A Simulation Model Based on Artificial System to Analyze Correlation between Journal Impact Factor and Article Quality

Ke Dong*
School of Computer Science and Information
Hefei University of Technology
Hefei, China
e-mail: coliapaston@163.com
*corresponding author

Yiran Wang
School of Network and New Media
Hebei University
Baoding, China
e-mail: ronnieyiranwang@163.com

Wen Fan
School of Engineering
University of Manchester
Manchester, UK
e-mail: fanwen2020@163.com

2021 International Wireless Communications and Mobile Computing (IWCMC) | 978-1-7281-8616-0/21/$31.00 ©2021 IEEE | DOI: 10.1109/IWCMC51323.2021.9498612

***Abstract*—As the most significant evaluation indicator in Journal Citation Report, Journal Impact Factor (JIF) plays a very important role in the process of assessing academic journals. However, the related study on JIF is only focused on bibliometrics and scientometrics. To remedy this issue, we proposed a social computing simulation model to analyze the correlation between JIF and quality of articles. In the model, a social problem can be evaluated from the perspective of simulation analysis and complex system, in which certain factors can be denoted by some parameters and agents. As an agent-based model, the JIF system consists of three agents (Author, Article and Journal). The simulation result and experimental analysis strongly demonstrated that the calculation of JIF can be modeled by the proposed model. Furthermore, it can be found that JIF can be reflected by the quality of the published articles.**

Keywords—JIF, Social Computing, Simulation Model, Correlation Analysis

I. INTRODUCTION

Journal Impact Factor (JIF) is the most famous important evaluation indicator in Journal Citation Reports (JCR) [1]. By the ratio of the citations received, JIF plays a very important role in the process of academic evaluation. JIF, for example, is commonly used to evaluate the scientific performance of individual scholars according to their publications. The impact of a journal can be measure by a citation indicator is firstly proposed by Garfield in 1955 [2]. In 1963, the academic level difference can be evaluated by JIF which is also introduced by Garfield [3]. With the further development and application, JIF is considered to have great value in library journal collection management [4]. However, the related study on JIF is only focused on bibliometrics [5] and scientometrics [6]. Such works tend to develop mathematical models from the perspective of metrology to explore the performance and reliability of JIF. Over the past decades, with the extensive application of the computer simulation technology [7, 8], an increasing number of scholars tend to utilize computer simulation technology to address social problems. Such as Tan et. al. proposed an agent-based simulation model to study analyze explore the relationship between JIF and the average review period [9], and Allesina introduced an agent-based model to analyze the peer review system [10]. As a method to deal with the network model of complex social systems, social computing was firstly proposed by Wang in 2004 [11]. While the artificial system method based on artificial society is used to model the complex system involved in social computing, it is considered the model as a possibility of a real complex system instead of regarding the degree of approximation of real complex system as the only criterion. Through extensive simulation experiments, artificial systems are easily to be manipulated [12]. Owing to the various features in real complex system can be quantified by artificial system, artificial system can be usually used to model various social problems such as urban rail transit [13], information security [14-16] and other fields [17, 18]. Based on social computing, this paper developed a simulation model for a comprehensive academic system which is composed of authors-papers-journals.

To explore the correlation between JIF and the quality of published article, this paper proposed an academic network which consists three subnetworks (journal network; article network; author network).

The key contributions of this work can be summarized as follow:

- This paper introduced social computing into the study of journal impact factor.
- A social simulation model is proposed, which can be used to analyze the journal impact factor.
- Through extensive simulation experiments, the correlation between JIF and quality of papers can be verified.

The rest of this paper is organized as follows. Sect. II introduces the structure of the simulation model in detail. Sect. III mainly analyzes the relationship between journal impact factor and the quality of journal papers and discusses the experimental results. Finally, the conclusion of this paper is presented in Sect IV.

II. MODEL FORMULATION

To analyze the correlation between JIF and the quality of articles, this paper utilized the major approach of artificial system modeling, and proposed an agent-based JIF system. As a bottom-up method. Agent-based model can decompose and normalize all kinds of interactive behaviors in the agent network. It mainly consists of three parts (Agent, Environment and Rule). Agents play role in the individual of

an artificial system. Behaviors of agents can be altered by changing over time, communication and the outside world. Environments are the conditions that affect the behavior and development of its own agents. Besides rules are the principle of agents, the environment itself, and the interaction between intelligent agents and the environment [19]. Then, we will explain the model of this paper from three aspects: agent, environment and rules in detail.

A. Agent-Based Model

By analyzing the main participants about JIF system, it can be found that the articles written by authors with higher reputation would hold higher impact [20]. In addition, Eghe et al. proposed that articles with higher impact would obtain more citations [21]. The citations of the articles decide the JIF of a journal. Taken together, three agents can be defined as follows:

- Author: Authors are the producer of the articles. Assume that the impact of an author is a constant value.
- Article: There exists a great level differences of the articles because the differences of different authors are greater. Articles can be cited from each other. Supposing that the citation rate is a constant value when the article is written by designated authors.
- Journal: Journal is a kind of agents used to publish articles. A journal can publish multiple articles. This paper assumes that Journal are published yearly.

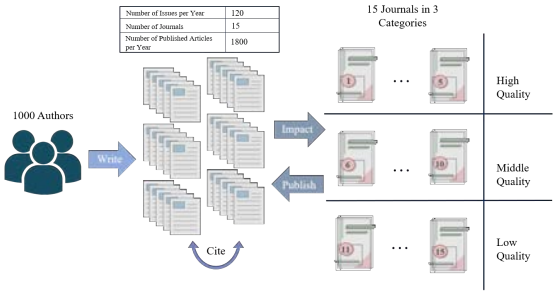


Fig. 1. The agent-based JIF system. The model contains 1000 author agents. 1440 article agents and 3 categories of 15 journal agents are generated per year. The interaction (write or publish) between agents is represented by the arrows.

It should be noted that the relationship between the authors and the papers is not one-to-one. One author may participate in the publication of multiple articles as different orders. An article can be co-published by 1-5 authors while each author has the probability to participate in the writing of any paper. There are five kinds of correspondence between the articles and the author: $1v1, 1v2, 1v3, 1v4, 1v5$.

There exists a closely correlation between the inherent states of agent models and simulation behavior. For the author agent, *Author Score* is an evaluation indicator of each author's own capability and reputation. The higher the score of authors, the higher the quality of their written articles, the more likely the author is to write high quality articles. For article agents, *Year* represents the year in which the paper was published. *Article Score* is a kind of evaluation indicator which is calculated by the corresponding *Author Score* of the article and is used to measure the quality of the article. As the citation probability of an individual paper, *Probability to Cite* is calculated by *Article Score*. All the references of a paper are recorded in *Citation List*. For journal agents, *Year* represents the year in which the journal was published. To avoid the confusion of journal information, *Journal ID* is set in which the journal published in different years can be categorized into corresponding list. All of the articles published in a journal are recorded in *Article List*. *JIF* represents the journal impact factor of the publication of the year. The main Features of the agents are listed in TABLE I.

TABLE I. THE FEATURES IN JIF AGENT-BASED SYSTEM.

Agent	Abbreviation of Features	Features
Author	AuS	Author Score
Article	Year	Year
	ArS	Article Score
	PtC	Probability to Cite
	CiL	Citation List
Journal	Year	Year
	JID	Journal ID
	ArL	Article List
	JIF	JIF

For expression concisely and clearly, this paper utilize the abbreviation of features in the formulas.

B. The Probability to Cite

Author Score is a random number in a closed interval of 0.1 to 1, which can be expressed as:

$$AuS = Random(0,1) \tag{1}$$

The *Article Score*, as the quality evaluation indicator of the article, is calculated by the corresponding *Author Score*:

$$AuC: \ Author \ Contribution$$
$$AuC = \sum_{i=1}^{n} \frac{AuS_i}{n} \quad n = 1,2,3,4,5 \tag{2}$$

However, it should be noted that the quality of an article not only influenced by the ability and the reputation of individual, but also strongly by the quality and extent of the cooperation. A great work, for instance, is more likely to be developed or proposed in a creative group. Thus, in the model, those factors can be represented by a random noise.

$$ArS = AuC \times Noise(0.9,1.1) \tag{3}$$

A normalize function is utilized to range the *Article Score*:

$$ArS' = Normalize(ArS) \tag{4}$$

To describe the correlation between paper quality and citation, the likelihood function $L(x)$ should follow characteristics:

- Boundedness: According to the characteristics of probability, the range of $L(x)$ should be $(0,1)$.

- Monotonicity: With the improvement of the article quality, the citation number of the corresponding article are increased. The quality of an article is positively related to the number of citation. Thus, $L(x)$ should be a monotone increasing function in its monotone $[0,1]$ interval.
- Slope: In order to follow the boundedness and monotonicity, the slope of $L(x)$ should decreases monotonously on $(0,+\infty)$.

To assess the relation between the score of an article and its citation number, a hyperbolic tangent function (see Fig. 2) is adopted as follows:

$$L(x) = \tanh(\frac{x}{\lambda}) \tag{5}$$

where x denotes the article score and λ is a parameter that can be used to adjust the correlation between the article score and the probability to cite other articles.

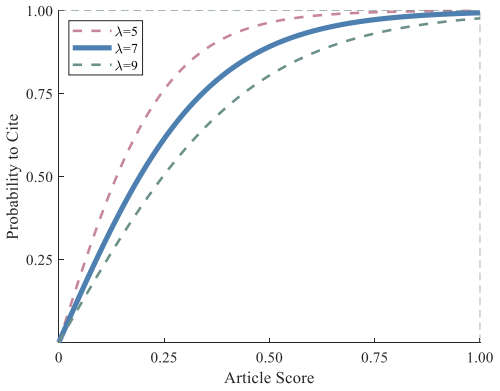


Fig. 2. The correlation between Article Score and Probability to Cite. This paper utilize the $L(x)$ at $\lambda = 7$ as the probability function of Article Score.

The citation probability of each article can be written as:

$$PtC = L(ArS') \tag{6}$$

JIF is computed by the citation number of journal in the first two years, which can be written by:

$$JIF_y = \frac{Citation_{y-1} + Citation_{y-2}}{Publication_{y-1} + Publication_{y-2}} \tag{7}$$

where $Citation_{y-1} + Citation_{y-2}$ is the total number of citations received by publications published by the journal in the $y-1$ year and $y-2$ year, and the $Publication_{y-1} + Publication_{y-2}$ is the total number of papers published by the corresponding journal in $y-1$ year and $y-2$ year.

Based on citation number, JIF of the corresponding journal can be easily calculated by the simulation model. Then, the score, which represent the quality of the article, can be calculated in each category. By comparing the JIF and article score in different categories, the correlation between JIF and quality of articles can be proved.

III. Simulation Experiment

This work is performed in the environment of Intel® Core™ is-7300HQ CPU @ 2.50GHz, 8.00GB RAM. The simulation code is written by Python 3.6.9.

Based on JIF, journals can be divided into three categories: High Quality Journals, Middle Quality Journals and Low Quality Journals. There are five journals in each category. In this paper, all the generated literatures are sorted by Article Score and divided into three categories: High Quality, Middle Quality and Low Quality. High quality articles will be contributed to high quality journals, medium quality articles will be issued by medium quality journals, while low quality articles will only be published by the journals with lower quality. The simulation results are shown in Fig. 3.

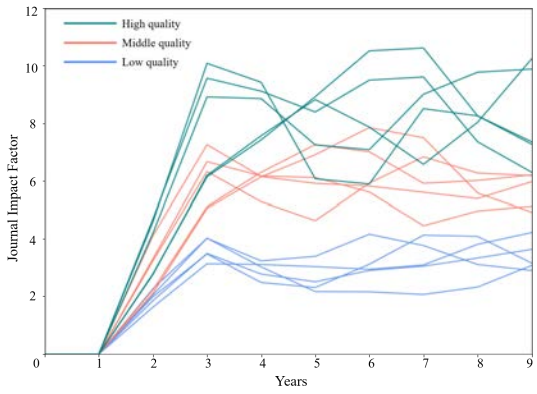
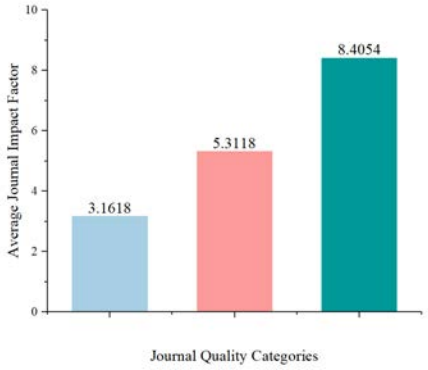


Fig. 3. The simulation result of JIF artificial systerm. The lines in different colors represent journals in different catagories. Each broken line represents an individual journal.

It can be found in Fig. 3 that the JIF is set to 0 at the first two years since the calculation of JIF. From the fourth year, the numerical change of JIF tends to be stable gradually. By the nineth year, there has been a clear distinction between the JIF of the three types of journals. It is clear that there exists a positive correlation between journal quality and JIF.

To further verify the correlation between JIF and article quality, the average of JIF and article quality since the third year in each category are shown in Fig. 4.



(a)

(b)

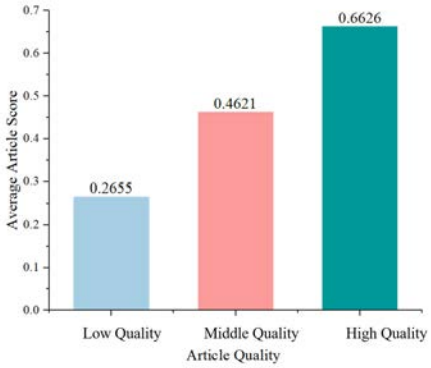


Fig. 4. The comparison of averge JIF and average article score. (a) represents the average JIF in 3 different catagories. The average article score of different qualities are painted in (b).

Similarly, the variations of JIF in different categories and article score can be observed in Fig. 4. The simulation results demonstrated that the academic level of a journal can be represented JIF and the quality of published article.

IV. Conclusion

As an indicator of assessing the academic journals, JIF is considered to have great value in library journal collection management. However, the studies on JIF mainly focused on bibliometrics and scientometrics in the past. With the development of computer technology, the increasing number of researchers attempt to analyze JIF with method of social computing. The development of wireless communication has promoted the real-time nature of data transmission [22-29]. This paper proposed a simulation model based on social computing, which can be utilized to analyze the correlation between JIF and the quality of published article. As an agent-based model, the JIF system consists of three agents (Author, Article and Journal). By assuming the relationship between Authors and Articles, this paper obtains the simulation result about Articles and Journals. According to the simulation result, the correlation between JIF and quality of articles can be validated.

In the future, other influencing factors can be added to the simulation model. Such as the length of the average paper review time, the impact of peer review system and the influence of the publication of an important paper on the reputation of corresponding authors. Through investigating JIF of real journals and related influencing factors, a real JIF dataset can be established. Based on the real JIF datasets, the simulation results utilizing real data can be obtained. By fitting the simulation results of real datasets with simulation results of artificial system, the correlation between JIF and quality of articles can be further evaluated. In addition, there are various evaluation indicators that can be considered in this model, so that an comprehensively parallel system about journal evaluation system can be developed.

References

[1] E. Garfield, "Journal Impact Factor: a Brief Review." *Canadian Medical Association Journal*, vol. 161, no. 8, pp. 979-980, 19 Oct. 1999

[2] E. Garfield, “Citation Indexes for Science”. *Science New Series*, vol. 122, no. 3159, pp. 108-111, 15 Jul. 1955

[3] E. Garfield, I. H. Sher, “New Factors in Evaluation of Scientific Literature Through Citation Indexing”. *American Documentation*, vol. 14, no. 3, pp. 195-201, 14 Jul. 1963

[4] E. Garfield, “Citation Analysis as A Tool in Journal Evaluation - Journals Can be Ranked by Frequency and Impact of Citations for Science Policy Studies”. *Science*, vol. 178, no. 4060, pp. 171-193, 3 Nov. 1972

[5] W. Glänzel, H. F. Moed, “Journal Impact Measures in Bibliometric Research”. *Scientometrics*, vol. 53, no. 2, pp. 171–193, Feb. 2002

[6] W. Glänzel, A. Schubert, “Predictive Aspects of a Stochastic Model for Citation Processes”. *Information Processing & Management*, vol. 31, no. 1, pp. 69-80, Jan. -Feb. 1995

[7] H. Gould, J. Tobochnik, W. Christian, “An Introduction to Computer Simulation Methods”, *New York: Addison-Wesley*, vol. 1, 1988

[8] M. Pidd, “Computer Simulation in Management Science”, *Chichester: Wiley*, vol. 4. 1998

[9] Z. Y. Tan, N. Cai, J. Zhou, “An Agent-based Simulation Model to Analyze Journal Impact Factor”. *In 2018 Chinese Control And Decision Conference (CCDC)*. IEEE, pp. 3681-3686, 9 Jul. 2018

[10] S. Allesina, “Modeling Peer Review: an Agent-based Approach”. *Ideas in Ecology and Evolution*, vol. 5, no. 2, pp. 27-35, 2012

[11] F. Y. Wang, “Social computing: Concepts, contents, and methods”. *International Journal of Intelligent Control and Systems*, vol. 9, no. 2, pp. 91-96, Jan. 2004

[12] F. Y. Wang, S. M. Tang, “Concepts and Frameworks of Artificial Transportation Systems”. *Complex Systems and Complexity Science*, vol. 1, no. 2, pp. 52-59, Feb. 2004

[13] S. Li, F. Bi, W. Chen, X. Miao, J. Liu, C. Tang, “An Improved Information Security Risk Assessments Method for Cyber-physical-Social Computing and Networking”. *IEEE Access*, 10311-10319 vol. 6, pp. 10311 - 10319, 31 Jan, 2018

[14] H. Chen, F. Y. Wang, D. Zeng, “Intelligence and Security Informatics for Homeland Security: Information, Communication, and Transportation”. *IEEE Transactions on Intelligent Transportation Systems*, vol. 5, no. 4, pp. 329-341, 6 Dec. 2004

[15] H. Chen, F. Y. Wang, “Guest Editors' Introduction: Artificial Intelligence for Homeland Security”. *IEEE Intelligent Systems*, vol. 20, no. 5, pp. 12-16, 14 Jul. 2005

[16] M. Navarro-Cáceres, J. Bajo, J. M. Corchado, , “Applying Social Computing to Generate Sound Clouds”. *Engineering Applications of Artificial Intelligence*, vol. 57, pp. 171-183, Jan. 2017

[17] K. Wang, Z. Shen, “Artificial Societies and GPU-based Cloud Computing for Intelligent Transportation Management”. *IEEE Intelligent Systems*, vol. 26, no. 43, pp. 22-28, 1 Aug. 2011

[18] N. R. Jennings, "On Agent-based Software Engineering." *Artificial Intelligence*, vol. 117, no. 2, pp. 277-296, March 2000

[19] X. Li, W. Mao, D. Zeng, F. Y. Wang, “Agent-based Social Simulation and Modeling in Social Computing”. *In International Conference on Intelligence and Security Informatics*, Springer, Berlin, Heidelberg, pp. 401-412, Jun. 2008

[20] L. I. Meho, "The Rise and Rise of Citation Analysis." *Physics World* vol. 20, no. 1, pp. 32-47, Jan. 2007

[21] L. Eghe, R. Rousseau. "Quantitative Methods in Library, Documentation and Information Science." *Elsevier Science Publishers* ,1990

[22] S. Sun, M. Kadoch, L. Gong and B. Rong, "Integrating network function virtualization with SDR and SDN for 4G/5G networks," IEEE Network, vol. 29, no. 3, pp. 54-59, May-June 2015.

[23] N. Zhang, N. Cheng, A. T. Gamage, K. Zhang, J. W. Mark and X. Shen, "Cloud assisted HetNets toward 5G wireless networks," in IEEE Communications Magazine, vol. 53, no. 6, pp. 59-65, June 2015.

[24] B. Rong, Y. Qian, K. Lu, H. Chen and M. Guizani, "Call Admission Control Optimization in WiMAX Networks," IEEE Transactions on Vehicular Technology, vol. 57, no. 4, pp. 2509-2522, July 2008.

[25] N. Chen, B. Rong, X. Zhang and M. Kadoch, "Scalable and Flexible Massive MIMO Precoding for 5G H-CRAN," in IEEE Wireless Communications, vol. 24, no. 1, pp. 46-52, February 2017.

[26] S. Sun, L. Gong, B. Rong and K. Lu, "An intelligent SDN framework for 5G heterogeneous networks," in IEEE Communications Magazine, vol. 53, no. 11, pp. 142-147, November 2015.

[27] M. Giordani, M. Polese, M. Mezzavilla, S. Rangan, and M. Zorzi, "Toward 6G networks: Use cases and technologies," IEEE Commun. Mag., vol. 58, no. 3, pp. 55-61, Mar. 2020.

[28] Z. Zhang et al., "6G wireless networks: Vision, requirements, architecture, and key technologies," IEEE Veh. Technol. Mag., vol. 14, no. 3, pp. 28-41, Sep. 2019.

[29] W. Saad, M. Bennis, and M. Chen, "A vision of 6G wireless systems: Applications, trends, technologies, and open research problems," IEEE Netw., vol. 34, no. 3, pp. 134-142, May/June 2020.