose a novel framework, namely ACTION to simultaneously model relationships among journals, papers and related authors in the heterogeneous academic network for anomalous citation detection. Three independent modules modeling the different relations of the afore-mentioned entities are designed. The proposed framework has achieved better performance compared with baselines.
- We construct a database of anomalous citations, in combination with three academic datasets, i.e., Microsoft Academic Graph (MAG), Digital Bibliography & Library Project (DBLP) and CiteseerX.

The rest of the paper is organized as follows. Section 2 gives a brief introduction to the related work. In Sect. 3, we explain the definition of the problem and the mathematical preliminaries. In Sect. 4, we describe each part of the framework in detail. Section 5 depicts the process of the experiments and presents the results of the experiments. Finally, we conclude our work and give the future directions in Sect. 6.

2 Related work

Previous work on anomalous citation detection is limited. The most related topics focus mainly on citation network analysis. The proposed framework takes advantage of non-negative matrix factorization. In the contexts of citation network analysis and non-negative matrix factorization, we provide a brief review of the related work as follows.

2.1 Citation network analysis

Citation network analysis is a well-established research topic which utilizes multiple kinds of techniques: bibliometrics, machine learning algorithms, and complex network analysis (McLaren and Bruner 2022; Liu et al. 2021b; Xia et al. 2020; Liu et al. 2020). By analyzing the citation relationships between academic entities including journals, papers, and authors, researchers could reveal the quantitative characteristics and inherent laws within science of science (Jiang and Liu 2023; Singh et al. 2020).

Citation network analysis consists of three types from the perspective of contents, including number of citations, structure of the citation network, and relevance of topics reflected in citation relationships. Jiang and Liu (2023) generate a precise and comprehensive picture of domain evolution and uncover more coherent development pathways based on the semantic main path network analysis approach. Liu et al. (2018) trace the scientific publications the scientists produced and quantitatively describe the hot streak phenomenon in their careers. Siudem et al. (2020) propose a model to recreate citation records from three perspectives, i.e., the number of publications, citations, and the degree of randomness in the citation patterns. In the field of medicine, Liao et al. (2018) explore the current status of medicine through visualizing and analyzing the citation network constructed by publications related to medical big data.

The recent studies provide new insight into co-citation network analysis. Co-citation networks are constructed based on publication co-citation relationships, which means that two publications are cited in the same article. It has been proved that co-citation analysis is a useful way to help researchers identify key literature for cross-disciplinary ideas (Trujillo and Long 2018). Shiau et al. (2023) explore the research and identify eight core knowledge groups based on the citation and co-citation analysis in the field of information security. Kim et al. (2016) consider both citation contents and proximity to represent the authors' subject relatedness. Similarly, taking advantage of author co-citation analysis, Bu et al. (2018) propose a new approach incorporating new pieces of information, i.e., the number of mentioned times and the number of context words to depict scientific intellectual structures. By using the data from CiteSpace, Fang et al. (2018) tend to identify the intellectual landscape of climate change and tourism based on analyzing and visualizing the collaboration network and the co-citation network of the related areas.

Research on anomalous citation is limited. Anomalous citations can be regarded as a kind of improper behavior (Yu et al. 2020), including miscitations, missing citations, mandatory citations, and inappropriate self-citations. The citation behavior is arbitrary because the authors have the right to decide how to cite references. The research pointed out that less than 1/3 of the references in each article must be cited (Iqbal et al. 2021), which also confirms the randomness of citing behavior in the process of writing the paper. The arbitrariness also brings opportunities for academic misconduct. Moustafa (2016) explains different kinds of aberration of the citation including self-citations (to inflate the authors' h-index), discriminatory citations (citations only from specific journals resulting in a substantial increase in their impact factors), reciprocal citations (citations from people who cite their own work), etc. Elsevier has investigated more than 55,000 paper reviewers in more than 1,000 journals, and the evidence showed that 433 of them might have compulsory citations (Chawla 2019). The journal citation stacking is regarded as the behavior pattern of anomalous citations. Some studies focus on

analyzing journal self-citations and assuming that journal self-citation malpractices are aimed at inflating a journal's impact factor (Chorus and Waltman 2016; Humphrey et al. 2019; Gazni and Didegah 2021; Kojaku et al. 2021; Krell 2014). Besides, inappropriate co-author and collaborative self-citations may be misleading and may distort the scientific literature (Vercelli et al. 2023). Although scholars have put forward the concept and tried to divide anomalous citations, there is still room for further deepening and expansion. Due to the difficulty of obtaining annotation data for anomalous citations, the existing research is mainly based on a small amount of data, combined with specific cases to classify different anomalous citations behaviours. There is little research on anomalous citations identification methods in academic networks. The citation behaviour is complex and diverse in reality, and the identification of anomalous citations should not only consider the content similarity between the cited paper and the citing paper, but also the implicit conflict of interest relationship between the papers and the authors. Scientometric methods based on statistical analysis are not capable of representing and processing complex networks and text contents, so it is difficult to fully explore the patterns behind the data.

2.2 Non-negative matrix factorization

As one of the popular machine learning techniques (Xia et al. 2021a; Sun et al. 2020), Non-negative Matrix Factorization (NMF) is proposed by Lee and Seung (Lee and Seung 1999). NMF aims to make all components after decomposition non-negative, at the same time realizes dimension reduction. Mathematically, for any non-negative matrix \mathbf{V} , it aims to find a non-negative matrix \mathbf{W} and a non-negative matrix \mathbf{H} , which satisfies the condition $\mathbf{V} = \mathbf{W} * \mathbf{H}$. Thus it decomposes a non-negative matrix into the product of two non-negative matrices, which can realize the dimension reduction and feature extraction of the original matrix. Taking advantage of its convenient and reasonable explanation of the data, NMF has gradually become one of the most popular multidimensional data processing tools in many fields (Lin and Boutros 2020), such as signal processing (Puigt et al. 2021), biological processes (Hamamoto et al. 2022; Wang et al. 2021), and computer vision (Dai et al. 2020; Ji et al. 2022).

A class of methods based on NMF is graph-based matrix factorization that considers a priori information on a local structure of data. Since the matrix to be decomposed can be an asymmetric matrix and the decomposition results are all non-negative matrices, this type of method has good adaptability and interpretability, and has received extensive attention in recent years. For example, Wang et al. (2011) propose the asymmetric non-negative matrix factorization (Asymmetric NMF, ANMF), which decomposes the asymmetric adjacency matrix of the directed graph into non-negative matrix, and then divides the nodes into clusters based on the membership matrix. On this basis, Tosalı et al. (2019) propose Regularized Asymmetric NMF (RANMF) and improve the accuracy and robustness of directed graph clustering. One of the most successful applications of NMF is in the field of computer vision. Generally, the computer stores the image information in the form of a matrix. Thus researchers can conduct recognition, analysis, and processing based on the matrix, which makes NMF well applied in computer vision. Zhao et al. (2022) propose a novel progressive deep non-negative matrix factorization (PDNMF) for image reorganization. They add a basis image reconstruction step to the successive basis image factorization steps to improve accuracy. Li et al. (2016) propose NLMF method for obtaining an effective low-rank data representation and apply it to image clustering.

In recent years, there have been increasing interests in social network mining and analysis with the application of NMF (Ren et al. 2021; Xia et al. 2021b). Shu et al. (2019) utilize relationships among publishers, news content, and users who spread the news on social networks to identify fake news. Li et al. (2020) propose a multi-view clustering method based on deep graph regularized non-negative matrix factorization. Luo et al. (2020) propose a new framework named PGS for community extraction taking advantage of four different non-negative matrix factorization models. Inspired by these studies, we aim to detect anomalous citations based on NMF.

3 Preliminaries and definitions

We focus on mining anomalous citations hidden in the academic network which consists of complex relations among journals, authors, and publications. In this section, we first define anomalous citations in our work. Then we describe the mathematical preliminaries used throughout the paper.

3.1 Definition of anomalous citation

Although studies have proved the existence of anomalous citations by analyzing the citation network and stating on publication ethics, there is not a clear definition of anomalous citations. In this paper, in order to study the anomalous citation identification problem, we have carefully searched and read research on anomalous citations. The specific forms of anomalous citations include:

Citation stacking refers to anomalous citation behavior in which two or more journals exchange citations with each other. To improve the impact factor of journals, some editors or reviewers encourage the authors to cite papers published in journals or other journals related to their own interests, which leads to journal citation stacking (Chorus and Waltman 2016; Krell 2014; Mimouni et al. 2016). Under this circumstance, there are usually “donor” journals (journals that initiate the citations) and “recipient” journals (journals that are cited). Through giving and receiving, the impact factor can be increased without increasing the self-citation rate. In 1999, a paper published in Science proposed the concept of “citation cartel” for the first time (Franck 1999), defining this phenomenon as a mutually beneficial behavior in which editors and journal groups frequently exchange citations. The phenomenon has become more prevalent in recent years as the scholarly publishing market has increased competitively. Clarivate Analytics has announced and suspended journals whose impact factors were boosted and distorted by different anomalous citation patterns. Scholarly Kitchen also reported some cases of citation cartel. At present, scholars have extended the citation cartel to other relationships, such as editor-author or author-author (Fister Jr et al. 2016).

Relational citations include conflicts of interest between the citing authors and the cited authors, the most common of which is self-citation. In fact, proper self-citation is necessary. However, some journals or scholars increase their influence and impact factor by manipulating self-citations (Szomszor et al. 2020). Some scholars have analyzed the

citations of Nobel Prize winners from 2002 to 2007 and found that the self-citation rate of these authors is lower than the average level. At the same time, the self-citation rate of influential papers is relatively low. Research has found that papers with the citation in the top 10% have a lower self-citation rate, which is only 2%. Normal self-citation can provide support for the paper, but artificial self-citation will damage the fairness of academic evaluation.

Journal self-citation behavior is closely related to its impact factor. Through the analysis, Campanario (2011) has found that the rapid growth and decay of journal impact factor within a year depends on the amount of journal self-citation. Self-citations make up a significant portion of citations in some journals. In this case, self-citation can distort the true impact of a journal in a given field. Some authors may also prefer to cite papers from team members, colleagues, supervisors, students, or other collaborators. The proportion of such citations is as high as 25% in research papers, and 15% in review papers (Greenland and Fontanarosa 2012). This kind of citation behavior also needs to be paid attention to. Some scholars also regard citations from collaborators and the same institution as conflicts of interest (Cai et al. 2019). Reciprocal citation can also be regarded as a kind of relational citation. Some scholars have studied the papers published in Science and found that the authors tend to cite the papers published by scholars who have cited their own articles (Corbyn 2010).

Normal relational citations are excusable. However, some scholars use personal and interpersonal resources to make some papers that have no (or little) academic relevance to be cited, thereby expanding the influence of academic research (Biagioli et al. 2019).

We acknowledge that not all relational citations are anomalous citations, so textual semantic information should be considered in the identification of anomalous citations. In summary, we believe that anomalous citations have the following characteristics:

- The content of the citing paper is irrelevant to references.
- The citations between the cited papers and citing papers are relational citations.

In this paper, we use the abstract of the paper to measure the similarity between the cited paper and the citing paper. Besides, if the citing paper and the cited paper are related (i.e., self-citations, discriminatory citations, reciprocal citations, co-author citations), it belongs to relational citations. In this paper, we model the relationship between the authors, papers, and journals to help determine whether the citations are relational citations.

3.2 Mathematical preliminaries

Next, we explain the meaning of the matrices and notations that appear in this paper. In the heterogenous academic network G , $\mathcal{P} = \{p_1, p_2, \dots, p_n\}$ is the set of papers and $\mathcal{A} = \{a_1, a_2, \dots, a_m\}$ is the set of authors, where m represents the number of authors and n is the number of publications. $\mathcal{J} = \{j_1, j_2, \dots, j_l\}$ is the set of journals, where l is the number of journals. We define $\mathbf{X} \in \mathbb{R}^{n \times t}$ as the paper feature matrix, where n is the number of publications and t is the feature dimension (the dimension of vectors into which abstracts are converted). For the collaboration relationships between authors, we use $\mathbf{A} \in \mathbb{R}^{m \times m}$ to denote the author collaboration times adjacency matrix. The author-paper citing matrix in the citation network is denoted as $\mathbf{C} \in \{0, 1\}^{m \times n}$, where $\mathbf{C}_{ij} = 1$ indicates that the author a_i has cited the paper p_j ; otherwise $\mathbf{C}_{ij} = 0$. We define $\mathbf{cr} \in \mathbb{R}^{m \times 1}$ as the author credibility. $\mathbf{B} \in \mathbb{R}^{l \times n}$ is the journal-paper relation matrix, where $\mathbf{B}_{kj} = 1$ means that the paper p_j is

Table 1 The description of notations

Notation	Description
G	The heterogenous academic network
\mathcal{P}	The set of papers
\mathcal{A}	The set of authors
\mathcal{J}	The set of journals
$\mathbf{X} \in \mathbb{R}^{n \times l}$	The paper feature matrix
$\mathbf{A} \in \mathbb{R}^{m \times m}$	The author collaboration times matrix
$\mathbf{C} \in \mathbb{R}^{m \times n}$	The author-paper citing matrix
$\mathbf{cr} \in \mathbb{R}^{m \times 1}$	The author credibility
$\mathbf{B} \in \mathbb{R}^{l \times n}$	The journal-paper relation matrix
$\mathbf{j} \in \mathbb{R}^{l \times 1}$	The journal grade vector
$\mathbf{Y} \in \mathbb{R}^{n \times n}$	The label matrix
$\mathbf{N} \in \mathbb{R}^{n \times d}$	The low-dimensional representations of papers
$\mathbf{N}_L \in \mathbb{R}^{r \times d}$	The papers' latent feature matrix for labeled cited papers
$\mathbf{N}_U \in \mathbb{R}^{(n-r) \times d}$	The papers' latent feature matrix for unlabeled cited papers
$\mathbf{K} \in \mathbb{R}^{l \times d}$	The low-dimensional representations of words
$\mathbf{D} \in \mathbb{R}^{m \times d}$	The authors' latent matrix
$\mathbf{U} \in \mathbb{R}^{d \times d}$	The author association matrix
n	The number of publications
t	The feature dimension of the paper abstract
m	The number of authors
l	The number of journals
d	The dimension of the latent topic space
r	The number of labeled citation papers

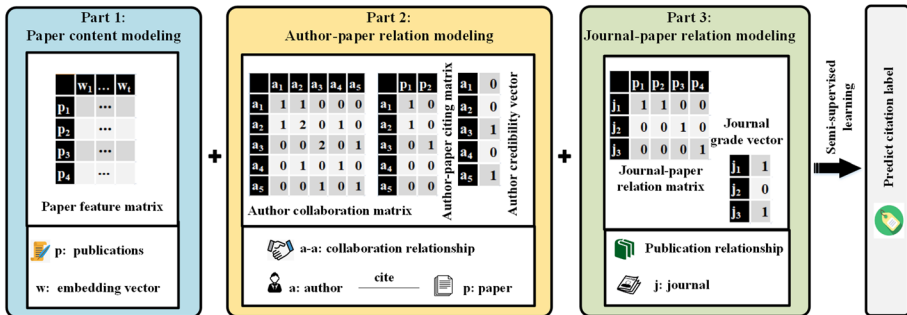


Fig. 2 The overall model framework of ACTION which contains three parts: paper content embedding, author-paper relationship modeling, and journal-paper relationship modeling

published in the journal j_k ; otherwise $\mathbf{B}_{kj} = 0$. $\mathbf{j} \in \mathbb{R}^{l \times 1}$ represents the journal grade. We will introduce how to get the author credibility and the journal grade in the next section.

We treat anomalous citation detection as a binary classification problem. So each citation is either a real citation (non-anomalous citation) or a false one (anomalous citation). We use $\mathbf{Y} \in \mathbb{R}^{n \times n}$ to represent labels for the citations. $\mathbf{Y}_{ij} = 1$ indicates that the citation from the target paper p_i to paper p_j is an anomalous citation; $\mathbf{Y}_{ij} = -1$ indicates that the citation is a normal citation. The description of notations involved in this section is shown in Table 1.

3.3 Problem definition

Based on the notations explained above, the input of the task is composed of paper feature matrix \mathbf{X} , author collaboration times matrix \mathbf{A} , author-paper citing matrix \mathbf{C} , author credibility vector \mathbf{cr} , journal-paper relation matrix \mathbf{B} , journal grade vector \mathbf{j} , and partial labeled citations \mathbf{y}_L . Note that, in this paper, the labeled citations equal to the labeled cited papers. The main goal of the task to predict the label \mathbf{y}_U of remaining unlabeled citations (cited papers). Based on the judgement of citation labels, we can judge whether a paper is a paper containing anomalous citations (defined as anomalous papers).

4 The ACTION framework

In this section, we will give the details of the proposed framework. Specifically, we use a semi-supervised framework to explore the relationships among journals, papers, and authors. The overall architecture of the framework is shown in Fig. 2. ACTION contains three critical parts, including paper content embedding, author-paper relationship modeling, and journal-paper relationship modeling. First, we will introduce the paper latent feature embedding for paper contents. Then, we illustrate how to model the author-paper relationship and journal-paper relationship, respectively. Last, we will emphasize on how to integrate these three parts.

4.1 Paper content embedding

It's important to extract feature representations of paper contents. We utilize the abstract of each paper to represent the paper content for two reasons. (1) Abstract is the most refined part of a paper which can effectively express its central theme. (2) Abstract is brief and easy to obtain. Inspired by previous work, we use Doc2Vec (Le and Mikolov 2014) to map the abstract of the paper to a vector. Doc2Vec is an unsupervised algorithm that can obtain vectors of sentences, paragraphs, and documents, which is an extension of the Word2Vec (Mikolov et al. 2013). It can overcome the shortcomings of traditional bag of words models. There are mainly two steps to obtain the representation:

1. Training stage. The first step is to get word vectors, parameters of softmax, and paragraph vectors or sentence vectors from the training data by the learning process.

2. Inference stage. This step aims to obtain the vector expression for new paragraphs. Specifically, the paragraph vector is initialized randomly. The model will conduct the process of iterative learning to get the final stable sentence vector according to the random gradient descent.

Thus we can obtain the paper content representation matrix \mathbf{X} by transforming paper abstracts into the vector representations by Doc2Vec. It's worth mentioning that after transformation there may be negative values. To solve this problem, we use a linear transformation to make the values positive. After the transformation, the new matrix will preserve the content features and not affect the following steps. Thus we can get a non-negative matrix to represent paper contents for the entire network.

Then we use NMF to get a low-dimensional matrix. As mentioned before, NMF is a matrix decomposition method that can make all decomposed components non-negative. Based on the paper content matrix $\mathbf{X} \in \mathbb{R}^{n \times t}$, we try to find two non-negative matrices $\mathbf{N} \in \mathbb{R}^{n \times d}$ and $\mathbf{K} \in \mathbb{R}^{t \times d}$ by solving the following optimization problem:

$$\mathcal{L}_{\{paper\}} = \min_{\mathbf{N}, \mathbf{K} \geq 0} \|\mathbf{X} - \mathbf{N}\mathbf{K}^T\|_F^2 \quad (1)$$

where d is the dimension of the latent topic space. In addition, \mathbf{N} and \mathbf{K} are non-negative matrices indicating low-dimensional representations of papers and words. Besides, $\mathbf{N} = [\mathbf{N}_L; \mathbf{N}_U]$, where $\mathbf{N}_L \in \mathbb{R}^{r \times d}$ is the papers' latent feature matrix for labeled cited papers and r is the number of labeled citation papers. $\mathbf{N}_U \in \mathbb{R}^{(n-r) \times d}$ denotes the papers' latent feature matrix for unlabeled cited papers.

4.2 Author-paper relationship modeling

With representations for the paper contents, the next step is to model the relationship between papers and the authors' citing behavior. It is based on the assumption that the relationships between authors can reflect the learning process of citations' latent features. Corresponding to the yellow part of Fig. 2, we try to explore the author-paper relationship from the following two aspects.

- We try to use author collaboration times to learn the basic author's potential characteristics because the collaboration can lead to conflicts of interest between authors.
- Based on the labels of the citation and author citing behavior, we try to encode the relationship between author credibility and citation features.

4.2.1 Author feature representation

In academic networks, there are multiple relationships existing among authors such as collaboration relationships. In the collaboration network, scholars tend to form relationships with like-minded friends, rather than those users who have opposing preferences and interests. Scholars who are connected are more likely to share similar latent interests towards papers. Based on Wang et al. (2017), both positive and negative links are related

to scholars' preferences, which indicates that we can model the independent and dependent information of positive and negative links by learning the scholar's preference. We also use NMF method to learn the potential representation of authors. Given the author collaboration times matrix $\mathbf{A} \in \mathbb{R}^{m \times m}$, we finally obtain the non-negative matrix $\mathbf{D} \in \mathbb{R}_+^{m \times d}$ by solving the following optimization problem:

$$\mathcal{L}_{\{author\}} = \min_{\mathbf{D}, \mathbf{U} \geq 0} \|\mathbf{V} \odot (\mathbf{A} - \mathbf{D}\mathbf{U}\mathbf{D}^\top)\|_F^2 \tag{2}$$

where \mathbf{D} is the author's latent matrix. $\mathbf{U} \in \mathbb{R}_+^{d \times d}$ is the author association matrix and $\mathbf{V} \in \mathbb{R}^{m \times m}$ controls the contribution of matrix \mathbf{A} . Since there are only positive samples in \mathbf{A} , we first set $\mathbf{V} = \text{sign}(\mathbf{A})$ and then perform negative sampling and generate the same number of unobserved links and set weights as 0. If $\mathbf{A}_{i,j} = 0$, then $\mathbf{V}_{i,j} = \text{sign}(\mathbf{A}_{i,j}) = 0$. Hence, we can mathematically model positive and negative links in the signed network by adding the matrix \mathbf{V} .

4.2.2 Capturing relations of author collaboration

Author collaboration information could provide additional information for detecting anomalous citations. The study has proved that once the collaboration has been established between scholars, one author prefers to cite papers published by his/her co-authors in his/her subsequent papers (Zhou et al. 2018).

In order to model the author's citing behavior, we consider the inherent association between authors' credibility and the papers they publish. According to labeling theory (Bernburg 2019), deviants are people who have been labeled by a labeler. The calibrator can be a social organization or other individual. Once a person is labeled, he is regarded as a dangerous person. At this time, the self-concept of the deviant also changes, and he considers himself a deviant, which prompts him to continue to have deviant behavior. Wen (2019) analyze the citations of Nobel Prize winners' papers and found that the self-citation rate of these authors is extremely low, which also shows that academic credit is related to anomalous citation behavior. Intuitively, we assume that authors with low reputations are more likely to make anomalous citations, and vice versa. For example, if a_1, a_2, a_4 have low credit, they are more likely to make anomalous citations in comparison with authors with high credit.

We assume that each author has a virtual credibility score. We use $\mathbf{cr} = \{\mathbf{cr}_1, \mathbf{cr}_2, \dots, \mathbf{cr}_m\}$ to represent the credibility for each author. The value of \mathbf{cr}_i is between 0 and 1. \mathbf{cr}_i is obtained by calculating the normal references proportion of the author a_i :

$$\mathbf{cr}_i = \frac{R_{normal}}{R} \tag{3}$$

where R_{normal} and R represent the number of normal references and the total number of the references of the author a_i , respectively.

The definition of anomalous citations brings the problem of identifying the content similarity between the citing paper and its references. That is, given a target paper p_i and its reference p_j , we need to identify the content similarity between p_i and p_j . However, judging the similarity of papers only from the abstract is not comprehensive. Many citations, which should be considered content related, are for referencing methods, providing evidence, etc. Such content relevance cannot be found and measured in abstracts. Citation context can provide important semantic information about the relation between the paper and its references, i.e., the authors' intention. Therefore, we also use citation context to

model the authors' intention in this paper. Researchers also have come up with different schemes composed of citation purpose categories. In this paper, we adopt a scheme that contains six purpose categories including criticizing, comparison, use, substantiating, basis, and others after studying the previously used citation taxonomies (Abu-Jbara et al. 2013). We use the CPU algorithm (Liu et al. 2022) to identify the citation purpose based on the citation context. We can get the probability $CP_{p_i p_j}$ that there is a clear citation purpose between paper p_i and p_j after applying the algorithm. The larger the value, the greater the probability that there is a clear citation purpose between the two papers. Therefore, the author's credibility should also be higher. Hence we further modified the \mathbf{cr}_i of the author a_i :

$$\mathbf{cr}_i = \mathbf{cr}_i * \frac{\sum_{p_i \in P_{a_i}} \sum_{p_j \in R_{p_i}} CP_{p_i p_j}}{R} \tag{4}$$

where P_{a_i} is the set of publications of the author a_i , and R_{p_i} is the set of references of paper p_i . By introducing the author's intention to modify the author's credibility matrix, we can solve the problem that the abstract similarity cannot fully reflect the correlation between the paper and the citation.

We believe that the latent features of low-credit authors are closer to the latent features of anomalous citations, while the latent features of high-credit authors are closer to the latent features of real citations. We get the following optimization formula:

$$\begin{aligned} \mathcal{L}_{\{ACrelation\}} = \min & \sum_{i=1}^m \sum_{j=1}^r \mathbf{C}_{ij} \mathbf{cr}_i \left(1 - \frac{1 + \mathbf{y}_{L_j}}{2}\right) \|\mathbf{D}_i - \mathbf{N}_{L_j}\|_2^2 \\ & + \sum_{i=1}^m \sum_{j=1}^r \mathbf{C}_{ij} (1 - \mathbf{cr}_i) \left(\frac{1 + \mathbf{y}_{L_j}}{2}\right) \|\mathbf{D}_i - \mathbf{N}_{L_j}\|_2^2. \end{aligned} \tag{5}$$

In Eq. (5), $\mathbf{y}_L \in \mathbb{R}^{r \times 1}$ is the label of the partially labeled citations. Then, we consider the following two situations:

- For real normal citations, we set $\mathbf{y}_{L_j} = -1$ and ensure that the latent features of high reputation authors are close to the latent features of real citations.
- For anomalous citations, we set $\mathbf{y}_{L_j} = 1$ and make the latent features of authors with low reputation are close to the latent features of the anomalous citations.

Actually, Eq. (5) can be convert to:

$$\mathcal{L}_{\{ACrelation\}} = \min tr(\mathbf{H}^T \mathbf{L} \mathbf{H}). \tag{6}$$

The detailed derivation process is presented as follows. If $\mathbf{H} = [\mathbf{D}; \mathbf{N}_L] \in \mathbb{R}^{(m+r) \times d}$ and $\mathbf{O}_{ij} = \mathbf{C}_{ij} (\mathbf{cr}_i (1 - \frac{1 + \mathbf{y}_{L_j}}{2}) + (1 - \mathbf{cr}_i) \frac{1 + \mathbf{y}_{L_j}}{2})$, Eq. (5) can be simplified to:

$$\begin{aligned}
 \mathcal{L}_{\{ACrelation\}} &= \min \sum_{i=1}^m \sum_{j=1}^r C_{ij} \mathbf{c}r_i \left(1 - \frac{1 + \mathbf{y}_{L_j}}{2}\right) \|\mathbf{D}_i - \mathbf{N}_{L_j}\|_2^2 + \\
 &\sum_{i=1}^m \sum_{j=1}^r C_{ij} (1 - \mathbf{c}r_i) \left(\frac{1 + \mathbf{y}_{L_j}}{2}\right) \|\mathbf{D}_i - \mathbf{N}_{L_j}\|_2^2 \iff \\
 &\min \sum_{i=1}^m \sum_{j=1}^r \mathbf{O}_{ij} \|\mathbf{D}_i - \mathbf{N}_{L_j}\|_2^2 \iff \tag{7} \\
 &\min \sum_{i=1}^m \sum_{j=1}^r \mathbf{O}_{ij} \|\mathbf{H}_i - \mathbf{H}_j\|_2^2 \iff \\
 &\min \sum_{i,j=1}^{m+r} \mathbf{E}_{ij} \|\mathbf{H}_i - \mathbf{H}_j\|_2^2
 \end{aligned}$$

where \mathbf{E}_{ij} is computed as:

$$\mathbf{E}_{ij} = \begin{cases} 0, & i, j \in [1, m] \text{ or } i, j \in [m + 1, m + r] \\ \mathbf{O}_{i(j-m)}, & i \in [1, m], j \in [m + 1, m + r] \\ \mathbf{O}_{(i-m)j}, & i \in [m + 1, m + r], j \in [1, m]. \end{cases} \tag{8}$$

We use $\mathbf{L} = \mathbf{S} - \mathbf{E}$ to represent the Laplacian matrix and \mathbf{S} is a diagonal matrix with diagonal element $S_{ii} = \sum_{j=1}^{m+r} \mathbf{E}_{ij}$, where \mathbf{E} is symmetric. Then Eq. (7) can be finally rewritten as:

$$\mathcal{L}_{\{ACrelation\}} = \min \sum_{i,j=1}^{m+r} \mathbf{E}_{ij} \|\mathbf{H}_i - \mathbf{H}_j\|_2^2 \iff \min \text{tr}(\mathbf{H}^T \mathbf{L} \mathbf{H}). \tag{9}$$

4.3 Journal-paper relationship modeling

Strain theory suggests that lower status individuals and organizations are more likely to engage in inappropriate behavior to achieve goals that cannot be achieved through legitimate means (Greve et al. 2010; Hall and Martin 2019). Individuals and organizations will select reference groups to measure their achievements and performance (Frank 1985). Both low-status and middle-status imply relative deprivation compared with relatively higher-status peers. Therefore, lower-status journals may be more willing to engage in index inflation in pursuit of improved scholarly metrics. Siler and Larivière (2022) also propose and verify that low and middle-status journals will be more likely to engage in JIF inflation than high status journals. Hence, we assume that papers published in lower-impact journals have a higher probability of having anomalous citations than papers published in higher-impact journals. Therefore, exploring the impact of the journals can help us detect anomalous citations correctly. Journal impact factor (JIF) (Garfield 1972) is an internationally accepted journal evaluation index. According to journals’ JIF, we regard \mathbf{j} as the journal grade vector. Because journal impact factors of different journals are quite different, we normalize them to $[0, 1]$. As shown in the green part of Fig. 2, the basic idea is to use the journal’s grade matrix $\mathbf{j} \in \mathbb{R}^{l \times 1}$ and the journal-paper relation matrix $\mathbf{B} \in \mathbb{R}^{l \times n}$ to optimize the feature representation learning of the citations:

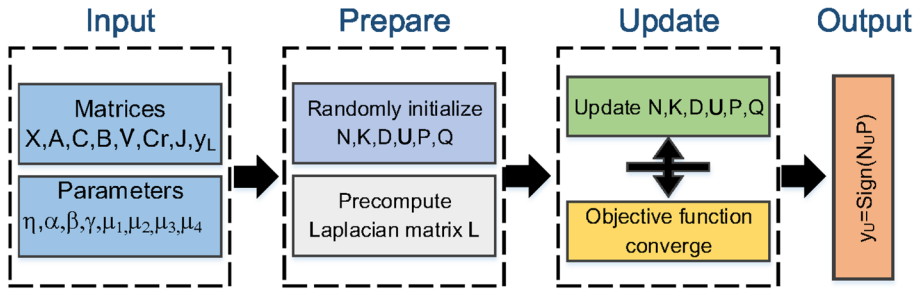


Fig. 3 The calculation process of ACTION

$$\mathcal{L}_{\{JPrelation\}} = \min \|\bar{\mathbf{B}}\mathbf{N}\mathbf{Q} - \mathbf{j}\|_2^2. \tag{10}$$

We suppose that the latent features of a journal can be represented by the papers it publishes such as $\bar{\mathbf{B}}\mathbf{N}$. $\bar{\mathbf{B}}$ is a normalized journal-paper relation matrix. $\mathbf{Q} \in \mathbb{R}^{d \times 1}$ is a weight matrix that maps the potential features of the journal to the corresponding journal grade vector \mathbf{j} .

4.4 Classification model

We introduce how to capture the latent feature representations of the cited papers by modeling relationships among journals, papers, and authors. We further use a semi-supervised linear classification model to learn the latent features. We try to minimize the following equation:

$$\mathcal{L}_{\{classification\}} = \min \|\mathbf{N}_L \mathbf{P} - \mathbf{y}_L\|_2^2 \tag{11}$$

where $\mathbf{P} \in \mathbb{R}^{d \times 1}$ is a weighting matrix that maps the potential features of the paper to the anomalous citation labels. Thus we can identify the anomalous citations by the classification model.

4.5 Optimization

If we merge all the parts mentioned above, we get the objective function:

$$\begin{aligned} \mathcal{L}_{\{final\}} = & \min_{\theta} \|\mathbf{X} - \mathbf{N}\mathbf{K}^T\|_F^2 + \alpha \|\mathbf{V} \odot (\mathbf{A} - \mathbf{D}\mathbf{U}\mathbf{D}^T)\|_F^2 \\ & + \beta tr(\mathbf{H}^T \mathbf{L}\mathbf{H}) + \gamma \|\bar{\mathbf{B}}\mathbf{N}\mathbf{Q} - \mathbf{j}\|_2^2 + \eta \|\mathbf{N}_L \mathbf{P} - \mathbf{y}_L\|_2^2 \\ & + \lambda (\|\mathbf{N}\|_F^2 + \|\mathbf{K}\|_F^2 + \|\mathbf{D}\|_F^2 + \|\mathbf{U}\|_F^2 + \|\mathbf{P}\|_2^2 + \|\mathbf{Q}\|_2^2) \\ & s.t. \ \mathbf{N}, \mathbf{K}, \mathbf{D}, \mathbf{U} \geq 0 \end{aligned} \tag{12}$$

where $\beta tr(\mathbf{H}^T \mathbf{L}\mathbf{H})$ is the simplified form of Eq. (5). The last term is introduced to avoid over-fitting. $\mathbf{N} = [\mathbf{N}_L; \mathbf{N}_U]$ consists of labeled and unlabeled parts. In the process of computing derivatives of \mathbf{N} , we should first compute the derivatives of these two parts separately. Similarly, \mathbf{H} and \mathbf{X} also contain two parts, including labeled part and unlabeled part. $\mathbf{H} = [\mathbf{D}; \mathbf{N}_L]$, $\mathbf{X} = [\mathbf{X}_L, \mathbf{X}_U]$. \mathbf{L} is the Laplacian matrix and we rewrite

$\mathbf{L} = [\mathbf{L}_{11}, \mathbf{L}_{12}; \mathbf{L}_{21}, \mathbf{L}_{22}]$ in order to facilitate the separate derivation of the labeled and unlabeled parts. $\mathbf{L}_{11} \in \mathbb{R}^{m \times m}$, $\mathbf{L}_{12} \in \mathbb{R}^{m \times r}$, $\mathbf{L}_{21} \in \mathbb{R}^{r \times m}$, $\mathbf{L}_{22} \in \mathbb{R}^{r \times r}$.

Our goal is to get the optimal solution of the loss function. In this process, we adopt gradient descent to obtain the optimal solution. The calculation process is shown as Fig. 3. At the optimization process, we first randomly initialize \mathbf{D} , \mathbf{K} , \mathbf{U} , \mathbf{N} , \mathbf{P} , \mathbf{Q} and construct the Laplacian matrix \mathbf{L} . Then we repeatedly update related parameters until convergence. After training to obtain the optimal parameters, we predict the labels of unlabeled citations \mathbf{y}_U through calculating $\mathbf{y}_U = \text{Sign}(\mathbf{D}_U \mathbf{P})$. Hence we can get the labels \mathbf{y}_U for unlabeled citations.

Next, we show the process of deriving partial derivatives for each matrix variable.

1. Compute the partial derivative of the objective function with respect to \mathbf{N}_L and \mathbf{N}_U separately and obtain the partial derivative of \mathbf{N} .

$$\frac{1}{2} \frac{\partial l}{\partial \mathbf{N}_L} = (\mathbf{N}_L \mathbf{K}^\top - \mathbf{X}_L) \mathbf{K} + \beta \mathbf{L}_{21} \mathbf{D} + \beta \mathbf{L}_{22} \mathbf{N}_L + \gamma \bar{\mathbf{B}}_L^\top (\bar{\mathbf{B}}_L \mathbf{N}_L \mathbf{Q} - \mathbf{j}) \mathbf{Q}^\top + \eta (\mathbf{N}_L \mathbf{P} - \mathbf{y}_L) \mathbf{P}^\top + \lambda \mathbf{N}_L \tag{13}$$

$$\frac{1}{2} \frac{\partial l}{\partial \mathbf{N}_U} = (\mathbf{N}_U \mathbf{K}^\top - \mathbf{X}_U) \mathbf{K} + \gamma \bar{\mathbf{B}}_U^\top (\bar{\mathbf{B}}_U \mathbf{N}_U \mathbf{Q} - \mathbf{j}) \mathbf{Q}^\top + \lambda \mathbf{N}_U \tag{14}$$

$$\frac{1}{2} \frac{\partial l}{\partial \mathbf{N}} = (\mathbf{N} \mathbf{K}^\top - \mathbf{X}) \mathbf{K} + \gamma \bar{\mathbf{B}}^\top (\bar{\mathbf{B}} \mathbf{N} \mathbf{Q} - \mathbf{j}) \mathbf{Q}^\top + \lambda \mathbf{N} + [\beta \mathbf{L}_{21} \mathbf{D} + \beta \mathbf{L}_{22} \mathbf{N}_L + \eta (\mathbf{N}_L \mathbf{P} - \mathbf{y}_L) \mathbf{P}^\top; 0] \tag{15}$$

2. Calculate the partial derivative of l on \mathbf{D} .

$$\frac{1}{2} \frac{\partial l}{\partial \mathbf{D}} = \alpha (\mathbf{V} \odot (\mathbf{D} \mathbf{U} \mathbf{D}^\top - \mathbf{A})) \mathbf{D} \mathbf{U}^\top + \alpha (\mathbf{V} \odot (\mathbf{D} \mathbf{U} \mathbf{D}^\top - \mathbf{A}))^\top \mathbf{D} \mathbf{U} + \lambda \mathbf{D} + \beta (\mathbf{L}_{11} \mathbf{D} + \mathbf{L}_{12} \mathbf{N}_L) \tag{16}$$

3. Calculate the partial derivative of l on \mathbf{K} .

$$\frac{1}{2} \frac{\partial l}{\partial \mathbf{K}} = (\mathbf{N} \mathbf{K}^\top - \mathbf{X})^\top \mathbf{N} + \lambda \mathbf{K} \tag{17}$$

4. Calculate the partial derivative of l on \mathbf{U} .

$$\frac{1}{2} \frac{\partial l}{\partial \mathbf{U}} = \alpha \mathbf{D}^\top (\mathbf{V} \odot (\mathbf{D} \mathbf{U} \mathbf{D}^\top - \mathbf{A})) \mathbf{D} + \lambda \mathbf{U} \tag{18}$$

The update rules of \mathbf{N} can be obtained by gradient descent:

$$\mathbf{N}_{ij} = \mathbf{N}_{ij} - \mu_1 \left(2 * \frac{1}{2} \frac{\partial l}{\partial \mathbf{N}_{ij}} \right) = \mathbf{N}_{ij} - \mu_1 \frac{\partial l}{\partial \mathbf{N}_{ij}} \tag{19}$$

Similarly,

$$\mathbf{D}_{ij} = \mathbf{D}_{ij} - \mu_2 \frac{\partial l}{\partial \mathbf{D}_{ij}} \tag{20}$$

For \mathbf{K} and \mathbf{U} :

$$\mathbf{K}_{ij} = \mathbf{K}_{ij} - \mu_3 \frac{\partial l}{\partial \mathbf{K}_{ij}} \tag{21}$$

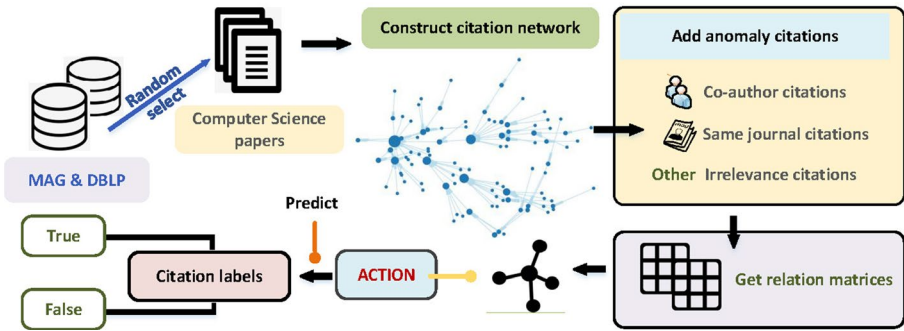


Fig. 4 The experimental process of identifying anomalous citations in the academic network

$$U_{ij} = U_{ij} - \mu_4 \frac{\partial l}{\partial U_{ij}} \tag{22}$$

Finally, we can update \mathbf{P} and \mathbf{Q} as:

$$\mathbf{P} = (\eta \mathbf{N}_L^T \mathbf{N}_L + \lambda \mathbf{I})^{-1} \eta \mathbf{N}_L^T \mathbf{y}_L \tag{23}$$

$$\mathbf{Q} = (\gamma \mathbf{N}^T \bar{\mathbf{B}}^T \bar{\mathbf{B}} \mathbf{N} + \lambda \mathbf{I})^{-1} \gamma \mathbf{N}^T \bar{\mathbf{B}}^T \mathbf{j} \tag{24}$$

where \mathbf{I} is the unit matrix.

4.6 Model complexity

In each iteration, the computational complexity for computing \mathbf{N} is $O(nd + nd^2 + rd + rm + n^2)$, where n , l , and m represent the number of papers, journals, and authors, respectively. d is the dimension of the latent space. Similarly, the cost for computing \mathbf{K} is $O(tnd)$, where t is the dimension of the paper content matrix. The computation cost for \mathbf{D} is $O(m^4 d^3 + md)$, for \mathbf{U} is $O(m^4 d^3 + m^2 d^2)$. The time complexity for updating \mathbf{P} and \mathbf{Q} are $O(d^3 + d^2 + dr)$ and $O(d^2 ln + d^3 + dl)$, where r represents the number of labeled papers.

5 Experiments

In this section, we will verify the effectiveness of ACTION. In particular, we discuss the following questions:

- Whether we can improve the performance of identifying anomalous citations by modeling the journal and author simultaneously?
- Whether learning of journal grades and author credibility play an important role in identifying anomalous citations?

Table 2 The statistics of datasets

Datasource	MAG	DBLP	CiteseerX
# Candidate papers	204	240	302
# True papers	102	120	151
# False papers	102	120	151
# Anomalous citations	739	706	688
# Authors	3741	5625	7489
# Citing Links	17431	17065	8612
# Collaboration Links	19095	29826	39847
# Journals	341	720	518

We first illustrate the datasets used in this paper. Then we analyze the performance of ACTION from the perspectives of anomalous boundary and parameter sensitivity. The whole experimental procedure is shown in Fig. 4.

5.1 Datasets

We can obtain citation relationships from different academic datasets, such as MAG and DBLP. However, there are no recognized datasets for anomalous citations. Especially, there is no authoritative expert to point out which papers exist anomalous citations. So, we manually add non-existent references in the original papers as anomalous citations and take the original references in the papers as real citations. The reason for constructing datasets based on the MAG and DBLP datasets is that they all contain abundant information that we need, e.g., paper's abstract, authors, journal, published year and references. We randomly extract some papers with complete information in MAG and DBLP in the field of Computer Science as central papers to establish citation networks respectively. When constructing the citation network based on these central papers, we ensure that the papers having the citing-cited relationship with these papers are included in the constructed dataset. The anomalous citation links we added are based on the constructed citation network, which ensures that both the citing paper and the cited paper are included.

We randomly select half of the papers to add anomalous references. Then the citation from the original paper to the added papers can be regarded as an anomalous citation. We add the same number of anomalous references as original references for each paper. We add three types of references: (1) Citing collaborators' publications; (2) Citing the same journal's publications; (3) Citing interdisciplinary publications with irrelevant contents. Table 2 lists the basic information of constructed datasets. We first choose 204 papers and 240 papers from two datasets respectively and then select half of the papers to add anomalous references. True papers represent the papers without anomalous citations and false papers represent the papers with anomalous citations. The number of citing links is obtained by adding each element in the author-paper citing matrix. Similarly, the number of collaboration links is obtained by adding each element in the author collaboration times matrix. From the table, we can see that the two

Table 3 Summary of the detection methods for comparison

Variants of ACTION	Paper content	Journal & paper relation	Author & paper relation	Author credibility	Journal grade
ACTION-JP	√	-	√	√	-
ACTION-AP	√	√	-	-	√
ACTION-JA	√	-	-	-	-
ACTION-Cr	√	√	√	-	√
ACTION-J	√	√	√	√	-
ACTION	√	√	√	√	√

datasets differ in the number of authors and journals. Therefore, there are differences in collaboration density and citing density.

Furthermore, we have artificially created another anomalous dataset motivated by previous research (Kojaku et al. 2021; Liu et al. 2022) based on CiteseerX dataset. The CiteseerX dataset is a commonly used dataset for academic network analysis. We first collect the paper titles which are reported as papers causing “journal cartels”. The set of anomalous edges consists of the edges between anomalous papers and their references/citations that meet the anomalous citation relationship. We also need to ensure that information for these papers and their references in the CiteseerX dataset is complete and the citation contexts between them are easy to obtain. Finally, the dataset contains 8612 citation links. The number of anomalous edges is 688, accounting for 7.98% of all edges, which is similar to the proportion of anomalous edges in the real world.

5.2 Evaluation metrics

We use Accuracy, Precision, Recall, and F1 to evaluate the performance of ACTION, which are common metrics in the classification tasks. The metrics are calculated as:

- $Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$
- $Precision = \frac{TP}{TP + FP}$
- $Recall = \frac{TP}{TP + FN}$
- $F1 = 2 * \frac{Precision * Recall}{Precision + Recall}$

where TP , TN , FP , FN represent the number of anomalous citations correctly classified (true positive), the number of normal citations correctly classified (false positive), the number of normal citations misclassified (false positive), and the number of anomalous citations misclassified (false negative), respectively.

5.3 Baselines and variants

Due to the lack of any existing baseline for the given task, we decide to use the existing dimension reduction method as baselines because in our framework, we utilize NMF

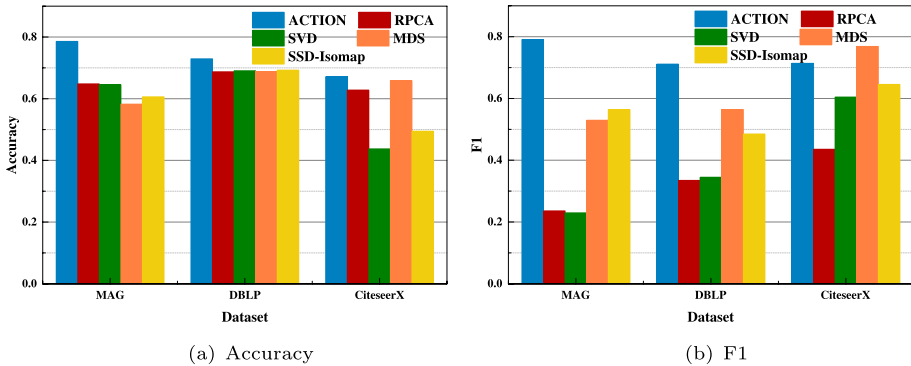


Fig. 5 Dimension reduction methods comparison results

to get the low-dimensional matrix. There are many ways to reduce the dimension of the matrix in machine learning. We choose some commonly used dimension reduction methods including Robust Principal Component Analysis (RPCA) (Lee and Choe 2018), Singular Value Decomposition (SVD) (Wang and Zhu 2017), Multidimensional Scaling (MDS) (Borg and Groenen 2010), and SSD-Isomap (Rui et al. 2019) as baselines.

- **RPCA:** Principal Component Analysis (PCA) is one of the most widely used data dimension reduction algorithms. By calculating the covariance matrix of data, the eigenvector of the covariance matrix can be obtained. The eigenvectors corresponding to k features with the largest eigenvalues are selected to form the matrix. In this way, the data matrix can be transformed into a new space to reduce the dimension of data characteristics. However, the robustness of PCA is not good. RPCA can solve the problem of poor robustness of PCA.
- **SVD:** SVD has great ability to extract information. It can simplify the data, remove noise and improve the performance. But the process of data conversion may be difficult to understand.
- **MDS:** MDS is a very classical method for Manifold Learning. It can simplify the research object (sample or variable) of multidimensional space to the low-dimensional space for positioning, analysis, and classification while retaining the original relationship between objects.
- **SSD-Isomap:** SSD-Isomap is a semi-supervised version of Isometric Feature Mapping (ISOMAP). ISOMAP's core algorithm is consistent with MDS. The difference lies in the calculation of the distance matrix in the original space.

We use SVM as the classifier to conduct classification for all baselines. Furthermore, to analyze the importance of each module, we compare ACTION with different variants.

- **ACTION-JP:** We eliminate the effect of journal-paper relation module and only utilize paper content module and author-paper relation module to identify anomalous citations.
- **ACTION-AP:** We ignore the effect of author-paper relation module and only use paper content module and journal-paper relation module to identify anomalous citations.

Table 4 Performance comparison for different anomalous citations

Metrics	Collaborators' publications	Same journal's publications	Irrelevant contents
Accuracy	0.66	0.63	0.65
Precision	0.4	0.64	0.66
Recall	0.56	0.97	0.98
F1	0.47	0.76	0.79

- ACTION-JA: We only use paper content module for identification anomalous citations.

We also analyze the role of author credibility and journal grade.

- ACTION-Cr: We remove the effect of the author credibility in author-paper relation module.
- ACTION-J: We eliminate the effect of the journal grade in journal-paper relation module.

Table 3 presents the summary of the variants.

5.4 Results comparison and analysis

We present our results together with some case studies from two perspectives, (1) comparison results with dimension reduction methods, (2) comparison results with ACTION variants.

5.4.1 Dimension reduction methods comparison

We first compare ACTION with baselines mentioned in Sect. 5.3. We run the models on the MAG dataset and DBLP dataset respectively. The experimental results are presented in Fig. 5. From the figure, we can see that ACTION outperforms all baselines, especially on the MAG dataset. On the DBLP dataset, RPCA, SVD, MDS, and SSD-Isomap achieve similar accuracy. The F1 of ACTION comes out to be 79% and 71% on the MAG and DBLP dataset respectively, which is higher than that of the baselines. Specifically, on the

Table 5 Performance comparison for anomalous citations detection

Datasets	Metric	ACTION	ACTION-JP	ACTION-AP	ACTION-JA	ACTION-Cr	ACTION-J
MAG	Accuracy	0.786	0.762	0.738	0.714	0.714	0.691
	Precision	0.773	0.923	1.000	0.909	0.765	0.833
	Recall	0.810	0.571	0.476	0.476	0.619	0.476
	F1	0.791	0.706	0.645	0.625	0.684	0.606
DBLP	Accuracy	0.729	0.646	0.667	0.604	0.646	0.646
	Precision	0.762	0.889	0.900	1.000	0.889	0.733
	Recall	0.667	0.333	0.375	0.208	0.333	0.458
	F1	0.711	0.485	0.529	0.345	0.485	0.564

Bold numbers represent the best results

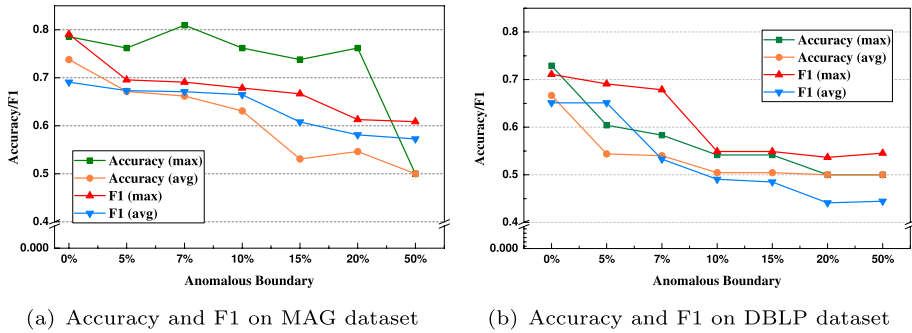


Fig. 6 Performance of anomalous paper identification under different anomalous boundaries

DBLP dataset, ACTION is 55%, 56%, 26%, and 23% higher than RPCA, SVD, MDS, and SSD-Isomap. For the MAG dataset, ACTION is 38%, 37%, 15%, and 23% higher than RPCA, SVD, MDS, and SSD-Isomap. The superiority of ACTION over baselines further demonstrates the effectiveness of the proposed framework. We also calculate the experimental results on the CiteseerX dataset, as shown in Fig. 5. From the experimental results, we can see that the model performance on the CiteseerX dataset is slightly worse compared with the MAG and DBLP datasets. One of the possible reasons is that the recognition accuracy for citation purpose cannot fully reach 100%. The Precision exceeds 74% and the F1 comes out to be 71%.

Finally, we split each dataset to test the model performance. In particular, we use the anomalous citations that only add the co-authors' publications, the same journal's publications, and interdisciplinary publications with irrelevant contents as the input of the model. The experimental results are shown in Table 4. From the experimental results, we can see that ACTION can better identify anomalous citations with irrelevant contents compared with the other types of anomalous citations. According to the F1 score, we can see that ACTION is less effective in identifying anomalous citations from collaborators.

5.4.2 ACTION variants comparison

Next, we evaluate the effectiveness of each module in ACTION on the task of anomalous citations detection. We evaluate and compare the variants of ACTION mentioned above on two different datasets. The results are shown in Table 5. Based on results from Table 5, the discoveries are listed as follows.

- (1) In terms of F1, ACTION is 8.48%, 22.64% higher than ACTION-JP on MAG and DBLP datasets, respectively. It shows that the journal-paper relation module is very important. In terms of F1, ACTION is 14.55%, 18.18% higher than ACTION-AP on MAG and DBLP datasets, respectively. The results suggest that the author-paper relation is indispensable.

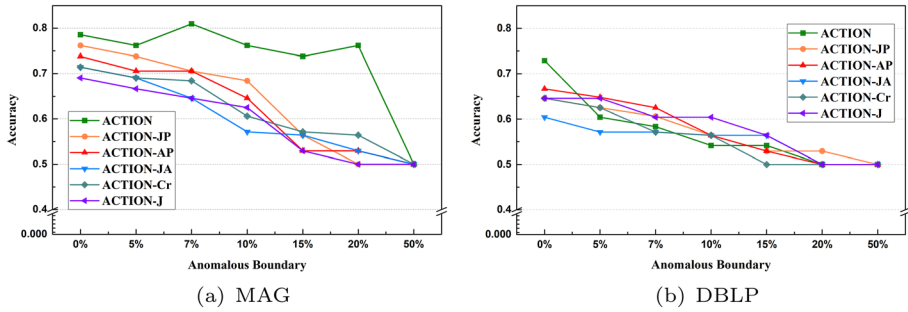


Fig. 7 Performance of anomalous paper identification results for variants of ACTION under different anomalous boundaries

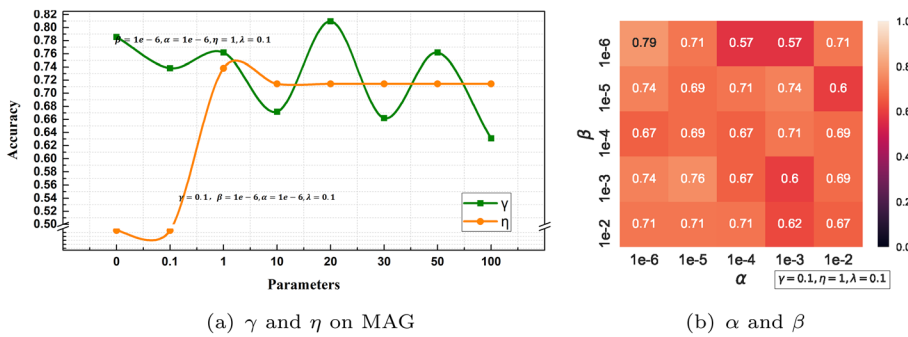


Fig. 8 Parameters analysis results

- (2) We can see that ACTION performs better than ACTION-JA which is only based on paper content. In terms of F1, ACTION is 16.57%, 36.64% higher than ACTION-JA on MAG and DBLP datasets, respectively. What's more, the ACTION-JA performs worse than ACTION-JP and ACTION-AP. Compared with ACTION-J, the accuracy reduces 4.76%, 4.16% on MAG and DBLP, respectively. Compared with ACTION-AP, the accuracy reduces 2.38%, 6.25% on MAG and DBLP, respectively. The F1 reduces 2.02%, 18.46% on MAG and DBLP, respectively. The above analysis shows that the joint modeling of journal and author plays a very important role in identifying anomalous citations.
- (3) ACTION performs better than ACTION-Cr which eliminates the effect of the author credibility. In terms of F1, ACTION is 10.65%, 22.64% higher than ACTION-Cr on MAG and DBLP datasets, respectively. It indicates that the author credibility provides additional information for identifying anomalous citations.
- (4) ACTION performs better than ACTION-J which removes journal grade vector. In terms of F1, ACTION is 18.46%, 14.71% higher than ACTION-J on MAG and DBLP datasets, respectively. We can conclude that the journal grade provides supplementary information for identifying anomalous citations.

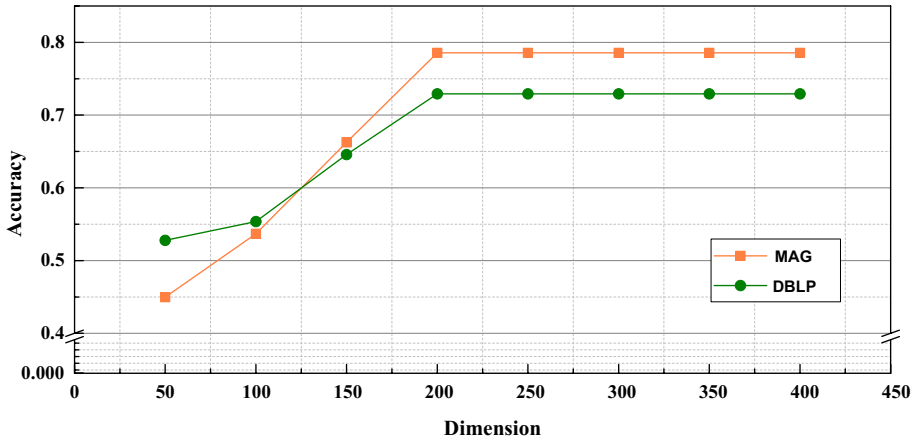


Fig. 9 Parameter sensitive in terms of dimensionality of word embedding

From the analysis of the experimental results, we conclude that: (1) joint modeling of the relationships between journals and authors contribute to the performance of ACTION; (2) author credibility and journal grade are necessary for identifying anomalous citations.

5.5 Anomalous paper identification results

In order to identify anomalous papers, we introduce the concept of anomalous rate. Anomalous rate refers to the percentage of a paper's anomalous citations to its total citations. It is calculated as:

$$\text{Anomalous rate} = \frac{\text{Anomalous citations}}{\text{Total citations}} \quad (25)$$

Anomalous boundary helps us to judge whether a paper is an abnormal paper. When a paper's anomalous rate is greater than the anomalous boundary, we regard it as an anomalous paper.

Figure 6 shows the anomalous paper classification results for ACTION under different anomalous boundaries. The blue line represents the highest accuracy, while the orange line represents the average accuracy. The results show a consistent downtrend both on the MAG dataset and the DBLP dataset. By observing and comparing the average accuracy of datasets, we find that the highest point is at the initial point. With the increase of anomalous boundary value, the accuracy tends to be 50%. The reason is that the number of papers with anomalous citations is the same as the number of papers without anomalous citations. The worst case is that all the papers are identified as papers without anomalous citations.

We also evaluate the performance of ACTION variants on the task of anomalous paper identification. The results are summarized in Fig. 7. As can be expected, all variants have decreasing performance with the increase of anomalous rate. Obviously, ACTION always has the best performance on the MAG dataset in comparison with its variants. In particular, when we set anomalous rate higher than 20%, ACTION is much superior to its variants.

5.6 Parameter sensitivity

We also explored how hyperparameters affect the performance of ACTION, including η , γ , α , β , and the embedding dimension.

The parameters of the experiment are obtained by cross-validation. There are four critical parameters in the framework. η controls the contribution of the semi-supervised classifier. γ controls the contribution of the journal-paper relation model. α and β control the contribution of the author-paper relation model.

Parameter η . As mentioned before, η controls the contribution of the classifier. We fix the value of γ , β , α and λ to verify the influence of η . Taking the parameter settings on MAG dataset as an example, we first fix $\{\gamma = 0.1, \beta = 10^{-6}, \alpha = 10^{-6}, \lambda = 0.1\}$. Then we vary η as $\{0, 1, 10, 20, 50, 100\}$, the performance is shown in Fig. 8a. Generally, the performance increases with the increase of η in the certain range. When η varies from 0 to 1, the model accuracy increases dramatically. The result of the experiment demonstrates the importance of semi-supervised classification under the framework. When we vary η as $\{10, 20, 50, 100\}$, the results keep stable.

Parameter γ . Parameter γ controls the weight of journal-paper relation model in the overall ACTION framework. Figure 8a shows the accuracy on the MAG dataset with different value of γ . We set $\{\beta = 10^{-6}, \alpha = 10^{-6}, \eta = 1, \lambda = 0.1\}$, then we vary γ as $\{0, 0.1, 1, 10, 20, 30, 50, 100\}$. From the results, we can easily find that when γ varies from 0 to 0.1, the performance of ACTION increases dramatically. It demonstrates the importance of journal-paper module under the framework.

Parameters (α , β). To explore the influence of α and β , we set $\{\gamma = 0.1, \eta = 1, \lambda = 0.1\}$, then we vary α as $\{10^{-6}, 10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}\}$ and β in $\{10^{-6}, 10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}\}$. The performance variations are depicted in Fig. 8b. We can see that when α and β vary from 0 to 10^{-2} , the performance in terms of accuracy tends to decrease first and then increase. It achieves relatively best performance when $\{\beta = 10^{-6}, \alpha = 10^{-6}\}$.

Embedding dimensions. In the process of paper content embedding, we use Doc2Vec to transform paper abstracts into low-dimensional representations. Generally, the dimension of the generated vectors is the same as that of the hidden layer in the neural network. In fact, empirical studies have shown that the learning process of word embedding will also lead to model under-fitting and over-fitting. Thus, it is important to select an appropriate embedding dimension. Figure 9 shows the sensitivity in terms of the dimensionality. From the results, we can easily find that the performance keeps increasing with the increase of embedding dimension in a certain region both on DBLP and MAG datasets. Furthermore, it is relatively stable when the embedding dimension $d > 200$.

6 Conclusion

In this work, we propose a novel framework, namely ACTION, to detect anomalous citations across heterogeneous academic networks. In the academic network, there are multiple types of relationships among different academic entities. ACTION can make full use of

information provided by journals, authors and contents of papers for identifying anomalous citations. The experiments are carried out on three artificially generated datasets from two aspects, including comparing ACTION with dimension reduction methods and comparing ACTION with ACTION variants. From the results, we observe that the performance of ACTION is at least 15% better than dimension reduction baselines on both DBLP and MAG datasets. Furthermore, by comparing ACTION with its variants, we can find that joint modeling of the relationships between journals and authors contributes to the performance of ACTION. At the same time, author credibility and journal grade are necessary for identifying anomalous citations.

There are some limitations to our work which must be borne in mind. For example, our research is focused on computer science area. It can be extended to other fields in future work. Furthermore, the construction of anomalous datasets and the labeling of anomalous citations are time-consuming. We can consider how to improve the efficiency of the algorithm so that it can be applied to a large-scale dataset. Finally, in this work, we treat the anomalous citation detection problem as a classification problem. In future work, we will consider solving this problem by using probabilistic modeling methods to help us determine whether the anomalous citations are acceptable.

Author contributions JL, FX, and XB designed the research; JL, XB, MW analyzed the data and performed the experiments; JL, FX and ST wrote the paper. All authors reviewed the manuscript.

Funding This work was partially supported by National Natural Science Foundation of China under Grant No. 72204037, MOE (Ministry of Education in China) Liberal arts and Social Sciences Foundation (Project No. 22YJC870009), Dalian Science and Technology Talent Innovation Support Program (Youth Science and Technology Star) (Project No. 2022RQ055), and the Fundamental Research Funds for the Central Universities under Grant DUT22RC(3)027.

Data availability The datasets generated and analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of interest The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abu-Jbara A, Ezra J, Radev D (2013) Purpose and polarity of citation: towards NLP-based bibliometrics. In: Proceedings of the 2013 conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pp 596–606
- Bai X, Xia F, Lee I et al (2016) Identifying anomalous citations for objective evaluation of scholarly article impact. *PLoS ONE* 11(9):e0162,364

- Bai X, Lee I, Ning Z et al (2017) The role of positive and negative citations in scientific evaluation. *IEEE Access* 5:17,607–17,617
- Bai X, Pan H, Hou J et al (2020) Quantifying success in science: an overview. *IEEE Access* 8:123,200–123,214
- Bernburg JG (2019) Labeling theory. *Handbook on crime and deviance* pp 179–196
- Biagioli M, Kenney M, Martin BR et al (2019) Academic misconduct, misrepresentation and gaming: a reassessment. *Res Polic* 48(2):401–413
- Borg I, Groenen P (2010) Modern multidimensional scaling: theory and applications. *J Educ Meas* 40(3):277–280
- Bu Y, Wang B, Wb Huang et al (2018) Using the appearance of citations in full text on author co-citation analysis. *Scientometrics* 116(1):275–289
- Cai L, Tian J, Liu J et al (2019) Scholarly impact assessment: a survey of citation weighting solutions. *Scientometrics* 118:453–478
- Campanario JM (2011) Large increases and decreases in journal impact factors in only one year: the effect of journal self-citations. *J Am Soc Inf Sci Technol* 62(2):230–235
- Chawla DS (2019) Elsevier investigates hundreds of peer reviewers for manipulating citations. *Nature* 573(7773):174–175
- Chorus C, Waltman L (2016) A large-scale analysis of impact factor biased journal self-citations. *PLoS One* 11(8):e0161,021
- Corbyn Z (2010) An easy way to boost a paper's citations. *Nature* 406:510–515
- Dai X, Su X, Zhang W et al (2020) Robust manhattan non-negative matrix factorization for image recovery and representation. *Inf Sci* 527:70–87
- Fang Y, Yin J, Wu B (2018) Climate change and tourism: a scientometric analysis using citespace. *J Sustain Tour* 26(1):108–126
- Fister I Jr, Fister I, Perc M (2016) Toward the discovery of citation cartels in citation networks. *Front Phys* 4:49
- Fortunato S, Bergstrom CT, Börner K, et al (2018) Science of science. *Science* 359(6379)
- Frank G (1999) Scientific communication—a vanity fair? *Science* 286(5437):53–55
- Frank RH (1985) *Choosing the right pond: human behavior and the quest for status*. Oxford University Press
- Garfield E (1972) Citation analysis as a tool in journal evaluation journals can be ranked by frequency and impact of citations for science policy studies. *Science* 178(4060):471–479
- Gazni A, Didegah F (2021) Journal self-citation trends in 1975–2017 and the effect on journal impact and article citations. *Learn Publ* 34(2):233–240
- Greenland P, Fontanarosa PB (2012) Ending honorary authorship. *Science* 337(6098):1019
- Greve HR, Palmer D, Pozner JE (2010) Organizations gone wild: The causes, processes, and consequences of organizational misconduct. *Acad Manag Ann* 4(1):53–107
- Hall J, Martin BR (2019) Towards a taxonomy of research misconduct: the case of business school research. *Res Polic* 48(2):414–427
- Hamamoto R, Takasawa K, Machino H et al (2022) Application of non-negative matrix factorization in oncology: one approach for establishing precision medicine. *Brief Bioinf* 23(4):246
- Humphrey C, Kiseleva O, Schleicher T (2019) A time-series analysis of the scale of coercive journal self-citation and its effect on impact factors and journal rankings. *Eur Account Rev* 28(2):335–369
- Iqbal S, Hassan SU, Aljohani NR et al (2021) A decade of in-text citation analysis based on natural language processing and machine learning techniques: an overview of empirical studies. *Scientometrics* 126(8):6551–6599
- Ji L, Song P, Zhang W (2022) Transferable discriminative non-negative matrix factorization for cross-database facial expression recognition. *Digit Signal Process* 123(103):424
- Jiang X, Liu J (2023) Extracting the evolutionary backbone of scientific domains: the semantic main path network analysis approach based on citation context analysis. *J Assoc Inf Sci Technol* 74(5):546–569
- Kim HJ, Jeong YK, Song M (2016) Content-and proximity-based author co-citation analysis using citation sentences. *J Inf* 10(4):954–966
- Kojaku S, Livan G, Masuda N (2021) Detecting anomalous citation groups in journal networks. *Sci Rep* 11(1):1–11
- Krell FT (2014) Losing the numbers game: abundant self-citations put journals at risk for a life without an impact factor. *Eur Sci Edit* 40(2):36–38

- Le Q, Mikolov T (2014) Distributed representations of sentences and documents. In: International Conference on Machine Learning, pp 1188–1196
- Lee DD, Seung HS (1999) Learning the parts of objects by non-negative matrix factorization. *Nature* 401(6755):788–798
- Lee J, Choe Y (2018) Robust PCA based on incoherence with geometrical interpretation. *IEEE Transact Image Process* 27(4):1939–1950
- Li J, Zhou G, Qiu Y et al (2020) Deep graph regularized non-negative matrix factorization for multi-view clustering. *Neurocomputing* 390:108–116
- Li X, Cui G, Dong Y (2016) Graph regularized non-negative low-rank matrix factorization for image clustering. *IEEE Transact Cybern* 47(11):3840–3853
- Liao H, Tang M, Luo L et al (2018) A bibliometric analysis and visualization of medical big data research. *Sustainability* 10(1):166
- Lin X, Boutros PC (2020) Optimization and expansion of non-negative matrix factorization. *BMC Bioinf* 21(1):1–10
- Liu J, Ren J, Zheng W, et al (2020) Web of scholars: A scholar knowledge graph. In: Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval, pp 2153–2156
- Liu J, Nie H, Li S et al (2021) Tracing the pace of Covid-19 research: topic modeling and evolution. *Big Data Res* 25(100):236
- Liu J, Xia F, Wang L et al (2021) Shifu2: a network representation learning based model for advisor-advisee relationship mining. *IEEE Transact Knowl Data Eng* 33(4):1763–1777
- Liu J, Xia F, Feng X et al (2022) Deep graph learning for anomalous citation detection. *IEEE Transact Neural Netw Learn Syst* 33(6):2543–2557
- Liu L, Wang Y, Sinatra R et al (2018) Hot streaks in artistic, cultural, and scientific careers. *Nature* 559(7714):396
- Liu XZ, Fang H (2020) A comparison among citation-based journal indicators and their relative changes with time. *J Inf* 14(1):101–107
- Luo X, Liu Z, Shang M et al (2020) Highly-accurate community detection via pointwise mutual information-incorporated symmetric non-negative matrix factorization. *IEEE Transact Netw Sci Eng* 8(1):463–476
- McLaren CD, Bruner MW (2022) Citation network analysis. *Int Rev Sport Exerc Psychol* 15(1):179–198
- Mikolov T, Chen K, Corrado G, et al (2013) Efficient estimation of word representations in vector space. *Comput Sci* pp 1–12
- Mimouni M, Ratmansky M, Sacher Y et al (2016) Self-citation rate and impact factor in pediatrics. *Scientometrics* 108(3):1455–1460
- Moustafa K (2016) Aberration of the citation. *Account Res* 23(4):230–244
- Puigt M, Yahaya F, Delmaire G et al (2021) In situ calibration of cross-sensitive sensors in mobile sensor arrays using fast informed non-negative matrix factorization. *ICASSP 2021–2021 IEEE international conference on acoustics, speech and signal processing (ICASSP)*, IEEE, pp 3515–3519
- Ren J, Xia F, Chen X et al (2021) Matching algorithms: fundamentals, applications and challenges. *IEEE Transact Emerg Top Comput Intell* 5(3):332–350
- Rui H, Zhang G, Chen J (2019) Semi-supervised discriminant ISOMAP with application to visualization, image retrieval and classification. *Int J Mach Learn Cybern* 10(6):1269–1278
- Shiau WL, Wang X, Zheng F (2023) What are the trend and core knowledge of information security? a citation and co-citation analysis. *Inf Manag* 60(3):103,774
- Shu K, Wang S, Liu H (2019) Beyond news contents: The role of social context for fake news detection. In: 12th ACM international conference on web search and data mining, WSDM 2019, Association for Computing Machinery, Inc, pp 312–320
- Siler K, Larivière V (2022) Who games metrics and rankings? institutional niches and journal impact factor inflation. *Res Polic* 51(10):104,608
- Singh CK, Filho DV, Jolad S et al (2020) Evolution of interdependent co-authorship and citation networks. *Scientometrics* 125(1):385–404
- Siudem G, Żogała-Siudem B, Cena A et al (2020) Three dimensions of scientific impact. *Proc Natl Acad Sci* 117(25):13,896–13,900
- Sun K, Wang L, Xu B et al (2020) Network representation learning: from traditional feature learning to deep learning. *IEEE Access* 8:205,600–205,617
- Szomszor M, Pendlebury DA, Adams J (2020) How much is too much? the difference between research influence and self-citation excess. *Scientometrics* 123(2):1119–1147
- Tosyali A, Kim J, Choi J et al (2019) Regularized asymmetric nonnegative matrix factorization for clustering in directed networks. *Pattern Recogn Lett* 125:750–757

- Trujillo CM, Long TM (2018) Document co-citation analysis to enhance transdisciplinary research. *Sci Adv* 4(1):e1701,130
- Vercelli S, Pellicciari L, Croci A et al (2023) Self-citation behavior within the health allied professions' scientific sector in Italy: a bibliometric analysis. *Scientometrics* 128(2):1205–1217
- Wang F, Li T, Wang X et al (2011) Community discovery using nonnegative matrix factorization. *Data Min Knowl Discov* 22:493–521
- Wang MN, You ZH, Wang L et al (2021) Ldgrnmf: Lncrna-disease associations prediction based on graph regularized non-negative matrix factorization. *Neurocomputing* 424:236–245
- Wang S, Aggarwal C, Tang J, et al (2017) Attributed signed network embedding. In: Proceedings of the 2017 ACM on conference on information and knowledge management. association for computing machinery, New York, NY, USA, CIKM '17, pp 137–146
- Wang Y, Zhu L (2017) Research and implementation of svd in machine learning. In: 2017 IEEE/ACIS 16th international conference on computer and information science (ICIS), pp 471–475
- Wen F (2019) Study on the research evolution of nobel laureates 2018 based on self-citation network. *J Doc* 75(6):1416–1431
- Xia F, Wang W, Bekele TM et al (2017) Big scholarly data: a survey. *IEEE Transact Big Data* 3(1):18–35
- Xia F, Liu J, Ren J, et al (2020) Turing number: how far are you to am turing award? *ACM SIGWEB Newsletter* (Autumn), pp 1–8
- Xia F, Sun K, Yu S et al (2021) Graph learning: a survey. *IEEE Transact Artif Intell* 2(2):109–127
- Xia F, Yu S, Liu C et al (2021) Chief: clustering with higher-order motifs in big networks. *IEEE Transact Netw Sci Eng* 9(3):990–1005
- Yu S, Xia F, Sun Y et al (2020) Detecting outlier patterns with query-based artificially generated searching conditions. *IEEE Transact Comput Soc Syst* 8(1):134–147
- Zhao Y, Deng F, Pei J et al (2022) Progressive deep non-negative matrix factorization architecture with graph convolution-based basis image reorganization. *Pattern Recogn* 132(108):984
- Zhou L, Amadi U, Zhang D (2018) Is self-citation biased? An investigation via the lens of citation polarity, density, and location. *Inf Syst Front* 1(1):1–14
- Zhu X, Turney P, Lemire D et al (2015) Measuring academic influence: not all citations are equal. *J Assoc Inf Sci Technol* 66(2):408–427

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.