Intelligent Automation & Soft Computing
DOI:10.32604/iasc.2021.016259
Article

Tech Science Press

Improved Short-video User Impact Assessment Method Based on PageRank Algorithm

Lei Hong[1,*], Jie Yin[1], Ling-Ling Xia[1], Chao-Fan Gong[1] and Qi Huang[2]

[1]Jiangsu Police Institute, Nanjing, 210000, China
[2]Purdue University, West Lafayette, IN, 47907, USA
*Corresponding Author: Lei Hong. Email: honglei@jspi.edu.cn
Received: 28 December 2020; Accepted: 02 March 2021

Abstract: The short-video platform is a social network where users' content accelerates the speed of information dissemination. Hence, it is necessary to identify important users to effectively obtain information. Four algorithms (Followers Rank, Average Forwarding, K Coverage, and Expert Survey and Evaluation) have been proposed to calculate users' influence and determine their importance. These methods simply take the number of a user's fans or posts as the standard of influence evaluation, ignoring factors such as the paid posters, which makes such evaluations inaccurate. To solve these problems, we propose the short-video user influence rank (SVUIR) algorithm, which combines direct and indirect influence to comprehensively measure the influence of short-video users, using reference factors such as the number of fans, likes, number of users' works, users' work quality, focus on behavior, comments, and forwarding behavior. An experiment verifies the algorithm on Douyin (i.e., TikTok), which is a typical short-video platform, and confirms that SVUIR is more comprehensive and objective than the above four algorithms.

Keywords: User impact assessment; Douyin; PageRank; entropy weight method; short-video user influence rank

1 Introduction

The American government has prohibited TikTok, the overseas version of Douyin, since early 2020, which highlights the importance of short-video sites. The development of Web 2.0 technology has helped this form of social platform to proliferate, and short videos have quickly become popular. Short videos stand out and have gained many users with original content. 5G technology has made the mobile internet more popular, and video can attract the attention of netizens more than static text and pictures, changing the era from "all people have microphones" to "all people have cameras." According to the 45th Statistical Report on China's Internet Development [1], the number of short-video users in China had reached 773 million by March 2020, accounting for 85.6% of all internet users. The popularity of short video increases the speed of information dissemination, reduces the threshold of knowledge acquisition, and promotes the virtuous circle of knowledge. However, the content is highly entertaining, the user group level is complex, supervision is difficult, the amount of false news and rumors is large, and the spread range is wide, which is a hidden danger to social security and stability.

User influence analysis is an important field in the research of social networks. It can identify opinion leaders and clarify interactions. User influence evaluation comprehensively considers the relationships between users in a complex social network, and finds metrics to influence their behavior. Based on Google's PageRank algorithm, which is used to rank webpages in search results, we propose the short-video user influence rank (SVUIR) algorithm to effectively measure the influence of short-video users in social networks while considering the characteristics of the platform and its users.

2 Reviews of Studies

A short-video platform enables users to upload, convert, store, and playback video content on the internet, often via a structured, large-scale system that may generate revenue. Little analysis has been done on the impact of short-video users. Most studies use sociology, psychology, computer science, artificial intelligence, and complex network theory to evaluate the influence and communication of social networks [2]. Influence refers to the ability to imperceptibly change other people's behavior and thinking [3]. The influence of users in social networks is mainly spread through their characteristics and the network topology, and these influencing factors are complex and changeable.

User characteristics in social networks are important standards to evaluate their influence. The influence of users can be quantified by the number of fans, likes, and forwards. Scholars have studied various user characteristics to explore their relationship with influence. More common user feature algorithms include the followers rank algorithm based on the number of fans [4], expert survey evaluation based on expert index [5], and average forwarding algorithm based on reposts [6]. There are also some multi-index fusion influence judgment methods. Wang et al. [7] extracted seven features, such as the number of user replies, number of replies, and number of posts, and calculated an influence index of users by conducting a clustering operation with an EM algorithm. Ghiassi et al. [8] proposed UilRank, an impact recognition algorithm based on parameters such as the number of views, replies, and posts on a social network. Romero et al. [9] proposed the influence-passivity algorithm based on the user information forwarding rate. Cha et al. [10] established a measurement standard of Twitter user influence based on the number of comments, attention, and information forwarding. User characteristics are the foundation of user behavior analysis and modeling in social networks, which can provide data support for research on user influence. With the rapid development of social networks, especially the increasingly obvious role of user-generated content in the transmission of user influence, it is necessary to dig deep into the post content as an important carrier of user influence and find its positive effect on the transmission of influence.

Users form a social network through user attention, browsing, and forwarding, and influence is disseminated through the network. The analysis of social network topology is important in the study of user influence. Scholars measure user influence by studying node locations and network structures. Ding et al. [11] proposed that the more important the node position, the more likely it is to be connected by other nodes in the propagation network, and the higher the node influence. Gu et al. [12] mined the relationship between similar nodes and sorted their influence. Zhang et al. [13] established a role-conformity model by studying the relationship between neighbor nodes in academic networks. The k-coverage algorithm focuses on the cascading relationship between users and is more common among methods based on network structure [14]. Point centrality and betweenness centrality are also effective in determining user influence. Lei et al. [15] used centrality to describe the power of users in a communication network, and ranked the influence of users. The PageRank algorithm is the most widely used method, is based on quantity and quality hypotheses, and establishes a random walk model of node influence [16]. According to the characteristics of a social network, scholars have improved upon the algorithm's defects, evolved its model, and derived many extended algorithms. Xing et al. [17] added a chain relationship to form a weighted PageRank algorithm. Wang et al. [18] quantified the spread ability of users and integrated it in PageRank to obtain the Sf-uir algorithm. Ku et al. [19] combined PageRank with the HITS algorithm to more quickly obtain influence rankings. Zhou et al. [20] combined user dynamic behavior with PageRank to more accurately identify users with high influence. Fialal et al. [21] combined PageRank with author contributions to explore the influence ranking of scholars in an academic network.

The influence measurement method based on network topology is clear and direct, and can grasp the overall structure of a social network. PageRank, in particular, has played an important role in theory and practice. This paper introduces the characteristics of short-video social networks, comprehensively considers the number and quality of users, and proposes a short-video user influence measurement algorithm based on an improved PageRank algorithm.

3 Principle of Proposed SVUIR Algorithm

The topological structure of a short-video social network is compared to the webpage topology in the World Wide Web (WWW). Users on a short-video platform are equivalent to webpages on the WWW, and the behavior of users' mutual attention is equivalent to webpage links. Thus PageRank can be used to measure the influence of short-video users.

3.1 Mathematical Model of PageRank Algorithm

PageRank is a classic webpage ranking algorithm; a webpage's rank in the search results is based on how many other pages link to it. The mathematical model of the PageRank algorithm is [16]

$$\mathrm{Page}Rank(p_i) = \frac{1-d}{N} + d\sum_{p_i}\frac{PageRank(p_j)}{L(p_j)}, \tag{1}$$

where p_i is the number of webpages in the internet, $M(p_j)$ is the number of linked in pages of p_j, $L(p_j)$ is the number of linked out pages of p_j, N is the total number of pages, and d is a damping coefficient, which takes the value 0.85 according to experience.

3.2 Shortcomings of PageRank Algorithm

Using PageRank to calculate the user influence of short-video, users are treated as webpages, which only considers the "follow" relationship between users, and has the following deficiencies.

3.2.1 Initial PR Value is Inaccurate

PageRank determines the initial PR value by averaging, which is reasonable when ranking webpages. However, it is unreasonable that the initial PR value is given to the user directly by using the average method in the ranking of users. Considering the different factors between different users, such as the number of fans and the activity of the user, the influence of the user's own attributes on the short-video spread cannot be ignored.

3.2.2 Unreasonable Proportion of PR Value Transfer

PageRank divides the PR value of a webpage equally over linked pages. This is not accurate in the transmission of user influence because most users do not pay equal attention to all users of interest.

3.2.3 Topology Does Not Conform to Actual Situation

PageRank is applied to user influence based on the "follow" relationship, so it is largely related to the number of a user's fans. Because these probably contain artificial followers, the number of fans cannot truly reflect the influence of users [22]. Furthermore, dynamic behavior between users, such as comments and forwarding, will greatly affect the scope and dissemination speed of work.

3.3 Improvement of Influencing Factors of SVUIR Algorithm

A user exerts short-video influence on the crowd through posted videos. It is generally believed that more fans imply greater influence [23]. However, it is neither objective nor comprehensive to judge the influence of users only by the number of fans. Factors affecting influence are complex, including not only the number of fans, works, and likes, among other objective indicators, but also subjective indicators such as fan activity. To measure the influence of users, one must consider various factors so as to objectively reflect the dynamic [18] of users.

To improve upon the PageRank algorithm, we propose SVUIR, which considers two factors.

3.3.1 Factors of Direct Influence of Short-video Users

Due to the differences of users, it is not reasonable for PageRank to give them all the same initial PR value. Four factors can be used to evaluate the direct influence of users—number of fans, number of likes, number of works, and quality of works—so as to assign users different initial PR values. We can preliminarily screen out silent powder and zombie powder without any activity on an account, and exclude them when constructing the topology.

3.3.1.1 Number of Fans

The number of fans is a basic indicator of user influence. Generally speaking, the more fans a user has, the more people can see the user's works. In the topology, the more links a node has, the more important it is.

3.3.1.2 Total Likes

The total number of likes reflects a user's degree of recognition. The more total likes a user's works have, the more popular the user is, and the greater the influence.

3.3.1.3 Number of User's Works

Short video is the fast food of entertainment, and the popularity of a video drops quickly. Publishing works is the most important way for users to maintain their influence. The short-video platform will provide a user pool for each video, and each video will have a guaranteed number of users to watch. The more works users have, the more influential they are.

3.3.1.4 User's Work Quality

If a user's works are well reflected upon after the initial user pool test, the platform will push these high-quality works to more users. The quality of the work determines how many people can see it, and this is reflected in the number of views, comments, and forwards. Larger numbers mean higher quality.

3.3.2 Factors of Indirect Influence

Users cannot equally like all the short-video users they come into contact with. They are often interested in one or several users and are willing to invest in the promotion of their works. This preference is manifested in the interaction between users, which can increase their indirect influence, as measured by the following factors.

3.3.2.1 Focus on Behavior

A user can put favorite accounts on an interest list, and the system will automatically push new works from those accounts to the user. In the network topology, it forms a pair of chain-in and chain-out relations, and the influence of users can be spread through the network.

3.3.2.2 Like, Comment, and Forward Behavior

A webpage distributes its influence equally to pages it links out in the PageRank algorithm. With short-video users, it divides their influence equally among the users concerned. But in real life, users have preferred accounts. The three interactive behaviors of likes, comments, and forwards show a user's degree of liking for a user. Forwarding is most important, followed by comments, and then likes. Through these three behaviors, the closer the relationship between fans and users, the higher the degree of closeness, and the greater the proportion of fans' contributions allocated to a user. By quantifying the closeness between users, the transfer of PR value in unreasonable proportions is solved.

In conclusion, SVUIR calculates the initial PR value of users through their own attributes and solves the problem of inaccurate initial PR value determination. Through the interaction between users, it obtains a user's interest intention, determines the user's indirect influence, optimizes the allocation of PR value when transmitting, and constructs a user's influence communication network, which effectively avoids the interference of zombie powder and silent powder on user influence calculation.

4 Implementation of SVUIR

4.1 Derivation

The total influence of a user consists of direct and indirect influence. Direct influence is based on behavior, while indirect influence is based on interaction with fans. The total influence of user *i* is

$$SVUIR(U_i) = SVUIR\_direct(U_i) + SVUIR\_indirect(U_i), \tag{2}$$

where $SVUIR\_direct(U_i)$ and $SVUIR\_indirect(U_i)$ are, respectively, the direct and indirect influence of user *i*.

The direct influence of users can be divided into two parts. One is radiated to the next level through the "follow" relationship, which is related to the number of users, fans, and works. The other is pushed by the platform through the video, which affects more users. This influence is reflected by the number of likes, comments, and forwards of a video. The direct influence is calculated as

$$SVUIR\_direct(U_i) = S(U_i) * M(U_i). \tag{3}$$

where $S(U_i)$ is the direct influence of users, and $M(U_i)$ is the direct influence of their videos.

The direct influence of users is reflected in the total number of works, likes, and fans,

$$S(U_i) = a_1 * W_{U_i} + b_1 * Z_{U_i} + c_1 * B_{U_i}, \tag{4}$$

where W_{U_i}, Z_{U_i}, and B_{U_i} are the total number of works, likes, and fans, respectively, of user *i*, and a_1, b_1, and c_1 are their respective weights.

The direct influence of video lies in the number of people watching and the influence of secondary transmission, which is mainly reflected in the number of likes, comments, and forwards. It is calculated as

$$M(U_i) = \sum_{m_j \in T_{U_i}} (a_2 * L_{m_j} + b_2 * C_{m_j} + c_3 * R_{m_j}), \tag{5}$$

where T_{U_i} is the collection of works published by user *i* within the statistical period, which is set to 30 days based on experience; m_j is any video work published by user *i* within the statistical period; L_{m_j} is the number of likes obtained by work m_j; C_{m_j} is the number of comments on work m_j; R_{m_j} is the number of reposts of work m_j; and a_2, b_2, and c_2 are respective weights.

The indirect influence of users is distributed by all fans to their favorite users through interactive behaviors,

$$SVUIR\_indirect(U_i)=(1-d)+d*\sum_{U_j\in F_{U_i}} affect(U_i,U_j)*SVUIR\_direct(U_j), \tag{6}$$

where d is a damping coefficient, F_{U_i} is the set of fans of user i, j is any fan of user i, and $affect(U_i,U_j)$ is the influence ratio assigned to user i by fan j based on their interaction,

$$affect(U_i,U_j)=\frac{\mathrm{int}eract(U_i,U_j)}{\sum\limits_{U_p\in force_{U_j}} \mathrm{int}eract(U_p,U_j)}, \tag{7}$$

where $\mathrm{int}eract(U_i,U_j)$ is the propagation ability of user i to fan i, $force_{U_j}$ consists of all admirers of fan j, and p is any admirer of fan j.

The propagation ability of fans to their admirers depends on their degree of preference for the user, which is mainly reflected in the number of likes, comments, and forwards of videos. This is calculated as

$$\mathrm{int}eract(U_i,U_j)=a_3*\frac{T_{like(U_i,U_j)}}{T_{like(U_j)}}+b_3*\frac{T_{comment(U_i,U_j)}}{T_{comment(U_j)}}+c_3*\frac{T_{forward(U_i,U_j)}}{T_{forward(U_j)}}, \tag{8}$$

where $T_{like(U_i,U_j)}$ is the number of likes that fan j gives to user i, $T_{like(U_j)}$ is the total number of likes that fan j gives to admirers, $T_{comment(U_i,U_j)}$ is the number of comments that fan j gives to user i, $T_{comment(U_j)}$ is the total number of comments that fan j gives to admirers, $T_{forward(U_i,U_j)}$ is the number of forwards that fan j gives to user i, and $T_{forward(U_i,U_j)}$ is the total number of forwards that fan j gives to admirers.

By substituting Eq. (6) into Eq. (2), we can obtain

$$SVUIR(U_i)=SVUIR\_direct(U_i)+(1-d)+d*\sum_{U_j\in F_{U_i}} affect(U_i,U_j)*SVUIR\_direct(U_j), \tag{9}$$

where $SVUIR\_direct(U_i)+(1-d)$ is a constant. The influence of short-video users can be obtained after multiple iterations to achieve Markov convergence.

4.2 Parameter Determination

Weight parameters appear in Eqs. (4), (5), and (8) in the mathematical model of the SVUIR algorithm. Existing data and experiments cannot provide reference values for weights. Hence, we adopt the entropy weight method to determine the parameters.

The entropy weight method is commonly used to determine weights in comprehensive evaluation [24]. Entropy represents the degree of disorder of information. An index with low entropy provides more information, has a greater role in comprehensive evaluation, and has a higher weight.

For any k variables $\{X_1, X_2, X_3, \ldots X_k\}$, the data in X_i, are $\{X_{i1}, X_{i2}, \ldots X_{in}\}$. The entropy weight method has three steps in the determination of the weights of k variables.

4.2.1 Data Standardization

The first step is to calculate the standardized data Y_{ij} for X_{ij}: $Y_{ij}=\dfrac{X_{ij}-\min(X_i)}{\max(X_i)-\min(X_i)}$.

4.2.2 Entropy of Information Calculation

The second step is to calculate the comentropy E_j of variable j, $E_j=-\ln(n)^{-1}\sum\limits_{i=1}^{n} p_{ij}\ln(p_{ij})$, where $p_{ij}=\dfrac{Y_{ij}}{\sum\limits_{r=1}^{n} Y_{ir}}$.

4.2.3 Weight Value Determination

The third step is to calculate the weight of variable i, $w_i = \dfrac{1 - E_i}{k - \sum E_i}$.

5 Analysis of Algorithm Results

5.1 Data Acquisition

The short-video platform Douyin was selected for data acquisition and analysis to evaluate the effect of the SVUIR algorithm. A mobile phone and computer were placed on the same LAN, and the mobile phone's proxy IP was changed to the computer's IP. Scripts were written to simulate mobile phone users receiving videos recommended by the Douyin platform. The mobile data traffic was captured and duplicate video authors were removed to obtain 573 Douyin users. A crawler was written to collect basic information and behavior information of users from time 0:00 on November 15, 2019, to 0:00 on December 15, 2019. The crawled data were divided into databases [25] of user information, video information, and interactive behavior. The collected basic fields of user, video, and interactive behavior information are shown in Tabs. 1–3, respectively. There were 573 pieces of user information, 16,456 pieces of video information, and 1,054,695 pieces of interactive behavior information.

Table 1: User information form

Field name	Description
uid	original ID of Douyin user
username	user name
sex	gender
sign	personalized signature
movie sum	total number of works
like sum	total number of likes
force	users this person pays attention to

Table 2: Video information form

Field name	Description
movie id	ID of this video
title	title of this video
uid	ID of author
username	name of author
time	release time
like	number of likes
comment	number of comments
forward	number of forwards
introduction	introduction of video

Table 3: Interactive behavior information form

Table name	Description
type	categories of actions (likes or comments)
movie id	video ID of action performed
start id	ID of implementer
end id	ID of person the behavior acted on
time	time to perform the action

5.2 Data Processing

To ensure the objectivity and authenticity of the experimental results, the data were processed, and all codes were converted to UTF-8 format to facilitate the import and export of the database [26]. Fields unrelated to the algorithm were deleted. Users were simply filtered to remove obvious silent and zombie users. The processed results were imported into the database.

The top 100 users with high attention were selected, and the network topology between users was drawn according to the relationship between followers as shown in Fig. 1. The influence value of short-video users was calculated and ranked by the SVUIR algorithm, and the influence ranking results were also calculated for the commonly used Followers Rank, Average Forward, K Coverage, and Expert Survey and Evaluation algorithms. To ensure a performance comparison of algorithm results under the same conditions, algorithms used the same data source and ran in the same computer environment.

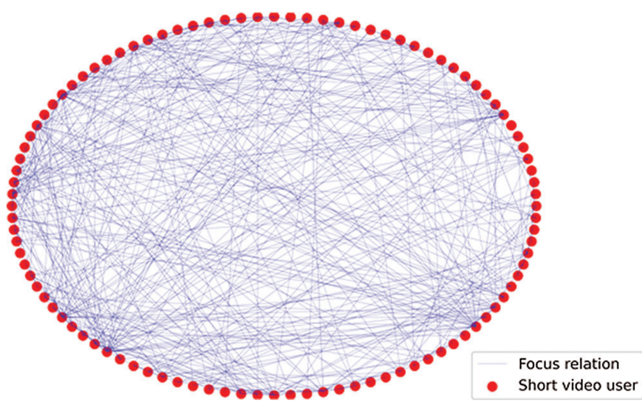

Figure 1: Short-video platform user network topology

5.3 Comparative Analysis of Results

Influence results of short-video platform users were compared as obtained by the SVUIR, Followers Rank, Average Forwarding [27], K Coverage, and Expert Survey and Evaluation [28] algorithms, and their pros and cons were judged. MATLAB was used to show the ranking of the top 10 users according to each algorithm. Fig. 2 shows the situation [29].

5.3.1 Comparison of SVUIR and Followers Rank Algorithms

Tab. 4 shows the ranking comparison of the top 10 users calculated by the SVUIR and Followers Rank algorithms. From Tab. 4 and Fig. 2, we can see that while there was a small difference in the rankings obtained by the two algorithms, there were users with big differences in rankings. Live in rizhao was ranked sixth by SVUIR and 34th by Followers Rank. This user has fewer followers, but each video receives many comments and forwards. Users like Zhu Xiaohan and Gui Ge were ranked highly by Followers Rank. Although they have a large number of fans, their works are few, and their response degree is low. The analysis shows that Followers Rank focuses on the number of fans without considering their quality. This ignores the particularity of fans, which leads to the result that Followers Rank is less convincing than SVUIR.

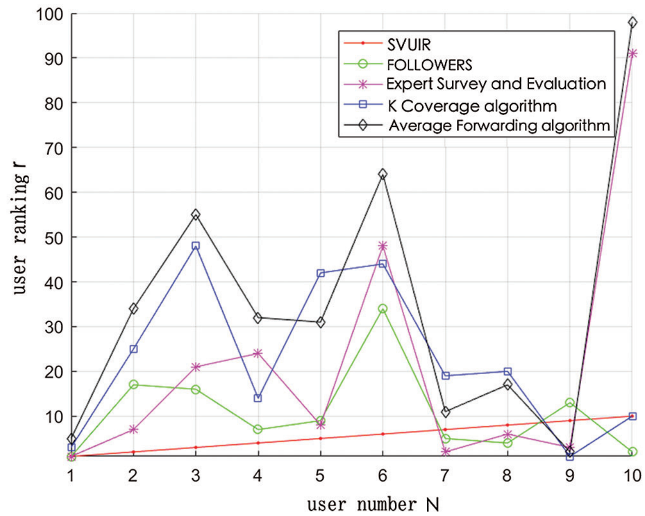

Figure 2: Ranking comparison of SVUIR and other algorithms

Table 4: Comparison of SVUIR and Followers Rank algorithms on top 10

SVUIR ranking	Username	Followers Rank ranking	Username
1	CCTV news	1	CCTV news
2	Economic Daily	2	People's Daily online
3	China Economic Net	3	Zhu xiaohan
4	The Xinhua News Agency	4	Tang tang
5	China Changan	5	Qi ayi(Qi haoluren)
6	Live in rizhao	6	Gui ge
7	Qi ayi(Qi haoluren)	7	The Xinhua News Agency
8	Tang tang	8	Xu juncong
9	Zhang baizhi	9	China Changan
10	People's Daily online	10	Biao ge(Tan jinzhan)

5.3.2 Comparison of SVUIR and Average Forwarding Algorithms

Tab. 5 shows the ranking comparison of the top 10 users calculated by the SVUIR and Average Forwarding algorithms. It can be seen from Tab. 5 and Fig. 2 that the results obtained by the two algorithms are quite different. Among the top 10 users from the SVUIR algorithm, only one user of CCTV news maintains top 10 influence in the Average Forwarding algorithm, and the rest are individual users. Although these users have high forwarding numbers, there are many zombie and silent fans, and their forwarding cannot prove effective for users' influence. People's Daily online was ranked 98th by the Average Forwarding algorithm, but it is still an influential government official account. This shows that the influence of users is not necessarily positively correlated with the number of video forwards.

Table 5: Comparison of SVUIR and average forwarding algorithms on top 10

SVUIR algorithm ranking	Username	Average Forwarding algorithm ranking	Username
1	CCTV news	1	Mr Xiaozhanglaile
2	Economic Daily	2	Zhang baizhi
3	China Economic Net	3	Biao ge(Tanjinzhan)
4	The Xinhua News Agency	4	Zhao gejiongshi
5	China Changan	5	CCTV news
6	Live in rizhao	6	Cheng ye
7	Qi ayi(Qi haoluren)	7	Piao ge???
8	Tang tang	8	Qi guoshaocai
9	Zhang baizhi	9	Xu juncong
10	People's Daily online	10	Jun ge ZNL

5.3.3 Comparison of SVUIR and K Coverage Algorithms

Tab. 6 compares the top 10 user rankings calculated by the SVUIR and K Coverage algorithms. From Tab. 6 and Fig. 2, we can find that although some users' influence rankings are relatively close, there are some differences. We can see that the K Coverage algorithm mainly considers the hierarchical relationship among users, but not the influence of users' own factors, such as The Xinhua News Agency and other government official account numbers, which themselves have an important influence. Therefore, K Coverage lacks the persuasiveness of SVUIR.

Table 6: Comparison of SVUIR algorithms and K Coverage algorithm on top 10

SUVIR algorithm ranking	Username	K Coverage algorithm ranking	Username
1	CCTV news	1	Zhang baizhi
2	Economic Daily	2	Biao ge(Tanjinzhan)
3	China Economic Net	3	CCTV news
4	The Xinhua News Agency	4	Zhu xiaohan
5	China Changan	5	Zhao gejiongshi
6	Live in rizhao	6	Mr Xiaozhanglaile
7	Qi ayi(Qi haoluren)	7	Xu juncong
8	Tang tang	8	Piao ge???
9	Zhang baizhi	9	Neng nengshu
10	People's Daily online	10	People's Dailyonline

5.3.4 Comparison of SVUIR and Expert Survey Evaluation Algorithms

Tab. 7 shows the ranking comparisons of the top 10 users calculated by the SVUIR and Expert Survey and Evaluation algorithms. It can be seen from Tab. 7 and Fig. 2 that their results are somewhat similar, but the Expert Survey and Evaluation algorithm has two defects. One is that the values of index weights depend on expert experience. Although the accuracy can be improved through discussion with experts, there is still a certain error in the actual situation. Second, the algorithm does not consider the hierarchical structure of users. As a result, People’s Daily online, which has a large number of active fans, does not rank highly. Therefore, SVUIR is more reliable.

Table 7: Comparison of SVUIR algorithms and expert survey and evaluation algorithm on top 10

SVUIR algorithm ranking	Username	Expert Survey algorithm Ranking	Username
1	CCTV news	1	CCTV news
2	Economic Daily	2	Qi ayi(Qi haoluren)
3	China Economic Net	3	Zhang baizhi
4	The Xinhua News Agency	4	Biao ge(Tanjinzhan)
5	China Changan	5	Xu juncong
6	Live in rizhao	6	Tang tang
7	Qi ayi(Qi haoluren)	7	Economic Daily
8	Tang tang	8	China Changan
9	Zhang baizhi	9	Mr Xiaozhanglaile
10	People's Daily online	10	Lan yan

In summary, personal accounts whose rankings are high in traditional algorithm have a lower ranking in the SVUIR algorithm compared with official accounts of the government. This is because many artificial followers exist among fans of personal accounts, while official government accounts of government tend to have high-quality, active fans. The SVUIR algorithm considers not only the user’s own factors but the contribution of the user’s fans, which can more comprehensively and objectively show the influence of users of short-video platforms.

6 Conclusions

Based on the PageRank algorithm and the characteristics of short-video platform users, the SVUIR short-video user influence evaluation algorithm was proposed. The method decomposes user influence into direct and indirect influence. Direct influence reflects the role of users themselves, as expressed by the total number of users’ followers likes, and forwards. Indirect influence is obtained through interaction between users and is calculated based on the number of likes, comments, and forwards. The SVUIR algorithm was compared to the Followers Rank, Average Forwarding, K Coverage, and Expert Survey and Evaluation algorithms. It was found that the SVUIR algorithm can more truly and objectively reflect the influence of short-video platform users.

Acknowledgement: I would like to express my gratitude to all those who have helped me during the writing of this thesis. I gratefully acknowledge the help of Professor Guan Gui. I do appreciate his patience, encouragement, and professional instructions during my thesis writing.

Funding Statement: This study has been supported by Jiangsu Social Science Foundation Project (Grant No: 20TQC005). Philosophy Social Science Research Project Fund of Jiangsu University (Grant No: 2020SJA0500). Jiangsu Province modern education technology research project (Grant No: 2017-R-59195). The key teaching reform project of Jiangsu Police Institute (Grant No: 2019A30). The National Natural Science Foundation of China (Grant No. 61802155).

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] M. M. Huang, "A statistical report on the development of the Internet in China shows that the number of Internet users has exceeded 900 million," *Youth Journalist*, vol. 13, pp. 46, 2020.

[2] Z. M. Han, Y. Chen, W. Liu, B. H. Yuan, M. Q. Li *et al.,* "Research on node influence analysis in social networks," *Journal of Software*, vol. 28, no. 1, pp. 84–104, 2017.

[3] J. Li, Z. Chen and W. J. Huang, "Micro-blog impact evaluation study," *Netinfo Security*, vol. 3, pp. 10–13, 2012.

[4] C. Hao, "MicroBlog User Ranking Research based on Hadoop," M.S. dissertation. East China University of Science and Technology, China, 2014.

[5] J. H. Zhang, Y. J. Gu and S. H. Zhang, "The Microblogging user influence assessment based on PageRank and user behavior," *Netinfo Security*, vol. 6, pp. 73–78, 2015.

[6] H. Chen, "MicroBlog user ranking research based on Hadoop," Ph.D. dissertation. East China University of Science and Technology, China, 2014.

[7] Y. Wang, J. P. Zeng, B. H. Zhou and C. R. Wu, "Online forum opinion leaders discovering method based on clustering analysis," *Computer Engineering*, vol. 36, no. 5, pp. 44–46+49, 2011.

[8] M. Ghiassi, J. Skinner and D. Zimbra, "Twitter brand sentiment analysis: A hybrid system using N-gram analysis and dynamic artificial neural network," *Expert Systems with Applications*, vol. 40, no. 16, pp. 6266–6282, 2013.

[9] D. M. Romero, W. Galuba, S. Asur and B. A. Huberman, *Influence and passivity in social media*. Berlin Heidelberg: Springer, 18–33, 2011.

[10] M. Cha, H. Haddadi, F. Benevenuto and K. P. Gummadi, "Measuring user influence in Twitter: The million follower fallacy," in *Proc. of Int. Conf. on Weblogs and Social Media*, California, USA, pp. 14, 2010.

[11] Z. Ding, Y. Jia, B. Zhou and Y. Han, "Mining topical influences based on the multi-relational network micro-blogging sites," *China Communications*, vol. 10, no. 1, pp. 93–104, 2013.

[12] Y. R. Gu and Z. Y. Zhu, "Node ranking in complex networks based on LeaderRank and modes similarity," *Journal of University of Electronic Science and Technology of China*, vol. 46, no. 2, pp. 441–448, 2017.

[13] J. Zhang, J. Tang, H. Zhuang, C. W. Leung and J. Li, "Role-aware conformity influence modeling and analysis in social networks," in *Proc. of the 28th AAAI Conf. on Artificial Intelligence (AAAI 2014)*, Menlo Park, California, USA, pp. 958–965, 2014.

[14] C. L. Gao, Z. Y. Sun, X. F. Xing and C. F. Li, "Energy efficient k-degree coverage algorithm in wireless sensor networks," *Computer Engineering and Applications*, vol. 52, no. 23, pp. 142–147+235, 2016.

[15] H. Lei, S. S. Nie, X. B. Huang and W. Ma, "The behavior characteristics of network disseminators: An empirical study based on social network analysis," *Journal of Intelligence*, vol. 34, no. 1, pp. 161–168, 2016.

[16] L. Page, "The PageRank citation ranking: Bringing order to the web," *Stanford Digital Libraries Working Paper*, vol. 1, pp. 1–14, 1988.

[17] W. P. Xing and A. Ghorbani, "Weighted PageRank algorithm," in *Communication Networks and Services Research, Proc. of Second Annual Conf.*, Fredericton, NB, Canada, pp. 305–314, 2004.

[18] D. Wang, J. Xu, C. Y. Duan, Y. Y. Wu and J. Sun, "Improved user influence evaluation algorithm based on PageRank," *Journal of Harbin Institute of Technology*, vol. 50, no. 5, pp. 60–67, 2018.

[19] S. Ku and Z. Liu, "An improved algorithm for page rank optimization based on PageRank and HITS algorithms," *Journal of Wuhan University of Science and Technology*, vol. 42, no. 2, pp. 155–160, 2019.

[20] F. Zhou and M. T. Gao, "Discovery algorithm of opinion leaders for network community based on PageRank," *Computer Engineering*, vol. 44, no. 2, pp. 203–209+219, 2018.

[21] D. Fiala, F. Rousselot and K. Jezek, "PageRank for bibliographic networks," *Scientometrics*, vol. 76, no. 1, pp. 135–158, 2008.

[22] F. Luo, Y. Xu, Q. M. Pu and Q. Z. Qiu, "Multi-dimensional measure of microblog user influence based on PageRank," *Application Research of Computers*, vol. 32, no. 3, pp. 1354–1358+1367, 2020.

[23] H. Zhang, G. S. Liu and B. Su, "A computer method for microblogging users influence," *Computer Applications and Software*, vol. 32, no. 3, pp. 41–44, 2015.

[24] Q. Y. Cheng, "Structure entropy weight method to confirm the weight of evaluating index," *Systems Engineering-Theory & Practice*, vol. 30, no. 7, pp. 1225–1228, 2010.

[25] Z. M. Xing, L. N. Wang, W. B. Xing, Y. J. Ren, T. Li *et al.,* "Application of ontology in the web information retrieval," *Journal on Big Data*, vol. 1, no. 2, pp. 79–88, 2019.

[26] J. Wang, X. Wang, Y. Yang, H. Zhang and B. Fang, "A review of data cleaning methods for web information system," *Computers Materials & Continua*, vol. 62, no. 3, pp. 1053–1075, 2020.

[27] L. Zhang, "The optimization analysis of the communication model of negative influence of the entrepreneur's social relationship change," *Intelligent Automation & Soft Computing*, vol. 26, no. 3, pp. 577–583, 2020.

[28] Y. Hua, B. Chen, Y. Yuan, G. Zhu and J. Ma, "An influence maximization algorithm based on the mixed importance of nodes," *Computers Materials & Continua*, vol. 59, no. 2, pp. 517–531, 2019.

[29] H. Ghasemi, H. S. Park, X. Zhuang and T. Rabczuk, "Three-dimensional isogeometric analysis of flexoelectricity with Matlab implementation," *Computers Materials & Continua*, vol. 65, no. 2, pp. 1157–1179, 2020.