Information Processing and Management 59 (2022) 102996

Contents lists available at ScienceDirect

Information Processing and Management

journal homepage: www.elsevier.com/locate/ipm

Quantifying the structural and temporal characteristics of negative links in signed citation networks

Duoqi Song [a], Wenpei Wang [a], Ying Fan [a], Yanmeng Xing [a,b,*], An Zeng [a,**]

[a] School of Systems Science, Beijing Normal University, Beijing, 100875, PR China
[b] Networks, Data, And Society (NERDS) Research Group, IT University of Copenhagen, Copenhagen, 2300, Denmark

ARTICLE INFO

Keywords:
Negative citations
Citation behavior
Signed network
Natural-language processing

ABSTRACT

Although the citation relationships among papers can help in tracking and understanding the development of knowledge, few studies have noted that the content and sentiments of citations of a paper differ. Here, we use sentiment-labeled citation data to construct a directed signed citation network, in which an author may agree with or criticize the cited paper and these represent different ways of inheriting knowledge. The dataset we use consists of 9,038 papers in the field of Computational Linguistics, including 25,275 citations, with 20.8% positive citations, 8.6% negative citations and 70.6% neutral citations. We systematically quantify the structural patterns of negative citations, impact assortativity of involved papers, occurrence time distribution and consequences of receiving negative attention. Remarkably, we find that papers with different impacts have a similar probability of receiving negative citations, and highly cited papers tend to give negative citations to low-impact papers around but avoid giving negative citations to high-impact papers. Our research also reveals the random occurrence rules and colocation patterns of negative citation distribution. In addition, we show that, in the short term, around 60% of multiple negative citations is positively related to the impact of the cited paper while more than 80% are negatively related to the impact in the long run. Our findings explain the pattern by which negative citations occur and deepen the understanding of negative citations.

1. Introduction

Scientific publications document the accumulated knowledge of humans. The citation relationship between scientific publications is one of the key elements explaining the processes of the creation and propagation of knowledge and the evolution of the scientific citation system (Zeng et al., 2017). Citations also play an important social role since they may indicate the associations of the article with theoretical or methodological schools and potential epistemology (Hilário, Martínez-Vila, Grácio, & Wolfram, 2018). Moreover, the number of citations contributes significantly to measuring the scientific impact of scientists and articles (Redner, 1998), such as C_f (Radicchi, Fortunato, & Castellano, 2008), C_{10} (Sinatra, Wang, Deville, Song, & Barabási, 2016), h-index (Hirsch, 2005) and other indicators.

However, the reasons and sentiments underlying citations differ. Citation behavior is the author's selection of references based on the motivation of knowledge expression (Guo, Ying, & Milojevic, 2013). Citations may reflect positive (neutral) or negative attitudes from the scientific community (Hilário et al., 2018). Because of this, all the links of citation network, in which the nodes

* Corresponding author at: School of Systems Science, Beijing Normal University, Beijing, 100875, PR China.
** Corresponding author.
E-mail addresses: yanmengxing@mail.bnu.edu.cn (Y. Xing), anzeng@bnu.edu.cn (A. Zeng).

https://doi.org/10.1016/j.ipm.2022.102996
Received 1 April 2022; Received in revised form 3 June 2022; Accepted 12 June 2022
Available online 23 June 2022

represent papers and directed links represent the citations from one paper to another, could be signed by the corresponding citations' sentiment, and it is called a signed citation network. Here, the negative citation behavior specifically means that the author cite a reference negatively. Bordignon (2022) thought that a negative citation carrying a negative connotation toward to the cited paper, and they identified 3 functions on which to base the definition of negative citation: "to criticize", "to compare" and "to question" other papers. This implies that negative citations indicate the limitations, inconsistencies or flaws of research.

Different citation emotions play different roles. Positive and neutral citations reflect the heritage and development of the knowledge in the cited paper during the process of knowledge transmission, while negative citations reflect improvement or discontinuance. In addition, when evaluating the quality of a scientific finding, if the reason why papers are cited has been ignored, it is necessary to determine whether the number of citations can be used as a relevant metric to accurately assess the quality of its scientific results (Geras, Siudem, & Gagolewski, 2020; Kumar, 2016). Different citation sentiments may have distinct effects on papers, especially negative citations. The criticism and questioning from negative citations may affect papers' scope and development, even play a role in the entire field. Therefore, negative citations should be discussed separately. Negative views sometimes play a positive role in the evolution and development of science (Zeng et al., 2017). In Karl Popper's view of science, a discovery or theory can be defined as science if it is falsifiable (Popper, 1983). Negative citations are essentially scientific self-correction, in which new knowledge can be generated (Bordignon, 2020). For instance, since the Enlightenment, criticism has played a crucial role in the understanding of Western society and its fundamental institutions (Raffnsøe, 2017). Criticism can not only reveal vague concepts and methodological flaws, but also provide constructive ideas on how to avoid these flaws and strengthen research methods (Bordignon, 2020).

Negative citations obviously occupy an important position in the system of science, but related research is scarce. In recent years, with the rapid development of information technology and the maturity of natural language processing technology (Cambria & White, 2014), complete datasets and convenient basic support have enabled quantifying negative citations. Some scholars (Catalini, Lacetera, & Oettl, 2015; Geras et al., 2020; Kumar, 2016) have established signed citation networks for quantitative analysis through citation sentiment analysis (Bo & Lee, 2008; Yousif, Niu, Chambua and Khan, 2019). Nevertheless, the previous research has only described the phenomenon of negative citations, but not yet compared quantitatively the characteristics of negative citations with ordinary citations, and has not dug into the underlying mechanisms. In quantifying the effect of negative citations, the correlation between the number of negative citations papers received and their effects was not investigated. Furthermore, Teufel, Siddharthan, and Tidhar (2006) pointed out that many researchers are concerned that negative citations may pose a danger to themselves. When discussing the effects of negative citations, previous researches only focused on the cited paper, but did not explore its effect on the citing paper. The above limitations and deficiencies will be addressed in our study.

The objects of this paper are: (1) to study the characteristics of the papers involved in negative citations and understand the possible connection between negative citation behavior and citing papers' citation counts; (2) to explore the impact assortativity of negative citations, namely, which kind of references (the low-impact or the high-impact) are preferred to be cited by high-impact/low-impact papers; (3) to understand how negative citations are distributed and reveal the sequentially occurrence pattern of negative citations; (4) to quantify the effect of receiving negative citations on the impact of a paper.

First, we construct a signed network based on citation sentiment to identify negative links and explore which kinds of papers are more likely to cause controversy. We find that the probability of receiving negative citations is similar for papers with different levels of impact, although the data intuitively show that negative citations tend to attach to high-quality papers (Catalini et al., 2015; Geras et al., 2020). That is, whether a paper receives negative citations is not related to its own quality. We also point out that negatively citing other papers will not have a significant impact on citing papers' citations. Second, in examining the homogeneity of the signed citation network, we find that in negative citations, high-citation papers tend to negatively cite low-citation papers and avoid negatively citing high-citation papers. Third, we further analyze the timing and citation sequence of negative citations and find that negative citations received by a paper are randomly distributed throughout its citation trajectory. In other words, negative citations can appear anywhere with a similar probability in the citation sequence of a paper —a negative citation could be the first citation, a mid-term citation, or the last citation of the paper. In addition, we discover that most of negative citations are located close to each other. Finally, we explore the effect of negative citations on a paper. In the long run, consistent with intuition, negative citations are associated with the inhibition of the citation counts. However, interestingly, receiving multiple negative citations may correspond to the short-term promotion of the impact of the cited paper.

The remaining of this paper is organized as follows. Section 2 reviews related works on citation networks and negative citations. Section 3 presents the description of the empirical datasets and related methods. Section 4 reports the distribution features and temporal dynamics of negative citations, and their effects. Section 5 presents the discussion, conclusions and implications of the results.

2. Related work

Scientific knowledge diffusion is of great significance to the progress of science (Hassan, Visvizi, & Waheed, 2019). The academic publications record that knowledge, which needs to be spread to have an impact, and the spread of a paper can be approximated by the citations it receives. With the development of modern information technology, the number of papers continues to surge, and scientific publications data is stored, formatted and made public in large quantities, which promotes the development of citation relationship (Kayumovich, 2020). A citation network is a collection of quotations and cited relationships between papers (Mariani, Medo, & Lafond, 2019). With the help of complex network analysis, many structures and evolution rules in scientific publications data are revealed.

In the studies of citation networks, an important question is how to measure the quality of academic publications. Academic publications are not only an intuitive reflection of the research achievement of researchers, but also an important reference factor for measuring the level of scientists (Bai, Zhang, & Lee, 2019; Chan, Mixon, & Torgler, 2018). For evaluating the quality of a paper, the most straightforward and popular indicator by far is the number of citations of the paper (Didegah & Thelwall, 2013; Thelwall, 2016; Zhang, Xie, & Song, 2021). Citation counts are also the basis for other metrics (such as C_{10} (Sinatra et al., 2016), h-index (Hirsch, 2005), journal impact factor (Garfield, 1972)). In addition to the strength of the citation impact of scientific publications, there are also several studies focusing on the width of it. For example, Bu, Lu, Wu, Chen, and Huang (2021) understood the width of citations by the ego-centered citation networks (ECCN). Studying the dynamic evolution characteristics of citations to provide scholarly predictions and recommendations is another important issue in citation network research. Siudem, Żogała-Siudem, Cena, and Gagolewski (2020) summed up the three influencing factors of citation records —productivity, total impact and luck. By comparing three types of citation behavior trends (papers were classified into highly-cited, medium-cited, and lowly-cited based on the citation count), Yang and Liu (2022) explored the impact of citation behavior on the diffusion of disciplinary knowledge and knowledge structure, and quantified the knowledge diffusion process. Zhao and Feng (2022) predicted citation counts from the perspective of information cascade prediction by using the structure and temporal information of the citation network. However, these methods neglect the difference in the sentiment of citations. Specifically, each citation emotion has different role in spreading scientific achievements.

In recent years, with the maturity of information technologies such as sentiment analysis, natural language processing (NLP), and machine learning (Kunnath, Herrmannova, Pride, & Knoth, 2021; Yousif, Niu, Tarus and Ahmad, 2019), citation sentiment analysis can more detailedly describe how a paper is cited. In previous studies, there are two different methods used to identify the sentiment of the citation: unsupervised (Lexicon-based) methods and supervised (machine learning framework) methods (Athar, 2014). Unsupervised methods mostly judge sentiment based on sentiment lexicon, which is easier to understand and implement. For example: (Pilar Salas-Zárate et al., 2020) implemented example research work in the life sciences and biomedical fields using this approach. Kumar (2016) classified citation sentiments using a simple but powerful keyword-based technique, categorizing citation sentiment as positive, negative, or neutral. Lexicon-based methods are also used in the context of co-citation analysis (Yaghtin, Sotudeh, Nikseresht, & Mirzabeigi, 2021), for citation ranking and quantification of scientific quality. Supervised methods can classify unknown data (Athar, 2014), but this method requires a sufficient amount of similar and labeled data as a training set, so it is less applicable to new data (Gonçalves, Araújo, Benevenuto, & Cha, 2013).

The negative citation generally refers to citation that questions or criticizes aspect of the cited paper. Negative citations can correct errors and improve existing theories and techniques (Xu, Ding, & Lin, 2022). Bordignon (2020) pointed out that scientists can express doubts through academic debates and provide evidence for new claims. Thus, the study of negative citations helps to understand patterns of dissemination and development of scientific knowledge. However, the related research is scarce. Some scholars have discussed and conducted qualitative analyses of negative citations. For example, White (2001) believed that it takes more effort for researchers to make criticisms than perfunctory citations, so they prefer to minimize costs through frequent use of perfunctory citations (requiring relatively little context) and infrequent use of negative citations (requiring more context and thus more labor). Mertens, Herberz, Hahnel, and Brosch (2022) also pointed out a publication bias toward positive results in the literature over negative results. Savolainen (2021) asserted that criticism put forward by peers is particularly important, and that the nature of the criticism indicates the maturity of the research field. Similar to that of negative citations, the purpose of retraction is not to punish the authors of the work, but to ensure scientific integrity and correct errors, to prevent other scholars' exposure to them. Some scholars have studied the impact of retractions on articles and scholars. Most studies assert that retraction is considered harmful to an author's academic lifespan, which directly affects their reputation and perceived reliability (Noorden, 2011). Negative citations indicate a kind of disagreement (Tahamtan & Bornmann, 2018).

There is very little empirical evidence to show the quantity, most common areas, and consequences of such disagreements (Murray et al., 2019). Catalini et al. (2015) studied the incidence, distribution in time, and locations of negative citations within a paper and concluded that negative citations will slightly suppress the number of long-term citations in a paper. However, the correlation between the number of negative citations received and their impacts has not been investigated. It has been shown that negative citations follow a power-law distribution (Kumar, 2016), and signed citation networks follow the theory of weak balance (Davis, 1977). Geras et al. (2020) explored the impact of negative citations on quantitative scientific publications, showing that negative citations can serve as an indicator of the allocation of attention in the scientific community. Xu et al. (2022) analyzed the relationship between negative citations and the impact of cited papers by focusing on citation sentiment and citation type attributes. That study selected the data of papers on SVM and showed that papers with a certain negative citation rate have a greater impact. Undoubtedly, the negative citations have an important role in science, but few systematic studies have been devoted to their structural patterns, effects for citing or cited papers and other characteristics.

3. Data and methods

3.1. Data

The dataset that we use consists of 9038 papers from the field of Computational Linguistics, which are part of the ACL Anthology Reference Corpus (ACL ARC), including 25,275 citations (Kumar, 2016). Srijan Kumar uses a keyword-based technique to classify the citation sentiment, and the algorithm automatically divides the citations into three categories: positive, negative or neutral.

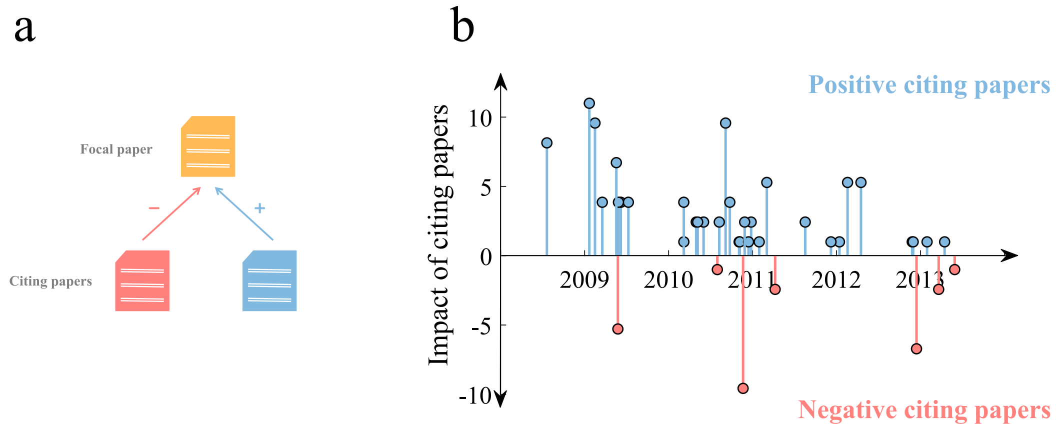

Fig. 1. Citation network and citation sequence. (a) Papers can be cited negatively (marked in red) or positively (marked in blue). (b) The time series displays the trajectory of a typical paper being cited. Each point represents the paper citing the focal paper, and the colors correspond to different sentiments. The heights are the impacts of citing papers. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 1
Basic characteristics of used data.

	Papers (citing/cited)	Citations
Positive	8789 (3814/6609)	23 098
Negative	2495 (1172/1458)	2177
Total	9038	25 275

Among the 25,275 citations, 5244 (20.8%) are classified as positive, 17,854 (70.6%) are neutral and 2177 (8.6%) are negative. The labeled dataset with the keywords is available for download on Kumar's project website.[1]

For neutral citations, we can assume that the authors agree with the points of reference if they do not express clear opposing views in their paper. We can even assume that they are quoting references neutrally out of approval of their conclusions. Therefore, it is reasonable that we artificially combine neutral and positive citations in the results as "positive" and still label negative citations "negative". Specifically, we group the citations in the obtained data into positive and negative categories (Fig. 1a), and the basic characteristics of the data used are presented in Table 1. Using the above dataset, we construct a signed network based on citation sentiment to identify negative links. In this network, the nodes represent papers and directed links represent the citations from one paper to another. The links have positive or negative sign that correspond to positive citation and negative citation, respectively. We also carry out the robustness check of combining neutral with positive citations in Supplementary Material. Positive citations indicate that the author supports, endorses and adopts the views in the reference, while negative citations indicate that the author cites the reference to point out deficiencies, defects or even errors in the cited paper, thus meaning that the views or methods in the original text are corrected or not adopted by that citation. Fig. 1b shows the citation trajectory, i.e., the sequence of citations of the focal paper, which we choose randomly, with different sentiments and impacts of citing papers.

3.2. Methods

Randomization of network topology. Here, we use the reshuffling method, which preserves the number of in-degree and out-degree for each paper as well as the sentiments attached to the edges. This method only changes the targets of the edge to achieve randomization of the network topology. We consider a random rewiring procedure in citation networks:

(i) Choose two edges with the same sign randomly;

(ii) Swap the ends of them and retain the sign of the edges.

Ensuring that the degree of each node remains unchanged, the rewiring procedure changes the nodes connected by edges while the degree distribution remains the same. By comparing the original and the random networks, we study the citing assortativity between papers with different citation impacts (Fig. 3).

Randomization of citation temporal trajectory. To study whether negative citation occurs randomly, regardless of publication time or order in the sequence of publications, we randomize the citation temporal information from two dimensions: citation time and sequence.

For randomization of citation time, we shuffle the cited year among all the citations of the same cited paper. The times of negative citations will be randomized through this operation, while the times of every citation that the target paper received can be

[1] Project website: http://cs.umd.edu/~srijan/citations/.

guaranteed after the publication of the cited paper. The time series of citations are randomized, while the total number of citations remains unchanged. The sign is moved with the corresponding edge during the randomization. The time distribution of the negative citations after randomization is shown by the gray line in Fig. 4c. To study the position of negative citations in the citation sequence, we rearrange the chronological order of citations of each focal paper. Based on the previous operation, the citations are sorted according to the new timestamp. As the data do not have detailed publication dates, we randomly sort the citations in the same year (Fig. 4d, gray line).

Relative changes in citations based on the benchmark. To quantify the effect of negative citations on focal papers, we denote a citation variation of negative citation i as δC_i:

$$\delta C_i = Y(T + n_1) - Y(T - n_2)$$

where the T means the negative citation i occurred at the Tth year of the focal paper's life cycle. $Y(n)$ is the yearly citation count of the cited paper at nth years. When calculating the short-term citation variation after papers are cited negatively, the situation is $n_1 = 1$ and $n_2 = 1$. The condition of $n_1 = 1$ and $n_2 = 10$ is for the analysis of long-term changes.

Generally, the annual citations of most papers tend to decrease after reaching the peak (Zhu, Wang, & Zhu, 2003). To eliminate the influence of the natural evolving rule of citations, for each negatively cited paper (focal paper), we match the control group consisting of other similar papers without any negative citation to set the benchmark.

The papers in the control group of each focal paper are selected by the following rules:

(i) Without any negative citation;

(ii) The life cycle is similar to that of the focal paper with an allowable error of ±2 years.

Thus, the benchmark for focal paper's citation variation is the mean of the citation variation of the papers in its control group over the corresponding time period.

We focus on the change in the number of citations of the focal paper relative to the benchmark before and after a given negative citation. For instance, a focal paper is cited negatively in the fifth year, and we study the relative changes in the negatively cited paper's impact in the short term. We compute the citation variation of the focal paper in the fourth and sixth year, and then in the same way, compute that of the fourth and sixth year of papers in the control group as benchmark. The relative variation is described by the difference between them. Therefore, we denote the relative variation as $\delta C_i'$:

$$\delta C_i' = \delta C_i - \overline{\delta C_{CG(i)}}$$

where $\overline{\delta C_{CG(i)}}$ is the benchmark, namely, the mean of citation variation of the papers in the control group of the focal paper i.

4. Results

Degree distribution and incidence of negative citations. The first question we ask is what the characteristics of the negatively cited papers are. To explore this problem, we compute the citation distribution of papers receiving negative citations and compare it with that of papers without negative in-citation. Here, we use "in-citation" to represent the citations received by a paper, namely, the in-degree of the node in the citation network. Similarly, the "out-citation" of a paper describes the references that a paper has. The total number of citations of a paper is usually used to approximate the impact of this paper. However, it is biased to the papers published earlier as they have longer time to accumulate citations. To correct this bias, we use a deformed C_{10} (Sinatra et al., 2016), the number of citations for a paper in ten years since it is firstly cited, to measure the impact of this paper. The larger C_{10} means the higher impact. Fig. 2a shows a log–log representation of the citation distribution of papers with negative citations (marked in red) and without negative citations (marked in blue). The fitting result suggests that both citation distributions are well described by fat tail such that $N(K) \propto k^{-\alpha}$ with $\alpha = -1.61$ and $\alpha = -2.59$, respectively. Therefore, papers cited negatively have larger exponents, which implies that their impact is higher compared with other articles.

To further explore the incidence of negative citations in papers with different impacts, we classify all papers cited negatively according to C_{10} and calculate the average number of negative citations of each type of paper (Fig. 2b, inset). We take the logarithmic coordinate, 2^k, to ensure the uniformity of the distribution of papers in each interval. We find that the average number of negative citations and C_{10} is a straight line with a positive slope (Fig. 2b, inset), indicating that the more influential an article is, the more negative citations it receives. This finding is consistent with previous studies (Catalini et al., 2015; Zhao & Feng, 2022). In particular, we calculate the proportion of negative in-citations in in-citation and the correlation with C_{10}, as shown in Fig. 2b. The proportion is stable with the increase of C_{10}, which means that the probability of receiving negative citations is almost the same in papers with different impacts. Why do articles with high impact appear to receive more negative citations? This occurs because of the large cardinality of citations.

We further ask, which types of papers tend to criticize (negatively cite) others. Researching the degree distribution of citing papers with and without negative out-citation, we find that both follow fat-tail distributions with exponents of −1.49 and −2.02, respectively (Fig. 2c). In contrast to the distribution of cited papers, there is no significant difference in the C_{10} distribution of citing papers. We also consider whether a paper can receive more attention if it cites others negatively. We investigate the correlation between the negative out-citations and C_{10} following the same calculation as in the cited papers. The result indicates that papers with different impacts have similar negative out-citations (Fig. 2d, inset). The results in Fig. 2d indicate that the probability of citing others negatively is not associated with their impact.

Impact assortativity in negative citations. We next study the impact assortativity of negative citations, that is, which kind of references (the low-impact or the high-impact) are preferred to be cited by high-impact/low-impact papers. Here we also use C_{10}

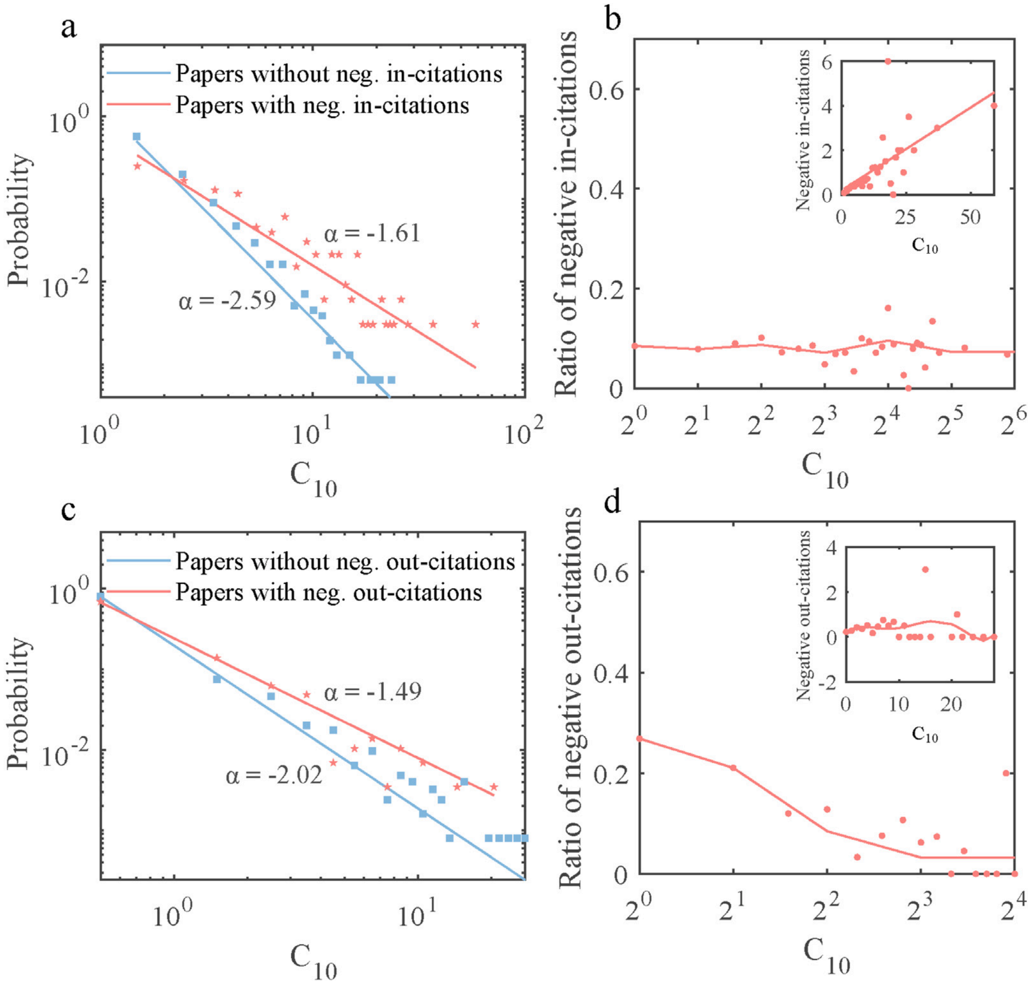

Fig. 2. Degree distributions and ratio of negative edges in negatively cited and positively cited papers. (a) Log–log representation of the distribution of citations of cited papers. The exponent α represents the slope. (b) Ratio of negative in-citations in papers with different C_{10}. Inset: correlation between the number of papers' negative in-citations and C_{10}. (c) Log–log representation of the distribution of citations of citing papers. (d) Ratio of negative out-citations in papers with different C_{10}. Inset: Correlation between the C_{10} and the number of papers' negative out-citations.

to measure the impact of the papers, the same as above, the larger C_{10} means the higher impact. First, we group papers into two categories by impact, i.e., high-impact and low-impact. Therefore, there are four different matches between papers and references. The citation relationships are divided into four types: high-impact papers cite high-impact references (H–H), low-impact papers cite high-impact references (L–H), high-impact papers cite low-impact references (H–L), and low-impact papers cite low-impact references (L–L). For simplicity, because citations follow a fat-tail distribution, we differentiate citing papers as high-impact and low-impact by a threshold of 2^3 (C_{10}), and the threshold can result in close sample sizes in the two groups. Similarly, the threshold for the cited papers is 2^4. Subsequently, we count the number of different kinds of edges in the citation network and compare it with the topology randomization network (see methods: randomization of network topology).

By comparing the numbers of four kinds of edges in a real network with those in a random network, we can obtain a ratio, i.e., relative impact assortativity. The ratio greater (less) than 1 indicates the high (low) citation preference between the citing papers and cited papers with corresponding impacts. In other words, a ratio greater than 1 means that there is a greater probability of observing citations between corresponding papers than the topology randomization network. Fig. 3 demonstrates the connection preference between high-impact and low-impact papers in negative and positive edges, respectively. Regardless of negative or positive citations, we find that high-impact papers tend to cite high-impact papers (red bubbles $\geq$ 1) rather than low-impact papers (green bubbles $\leq$ 1). However, compared with positive citations, high-impact papers have a lower preference to criticize (cite negatively) high-impact papers, as the ratio of the red bubble in Fig. 3a is smaller than that of the red bubble in Fig. 3b. Instead, since the green bubble in Fig. 3a is bigger than that in Fig. 3b, high-impact papers in negative concatenations have a

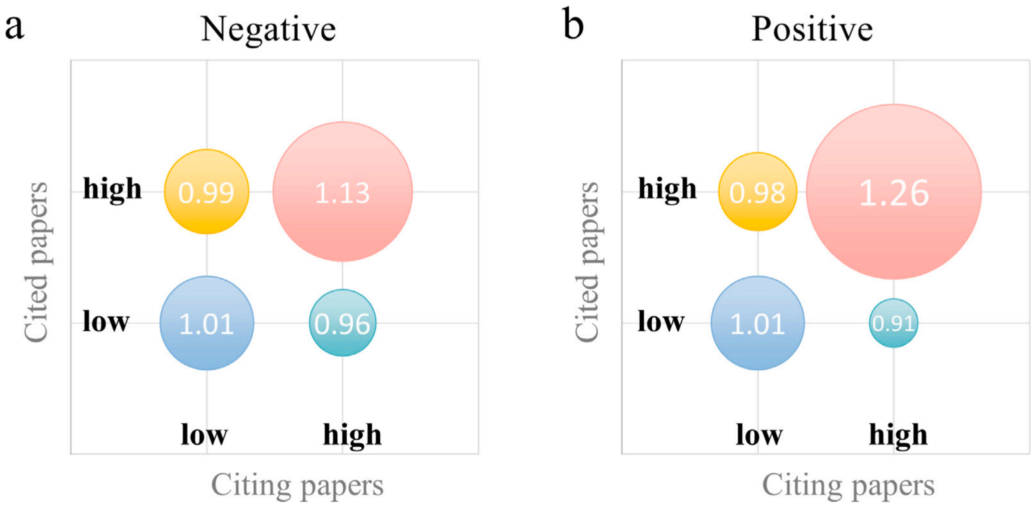

Fig. 3. Impact assortativity of citing and cited papers in two sentiments. The citation relationships are divided into four types: H–H (red bubble), H–L (green bubble), L–H (yellow bubble), and L–L (blue bubble). The sizes of the bubbles are encoded by the ratio. A ratio less than, equal to or greater than 1 indicates that the citation preference between the corresponding papers is low, similar to the random network, and high, respectively. (a) Impact assortativity of negative citations. (b) Impact assortativity of positive citations. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

higher preference to cite low-impact papers. We do not find that low-impact papers have different preferences in citing other papers compared to random networks (the ratios of yellow and blue bubbles are approximately equal to 1).

Random rules of negative citations. Understanding the distribution characteristics and potential generation pattern of negative citations is important. Fig. 4a shows that the distribution of negative citations is similar to that of positive citations in different periods of a paper's life cycle. Both kinds of citations are mostly accumulated in the first decade since the paper is cited for the first time.

To further study when a paper receives its negative citations, we specifically measure the distribution of negative citations on the citation trajectory. Hereafter, to characterize the citation sequence of each article more fairly, we only discuss citation records within 10 years after its publication to control the citation length of papers. We measure the probability $P(t^-)$ that every negative citation is generated at time t^- after a paper is firstly cited (Fig. 4c). With the increase in time, yearly negative citations gradually decrease. More than 40% of negative citations are concentrated in the firstly cited year, because there are some articles that have only one citation and it is negative. Therefore, we also show the distribution of citations with more than 5 negative citations in the inset. The drop in $P(t^-)$ suggests that it is less likely that a negative citation will come late in a paper's lifespan. To understand the origin of this pattern, we shuffle the cited year among all citations of the same cited paper, preserving the citation sentiment and only randomizing the time (see methods: randomization of citation temporal trajectory). The results show that negative citations are indistinguishable from the random result (Fig. 4c), indicating that variations in $P(t^-)$ are not due to specific impact sequences or other features but are entirely explained by variations in yearly citation count throughout a paper's lifespan.

These results prompt us to explore the position N^- of negative citations in the sequence of N citations by measuring $P(N^-/N)$ (Sinatra et al., 2016). We sort all the citations of a paper by time of occurrence (Fig. 4b), and then define N^-/N as the relative position where a negative citation occurs. A small N^-/N means that a negative citation appears early; otherwise, it appears later. Fig. 4d shows an interesting finding that the cumulative distribution function $P(\geq N^-/N)$ decreases linearly, which means that the negative citations can be anywhere in the sequence of citations with a similar probability. This conclusion could also be supported by the flat $P(N^-/N)$ (Fig. 4d, inset). We also randomize the citation sequence, and the distribution of N^-/N is not changed significantly (see methods: randomization of citation temporal trajectory). Ultimately, we draw a rather unexpected conclusion: negative citation is randomly distributed over a paper's lifespan, suggesting a temporal random rule in negative citation.

Temporal colocation patterns. Another interesting question is whether negative citations come in streaks over a paper's lifespan. To answer this question, we measure the correlation among the times of every negative citation, and compare that with the positive citations. First, for each negative citation, which serves as a center (so we call it a "central citation"), we sort all the negative citations (called as "circumjacent citations") pointing to the same cited paper in the five years before and after it. Then, we observe the distributions of the times of these citations. Fig. 5a shows the probability distribution of the time lag between the "central citations" and their "circumjacent citations". The zero-tick area indicates that both the central citations and the circumjacent citations occurred in the same year, and we find that nearly 60% of the citations are distributed in the zero-tick area. Nearly 20% of the circumjacent citations came in the previous year (in the -1-tick area) of the circumjacent citations, and it is similar in the following year (in the 1-tick area). Last, we study the positive citations in the same way (Fig. 5b). The temporal location distribution of positive citations is relatively even. These results support that negative citations tend to be located close to each other. In other words, negative citations are clustered over a paper's lifespan.

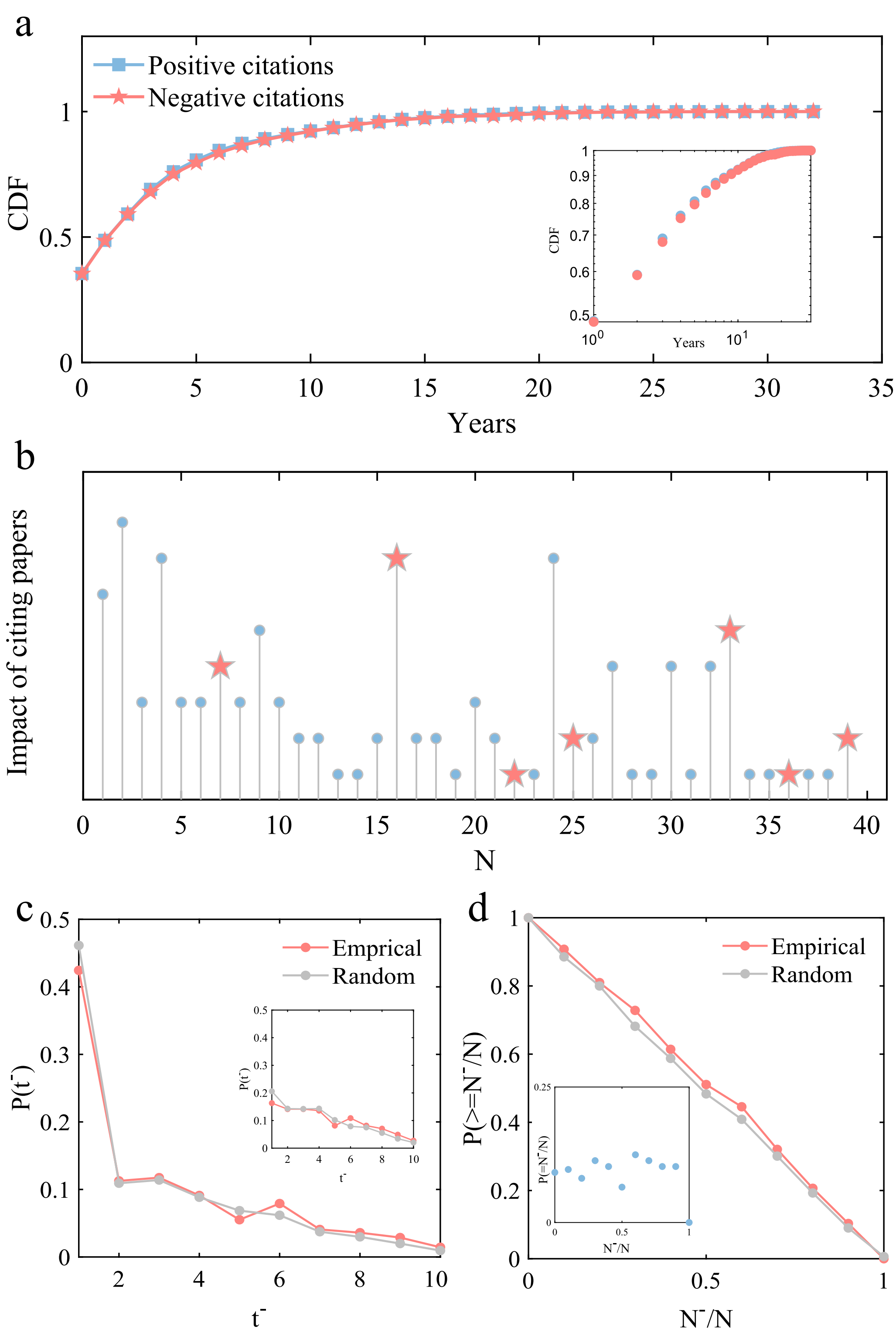

Fig. 4. Patterns of negative citations during a paper's lifespan. (a) Cumulative distribution of negative and positive citations. The horizontal axis indicates the number of years after the paper was cited firstly. CDF, cumulative distribution function. (b) Citation sequence of a focal paper. The horizontal axis indicates the rank in the citation sequence, and each vertical line corresponds to a citation. The height of each line is related to the impact of the citing papers. Blue circles correspond to positive citations, and red asterisks correspond to negative citations. (c) Time distribution of the negative citations for all the papers with at least one negative citation. The inset shows the distribution for papers with more than 5 negative citations. Both distributions are similar to the random case. (d) Cumulative distribution $P(\leq N^{-}/N)$ for papers with more than 5 negative citations, where N^{-}/N denotes the order of the negative citations in a paper's lifespan. The cumulative distribution is a straight line with slope 1, which is supported by the inset, where the probability of different positions in the sequence is presented.

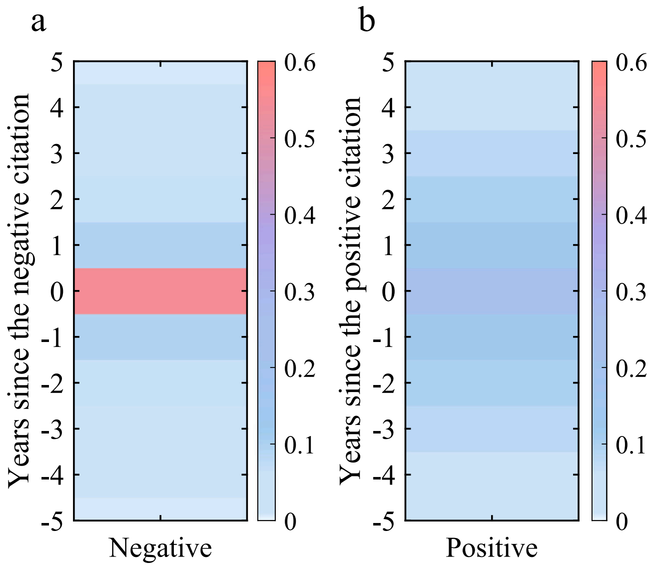

Fig. 5. Colocation patterns in negative citations. By centering on a citation (we call it the focal citation), we determine all the citations during 5 years forward and afterward, which have the same sentiment and are from the same cited paper as the central citation, to calculate the 11-year distribution of the circumjacent citations. Color coded, measures the mean of probability. (a) Negative citations; (b) Positive citations. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Relation between negative citations and impact of focal papers. Finally, we investigate the effect of receiving negative citations on a paper. In this part, we distinguish effects as short-term and long-term effects of a single negative citation and multiple citations, which was rarely done in previous related studies.

We quantify the effect of negative citations on papers by the citation variation, namely, the change in the yearly citation count before and after receiving negative citations. Short-term effects are characterized by the yearly citation variation between the 1st year after receiving negative citations and the 1st year before, and the long-term effect is characterized by that between the 10th year after and the 1st year before. However, the observed effect of negative citation on focal paper's citations could be a natural evolving rule of citations rather than the consequence of negative citation. To avoid the possible effect of the natural evolving rule of citations, we set the benchmark by matching the control group for each focal paper, in which all papers have the similar life cycle with the focal paper and have never received negative citation. The benchmark for focal paper's citation variation is the mean of the citation variation of the papers in its control group over the corresponding time period (see methods: relative changes in citations based on the benchmark).

Fig. 6 shows the distribution of the relative changes of citation $\delta C_i'$. The "+" mark in Fig. 6 means $\delta C_i' \geq 0$, and the negative citations belonging to this category correspond to promotion of the impact of the cited paper. Namely, compared with the benchmark, the focal paper's yearly citation count increased after receiving the negative citation belonging to this category. Similarly, the "–" mark corresponds to inhibition —the focal paper's yearly citation count decreased compared with the benchmark. We find that the proportions of inhibition and promotion effects brought by the single negative citation are similar form Fig. 6a. This means that a single negative citation has little effect on a paper's number of citations. Remarkably, nearly 60% of multiple citations lead to a promoting effect on the corresponding cited paper. The long-term effect is shown in Fig. 6b. We find that most of the single and multiple negative citations belong to "–", so both single and multiple negative citations have a "penalty". This result is consistent with previous research (Catalini et al., 2015), especially the multiple negative citations.

5. Conclusion and implications

5.1. Discussion and conclusions

In summary, we find that papers with different impacts have a similar probability of receiving negative citations. Our findings reveal the random rules of negative citations in the citation system of science, that is, the time when negative citations occur are random. Note that, our results do not mean that authors were randomly targeting any other study to make a negative citation. We also identify the colocation patterns and study the effect of receiving negative citations on a paper. These findings suggest the important role of negative citations and provide a new perspective on the study of scientific citation systems.

First, we explore the features of citing and cited papers with negative citations. Previous studies found that the citation value of papers receiving negative citations are somewhat larger than those without receiving negative citations. (Catalini et al., 2015). However, we find that this phenomenon is caused by the randomness of the critics, which means that each citation has a similar

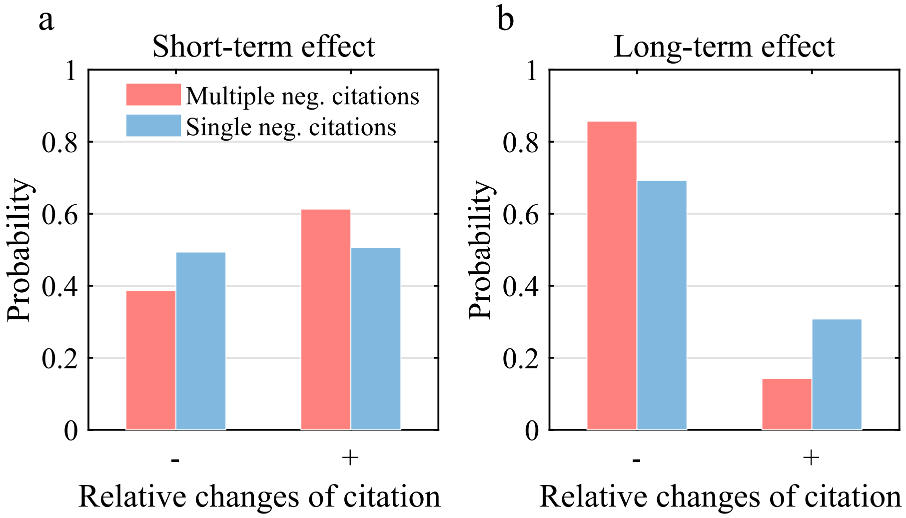

Fig. 6. Histogram of the relative citation changes from single negative citations and multiple citations. "+" corresponds to a promoting effect, that is, the relative citation variation $\delta C_i'$ is greater than 0, and "–" corresponds to inhibition. (a) Short-term effect; (b) Long-term effect.

probability of being negative. Naturally, papers receiving more citations will have more negative citations, rather than high-impact papers are more likely to be cited negatively. We also conclude that negative citation behavior of a paper and its citation counts are unrelated. That means that questioning peers' papers will not bring a crisis to own impact. Therefore, there is no need for scholars to feel pressured about giving negative citations. On the contrary, negative citations may promote the theoretical and technical development of the subject, it is worthy of being encouraged.

We then investigate the impact assortativity between the citing papers and cited papers with negative citations. We find that high-cited papers tend to cite high-cited papers, while low-cited papers have no preference in giving negative citations. This conclusion is consistent with our intuition. This may be because the high-cited article itself is more rigorous, and it costs more to determine its shortcomings. We then study the temporal pattern of negative citations and find they display a random rule. Moreover, we identify the colocation pattern in negative citations. These results indicate that negative citations tend to occur together, although the colocation positions are random. A similar colocation pattern has also been found in hit works within a career (Liu et al., 2018).

Finally, we assess the effect of receiving negative citations on a paper's impact and find that most of the multiple citations correspond to the promotion of the impact of the cited paper in the short run. That is, it is not necessarily bad for scholars to have their papers cited negatively. On the contrary, receiving negative citations may mean that the paper's impact will increase in a short period, and the controversy from negative citations can attract attention when they first occur. However, the "plenty" effect of negative citations requires a long time to function. This may be because negative citations go unnoticed, so the information that they carry takes a long time to spread. The negative evaluations of citations can provide scholars with critical understanding of the required theoretical basis (Bordignon, 2020), so as to develop the corresponding theory in a more comprehensive way.

Our study sheds light on the special role and underlying mechanisms of negative citations in the scientific citation system. We also point out the need to specifically introduce negative citations in research such as quantitative scientific outcomes, scholarly predictions and recommendations. Our approach has the potential to suggest to scholars how to better understand the meaning of negative citations and highlight the new dimensions of citation analysis.

This study also has limitations. In order to check the robustness of our results (see supplementary material), we have conducted a different set of analysis of dividing the citations into three categories (positive, negative, and neutral), and we found that the conclusions are consistent with the two-category setting (positive and negative). In fact, due to the small size of the dataset, the robustness of the conclusions can be further verified if more sufficient data can be collected. Moreover, the study is limited by the fact that we concern only the field of Computational Linguistics, and the generalizability of findings may be low. A more comprehensive comparison across disciplines is necessary before results can be generalized. If there are more suitable data sets in the future, we hope that the above conclusions can be tested. The time information of the data is in years and is rough, so we have to determine the sequence information of citations occurred in the same year in a random way, which may quantitatively influence the results. Moreover, we only study correlation analysis, and no related causality studies are conducted, which can be addressed in subsequent works. Further research could build a model to reveal in more details the rule of negative citation.

5.2. Implications

The theoretical and practical implications of this paper are as follows:

• Our research reveals the pattern by which negative citations occur and suggest that negative citations may play a special role in science. We construct a signed network based on the citation sentiment to identify negative links, and reveal the temporal random rules and colocation patterns of negative citations. We show that when quantifying the impact of a paper or a scientist with citations, it is necessary to develop a ranking method based on signed citation networks. This means that the existing evaluation methods still have room for improvement, so our research may promote a new research direction of scientific evaluation systems.

• Our research can help in understanding the characteristics of papers involved in negative citations, and we also quantify the effect of receiving negative citations on a paper. Our finding shows that the papers giving negative citations have no special characteristics with other citing papers. This means that negatively citing papers will not have a significant impact on their citations. Therefore, it is unnecessary to feel pressured when pointing out limitations in peers' papers. In the other hand, we point out that the large cardinality of citations leads to the accumulation of negative citations for high-impact papers, rather than citation preferences for high-impact papers. Actually, papers with different impacts have a similar probability of receiving negative citations. And receiving negative citations is not a worse thing for a paper in the short term, though it may bring "penalty" for a long term. These findings help people deepen the understanding of negative citations.

CRediT authorship contribution statement

Duoqi Song: Methodology, Software, Formal analysis, Writing – original draft. **Wenpei Wang:** Writing – original draft, Writing – review & editing. **Ying Fan:** Conceptualization, Supervision. **Yanmeng Xing:** Conceptualization, Methodology, Formal analysis, Writing – review & editing. **An Zeng:** Conceptualization, Methodology, Writing – review & editing.

Declaration of competing 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.

Acknowledgments

The authors are grateful to Prof. Srijan Kumar for his data. This work was sponsored in part by National Natural Science Foundation of China (71731002).

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.ipm.2022.102996.

References

Athar, A. (2014). *Sentiment analysis of scientific citations*: *Technical report*, University of Cambridge, Computer Laboratory.

Bai, X., Zhang, F., & Lee, I. (2019). Predicting the citations of scholarly paper. *Journal of Informetrics*, *13*(1), 407–418.

Bo, P., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends® in Information Retrieval*, *2*(1–2), 1–135.

Bordignon, F. (2020). Self-correction of science: a comparative study of negative citations and post-publication peer review. *Scientometrics*, (1).

Bordignon, F. (2022). Critical citations in knowledge construction and citation analysis: from paradox to definition. *Scientometrics*, *127*.

Bu, Y., Lu, W., Wu, Y., Chen, H., & Huang, Y. (2021). How wide is the citation impact of scientific publications? A cross-discipline and large-scale analysis. *Information Processing & Management*, *58*(1), Article 102429.

Cambria, E., & White, B. (2014). Jumping NLP curves: A review of natural language processing research [Review Article]. *IEEE Computational Intelligence Magazine*, *9*(2), 48–57.

Catalini, C., Lacetera, N., & Oettl, A. (2015). The incidence and role of negative citations in science. *Proceedings of the National Academy of Sciences*.

Chan, H. F., Mixon, F. G., & Torgler, B. (2018). Relation of early career performance and recognition to the probability of winning the nobel prize in economics. *Scientometrics*.

Davis, J. A. (1977). Clustering and structural balance in graphs. *Social Networks*, *20*(2), 27–33.

Didegah, F., & Thelwall, M. (2013). Which factors help authors produce the highest impact research? Collaboration, journal and document properties. *Journal of Informetrics*, *7*(4), 861–873.

Garfield, E. (1972). Citation analysis as a tool in journal evaluation. *Science*, *178*(4060), 471–479.

Geras, A., Siudem, G., & Gagolewski, M. (2020). Should we introduce a dislike button for academic articles? *Journal of the Association for Information Science and Technology*, *71*(2), 221–229.

Gonçalves, P., Araújo, M., Benevenuto, F., & Cha, M. (2013). Comparing and combining sentiment analysis methods. In *Proceedings of the First ACM conference on online social networks* (pp. 27–38).

Guo, Z., Ying, D., & Milojevic, S. (2013). Citation content analysis (cca): A framework for syntactic and semantic analysis of citation content. *Journal of the Association for Information Science & Technology*, *64*(7), 1490–1503.

Hassan, S.-U., Visvizi, A., & Waheed, H. (2019). The 'who'and the 'what'in international migration research: Data-driven analysis of Scopus-indexed scientific literature. *Behaviour & Information Technology*, *38*(9), 924–939.

Hilário, C., Martínez-Vila, D., Grácio, M., & Wolfram, D. (2018). Authorship in science: A critical analysis from a Foucauldian perspective. *Research Evaluation*, *27*.

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Ences of the United States of America*, *102*(46), 16569–16572.

Kayumovich, K. O. (2020). *Particular qualities use of social media in digital tourism, Vol. 28*. Gwalior Management Academy.

Kumar, S. (2016). Structure and dynamics of signed citation networks. In *Proceedings of the 25th international conference companion on world wide web* (pp. 63–64).

Kunnath, S. N., Herrmannova, D., Pride, D., & Knoth, P. (2021). A meta-analysis of semantic classification of citations. *Quantitative Science Studies*, 1–24.

Liu, L., Wang, Y., Sinatra, R., Giles, C. L., Song, C., & Wang, D. (2018). Hot streaks in artistic, cultural, and scientific careers. *Nature*, *559*(7714), 396–399.

Mariani, M. S., Medo, M., & Lafond, F. (2019). Early identification of important patents: Design and validation of citation network metrics. *Technological Forecasting and Social Change*, *146*, 644–654.
Mertens, S., Herberz, M., Hahnel, U. J., & Brosch, T. (2022). The effectiveness of nudging: A meta-analysis of choice architecture interventions across behavioral domains. *Proceedings of the National Academy of Sciences*, *119*(1).
Murray, D. S., Lamers, W., Boyack, K. W., Larivière, V., Sugimoto, C. R., van Eck, N. J., et al. (2019). Measuring disagreement in science. In *ISSI* (pp. 2370–2375).
Noorden, R. V. (2011). The trouble with retractions. *Nature*, *478*(7367), p.26–28.
Pilar Salas-Zárate, M. d., Alor-Hernández, G., García-Alcaraz, J. L., Colombo-Mendoza, L. O., Paredes-Valverde, M. A., & Sánchez-Cervantes, J. L. (2020). A sentiment analysis method for analyzing users opinions about drugs for chronic diseases. In *Data analysis and optimization for engineering and computing problems* (pp. 217–228). Springer.
Popper, K. R. (1983). The logic of scientific discovery.
Radicchi, F., Fortunato, S., & Castellano, C. (2008). Universality of citation distributions: towards an objective measure of scientific impact.
Raffnsøe, S. (2017). What is critique? Critical turns in the age of criticism. outlines. *Critical Practice Studies*, *18*(1), 28–60.
Redner, S. (1998). How popular is your paper? An empirical study of the citation distribution. *The European Physical Journal B*, *4*(2), 131–134.
Savolainen, R. (2021). Levels of critique in models and concepts of human information behaviour research. *Aslib Journal of Information Management*.
Sinatra, R., Wang, D., Deville, P., Song, C., & Barabási, A.-L. (2016). Quantifying the evolution of individual scientific impact. *Science*, *354*(6312), aaf5239.
Siudem, G., Żogała-Siudem, B., Cena, A., & Gagolewski, M. (2020). Three dimensions of scientific impact. *Proceedings of the National Academy of Sciences*, *117*(25), 13896–13900.
Tahamtan, I., & Bornmann, L. (2018). Core elements in the process of citing publications: A conceptual overview of the literature. *Journal of Informetrics*, *12*(1), 203–216.
Teufel, S., Siddharthan, A., & Tidhar, D. (2006). Automatic classification of citation function. In *Proceedings of the 2006 conference on empirical methods in natural language processing* (pp. 103–110).
Thelwall, M. (2016). Citation count distributions for large monodisciplinary journals. *Journal of Informetrics*, *10*(3), 863–874.
White, H. D. (2001). Authors as citers over time. *Journal of the Association for Information Science & Technology*, *52*(2), 87–108.
Xu, L., Ding, K., & Lin, Y. (2022). Do negative citations reduce the impact of cited papers? *Scientometrics*, *127*(2), 1161–1186.
Yaghtin, M., Sotudeh, H., Nikseresht, A., & Mirzabeigi, M. (2021). Modeling the co-citation dependence on semantic layers of co-cited documents. *Online Information Review*.
Yang, J., & Liu, Z. (2022). The effect of citation behaviour on knowledge diffusion and intellectual structure. *Journal of Informetrics*, *16*(1), Article 101225.
Yousif, A., Niu, Z., Chambua, J., & Khan, Z. Y. (2019). Multi-task learning model based on recurrent convolutional neural networks for citation sentiment and purpose classification. *Neurocomputing*, *335*(MAR.28), 195–205.
Yousif, A., Niu, Z., Tarus, J. K., & Ahmad, A. (2019). A survey on sentiment analysis of scientific citations. *Artificial Intelligence Review*, *52*(3), 1805–1838.
Zeng, A., Shen, Z., Zhou, J., Wu, J., Fan, Y., Wang, Y., et al. (2017). The science of science: From the perspective of complex systems. *Physics Reports*, *714*, 1–73.
Zhang, X., Xie, Q., & Song, M. (2021). Measuring the impact of novelty, bibliometric, and academic-network factors on citation count using a neural network. *Journal of Informetrics*, *15*(2), Article 101140.
Zhao, Q., & Feng, X. (2022). Utilizing citation network structure to predict paper citation counts: A deep learning approach. *Journal of Informetrics*, *16*(1), Article 101235.
Zhu, H., Wang, X., & Zhu, J. Y. (2003). Effect of aging on network structure. *Physical Review E*.