# Research on Network Node Influence in New Media Environments Based on Information Entropy: A Case Study of WeChat Official Accounts (Postprint)

**Authors:** Xing Yunfei, Wang Xiwei, Han Xuewen, Zhang Changliang

**Date:** 2023-08-26T00:00:00+00:00

## Abstract

[Objective/Significance] Research on network node influence in the new media environment enables in-depth analysis of information dissemination patterns, thereby facilitating the adoption of targeted measures for reasonable control of information propagation. [Method/Process] Based on information entropy theory, a network node influence model in the new media environment is constructed, with WeChat Official Accounts serving as a case study for node influence measurement. Direct influence, indirect influence, and comprehensive influence of nodes are analyzed in depth, and simulation analysis of the constructed model is performed using Matlab software. [Results/Conclusion] In the new media environment, the comprehensive influence of network nodes increases with the number of connected nodes and the frequency of inter-node interactions, with both direct and indirect influence growing at different rates. However, when the information entropy value of indirect influence exceeds 100, direct influence becomes the dominant factor affecting node comprehensive influence.

## Full Text

## Preamble

Vol. 62 No. 5, March 2018
ChinaXiv Cooperative Journal

**Research on Network Node Influence in New Media Environments Based on Information Entropy: A Case Study of WeChat Official Accounts**

Xing Yunfei[1], Wang Xiwei[1,2], Han Xuewen[1], Zhang Changliang[1]
[1] Department of Information Management, School of Management, Jilin University, Changchun 130022
[2] Big Data Management Research Center, Jilin University, Changchun 130022

## Abstract

[**Purpose/Significance**] Research on network node influence in new media environments enables deep analysis of information dissemination patterns, thereby facilitating targeted measures for reasonable control of information propagation. [**Method/Process**] Based on information entropy theory, this study constructs a network node influence model for new media environments, using WeChat official accounts as a case study to measure node influence and analyze direct influence, indirect influence, and comprehensive influence in depth. Finally, MATLAB software is employed for simulation analysis of the constructed model. [**Results/Conclusion**] In new media environments, the comprehensive influence of network nodes increases with the number of connected nodes and the frequency of interactions between nodes, with both direct and indirect influence growing at different rates. However, when the information entropy value of indirect influence exceeds 100, direct influence becomes the primary factor affecting node comprehensive influence.

## 1 Introduction

The latest China Internet industry report shows that as of December 2016, China's mobile internet user base reached 695 million, with growth rates exceeding 10% for three consecutive years. Mobile users accounted for 95.1% of total internet users, as netizens' devices further concentrated on mobile platforms, making mobile internet the primary medium for online public opinion dissemination. With the comprehensive penetration of mobile internet applications, netizens increasingly use new media tools such as WeChat, Weibo, and forums to disseminate information, expressing comments, intentions, attitudes, and emotions on topics closely related to their interests. Consequently, how to effectively manage and control network nodes in new media environments has become a new concern for information dissemination and regulatory agencies.

Both domestic and international scholars have conducted research on node influence in social networks. Foreign scholar D. Miorandi et al. proposed a K-core decomposition method for analyzing social network user influence and designed a framework to examine the characteristics and trends of network nodes over time in dynamic networks. L. Lü et al. introduced a parameter-free algorithm based on LeaderRank to quantify social network user influence and compared its ranking results with PageRank. K. Deanne et al. constructed a new competitive influence model for social network users based on game theory, analyzing network topology structure and node diffusion trends to identify opinion leaders. V. Arnaboldi et al. compared large-sample user data from Twitter and Facebook, concluding that network users with strong direct influence relationships play important roles in network information diffusion. Domestic scholars such as Cao Xueyan et al. proposed a complete technical method for mining and classifying network public opinion nodes to reveal structural complexity, scale-free properties, and sub-community structures in network public opinion dissemination. Kang Wei employed social network analysis to study network structural characteristics of public opinion dissemination in emergencies, exploring information propagation paths, speed, and scope. Jiang Kan et al. constructed a WSD-Rank diffusion influence measurement model from the perspective of information diffusion quality, analyzing distribution patterns and internal causes of key nodes. Wang Yuefen et al. designed a public opinion data crawling system and used social network analysis to study key node identification and application. However, while existing research has established models or employed social network analysis methods to study network node influence, few scholars have applied information entropy theory to analyze node influence in new media environments.

This study identifies and measures network node influence in new media environments based on information entropy theory, using MATLAB to simulate how node influence changes with node quantity and interaction frequency. The research addresses three key questions: (1) How can information entropy theory be applied to analyze node influence magnitude in new media environments? (2) How can a network node influence model for new media environments be constructed based on information entropy theory? (3) How can data from WeChat official accounts be obtained to calculate node influence and simulate the constructed model? This paper provides a new theoretical perspective for studying network node influence in new media environments and analyzes influence magnitude of WeChat official account network nodes, predicting influence variation patterns, which holds significant importance for relevant industry institutions to effectively manage information dissemination on new media platforms.

## 2 Theoretical Foundation

### 2.1 Information Entropy

In 1948, Shannon first proposed the concept of information entropy in his work *A Mathematical Theory of Communication*, solving the problem of information measurement. Information entropy defines the uncertainty of an information source and describes the amount of information for different values of random variables through its expression (see Formula (1)), where x_i represents random variables and p_i represents the set of all output results (i.e., probability function).

$$H(X) = K[-\log p_i] = -\sum p_i \log p_i \quad (0 \leq p_i \leq 1, i = 1, 2, \dots, n) \quad \text{Formula (1)}$$

The main idea embedded in Formula (1) assumes that the uncertainty function f is a monotonically decreasing function of probability p, and the probability of n independent uncertain events should be their sum. Information entropy theory has been primarily applied in physics and statistical mechanics, with some scholars using it to analyze information patterns and a limited number of studies applying it to computer science. R. Sangam et al. discussed the limitations of K-means algorithms based on information entropy theory and proposed a similarity coefficient to improve clustering accuracy. Y. Li et al. analyzed content similarity of user comments on search engines, calculated connection weights between users based on information entropy theory, and divided network communities. Wei Zhihui et al. used information entropy theory to verify the rationality of evaluation indicator systems, then applied an unascertained measurement model to comprehensively evaluate collected Weibo user data, ultimately identifying opinion leaders.

### 2.2 Network Node Influence

Network node influence is divided into direct influence, indirect influence, and comprehensive influence. Direct influence refers to the impact that a node A has on other nodes directly connected to it. The aggregated influence of all nodes directly connected to node A constitutes node A's direct influence. Indirect influence refers to the impact that node A has on node B through one or several intermediate nodes, even if A and B are not directly connected. The aggregated influence of node A on all indirectly connected nodes constitutes node A's indirect influence. Comprehensive influence represents the combined evaluation result of direct and indirect influence, measuring a node's overall role in information propagation. By assigning weights $\alpha$ and $\beta$ to direct and indirect influence respectively (with $\alpha + \beta = 1$), the comprehensive influence value of node A in information dissemination can be calculated as the final measure of network node influence.

### 2.3 Network Node Influence in New Media Environments Based on Information Entropy

New media encompasses new digital media forms supported by technologies including social media, digital media, mobile media, intelligent communications, and Internet television, such as WeChat, Weibo, QQ, blogs, and forums. These platforms serve as new channels where network users express opinions, participate in social activities, and share information, characterized by digitization, networking, diversification, real-time capabilities, and interactivity.

Network node influence in new media environments refers to the contribution magnitude of network nodes in information dissemination and exchange processes within new media contexts. Research on this topic helps relevant departments maintain network order and enables management agencies to better guide, monitor, warn, and control information dissemination and interaction on network platforms. Due to differences in node attributes and communication channels, these network nodes exhibit significant variations in influence during information propagation. Information entropy represents the occurrence probability of discrete random events. Given the conceptual ambiguity of information entropy, this study interprets it as the probability of specific information appearing. In new media environments, larger information entropy values of network nodes indicate greater probabilities of connection with other nodes, and larger information entropy values between nodes indicate closer connections.

## 3 Network Node Influence Model in New Media Environments

### 3.1 Influence Analysis Framework

Network nodes refer to points with information exchange behaviors in networks, which can be workstations, clients, network users, personal computers, or servers. The entire network consists of these nodes connected by communication lines forming specific geometric structures. Network nodes achieve information diffusion through information production, transmission, storage, and processing.

This study constructs a node influence model based on information entropy theory, obtains information dissemination samples of network nodes in new media environments over a period, calculates node influence magnitude through programmed compilation, and finally uses MATLAB software for simulation analysis of node influence variation trends. The specific steps are: (1) Model construction: Using information entropy theory to determine mathematical expressions for direct influence, indirect influence, and comprehensive influence of network nodes in new media environments. (2) Network construction: Setting the search period from March 14-24, 2017, programming for sample data collection from data sources, using Excel and Access for data processing to obtain closely connected network nodes, and drawing the node network. (3) Program compilation: Designing Java programming language on the Java platform to input node information and edge weight information from the constructed network, calculating information entropy values for direct, indirect, and comprehensive influence based on the model to obtain each node's influence magnitude. (4) Simulation analysis: Using MATLAB software to simulate the network node influence model, drawing cloud maps of node influence variation with node quantity and interaction frequency for visualization, and analyzing node influence change patterns during information propagation. The analysis framework is shown in Figure 1 [Figure 1: see original paper].

### 3.2 Node Influence Model Construction Based on Information Entropy

**3.2.1 Direct Influence** When calculating node direct influence, we first compute the information entropy of connected node quantity. Let f_{ij}(t) represent the relationship between node i and node j at time t. If node i connects to node j, f_{ij}(t) = 1; otherwise f_{ij}(t) = 0. K_i(t) denotes the set of all nodes connected to node i at time t, making K_i(t) an important indicator measuring node i's direct influence at time t. For example, if Xinhua News

Agency, NetEase News, and "I Was Shocked At That Time" all source from Xinhua Net within time t, then K_i(t) = 3 for Xinhua Net.

$$K_i(t) = \sum f_{ij}(t) \quad \text{Formula (2)}$$

The connected node information entropy I_c^i(t) for node i is expressed as:

$$I_c^i(t) = -\sum_{j=1}^{K_i(t)} \frac{1}{K_i(t)+1} \log_{10} \frac{1}{K_i(t)+1} \quad \text{Formula (3)}$$

Next, we calculate the interaction frequency information entropy between nodes. The interaction frequency between nodes is represented by F_{ij}(t), indicating the number of contacts between node i and node j at time t. For instance, if Xinhua Net's articles source from Beijing News twice within time t, then F_{ij}(t) = 2. Therefore, the interaction frequency information entropy I_f^i(t) between node i and node j at time t is:

$$I_f^i(t) = \sum_{j=1}^{K_i(t)} \frac{F_{ij}(t)}{\sum_{k=1}^{K_i(t)} F_{ik}(t)+1} \log_{10} \frac{F_{ij}(t)}{\sum_{k=1}^{K_i(t)} F_{ik}(t)+1} \quad \text{Formula (4)}$$

The direct influence information entropy MI_i(t) of node i should be the product of its connected node quantity information entropy I_c^i(t) and interaction frequency information entropy I_f^i(t) with each connected node:

$$MI_i(t) = I_c^i(t) \times I_f^i(t) \quad \text{Formula (5)}$$

**3.2.2 Indirect Influence** When calculating node indirect influence, we first consider whether nodes i and k connect through other nodes. Let K_{ik}(t) represent nodes connected to both i and k, i.e., K_{ik}(t) = K_i(t)  K_k(t). If no other nodes connect i and k, then K_{ik}(t) = 0. If one and only one node j connects i and k, then:

$$NI_{ik}(t) = (\theta + I_{ij}) \times \left[1/(\theta + I_{jk})\right] = (\theta + MI_i(t)) \times \left[1/(\theta + MI_j(t))\right]$$

where  represents the environmental influence of network nodes—the inherent influence when assuming a node is not connected to others. This study sets  = 1.

Paths connecting nodes i and k may pass through one node or several nodes. Therefore, when calculating indirect influence between nodes, we consider both single-node and multi-node connection scenarios.

**(1) Single-node connection.** If n nodes $j_1$, $j_2$, …, j  connect nodes i and k, and i and k can be connected through just one node (as shown in Figure 2 [Figure 2: see original paper]), then the indirect influence NI_{ik}(t) of node i on node k is:

$$NI_{ik}(t) = [\theta + I_{ij}(t)] \times \left[1/(\theta + I_{jk}(t))\right] + [\theta + I_{in}(t)] \times [1/(\theta + I_{mk}(t))] + [\theta + I_{in}(t)] \times [1/(\theta + I_{nk}(t))] + [\theta + I_{ip}(t)] \times \left[1/\right.$$

$$= [\theta + MI_i(t)] \times \left[1/(\theta + MI_j(t))\right] + [\theta + MI_i(t)] \times [1/(\theta + MI_n(t))] + [\theta + MI_i(t)] \times [1/(\theta + MI_n(t))] + [\theta + MI_i(t)] \times \left[\right.$$

$$= [\theta + MI_i(t)] \times \left[1/(\theta + MI_j(t))\right] \quad \text{Formula (8)}$$

**(2) Multi-node connection.** If nodes i and k cannot be connected through just one node and require multiple nodes for information transmission from i to k (as shown in Figure 3 [Figure 3: see original paper]), with possible paths being ($j_1$, $j_2$) or ($m_1$, $m_2$, $m_3$) or ($n_1$, $n_2$), then the indirect influence NI_{ik}(t) of node i on node k is:

$$NI_{ik}(t) = [\theta + I_{ij_1}(t)] \times \left[1/(\theta + I_{j_{1j}2}(t))\right] \times \left[1/(\theta + I_{j_{2k}}(t))\right] \quad \text{Formula (9)}$$

$$= [\theta + MI_i(t)] \times \left[1/(\theta + MI_{j_1}(t))\right] \times \left[1/(\theta + MI_{j_2}(t))\right] + [\theta + MI_i(t)] \times \left[1/(\theta + MI_{m_1}(t))\right] \times \left[1/(\theta + MI_{m_2}(t))\right] \times \left[1\right.$$

**3.2.3 Comprehensive Influence** Node comprehensive influence should represent the combined evaluation result of direct and indirect influence, indicating a node's overall role in information propagation. For example, if node i only connects to node j, but node j connects to hundreds or thousands of nodes, node i's direct influence may be small while its indirect influence is large, making its comprehensive influence substantial. Based on this, node i's comprehensive influence I_i(t) is expressed as:

$$I_i(t) = \alpha MI_i(t) + \beta NI_i(t) \quad \text{Formula (12)}$$

where weights $\alpha$ and $\beta$ for direct and indirect influence should satisfy $\alpha + \beta = 1$. Considering the overall evaluation capacity of direct and indirect influence, this study sets $\alpha = 0.6$ and $\beta = 0.4$.

### 3.3 Network Node Influence Construction Based on WeChat

According to iMedia Research data, WeChat official accounts led China's self-media platforms in 2016 with 63.4% market share, while Weibo self-media platforms became the second choice at 19.3%. Analysts believe that WeChat official accounts, as relatively closed self-media requiring following for content access, have stimulated the rise of new media platforms like Toutiao and UC. WeChat official accounts have become a major information dissemination channel and a development area for traditional media transformation. Since its launch in January 2011, WeChat official accounts have grown rapidly, exceeding 14 million by August 2016.

WeChat official account platforms provide business services primarily for enterprises, media, and governments. With functions including information push, comments, and related article links, these platforms have quickly become important channels for netizens to share and exchange information in new media environments, forming integrated information exchange platforms. Since article forwarding volumes, comments, and paths on WeChat official account platforms are only visible to operators, and considering privacy protection for WeChat (such as self-media or micro-businesses), relevant data is not completely public. Therefore, this study first uses the Sogou WeChat search engine with "NetEase News" as the keyword to obtain WeChat official accounts and article addresses, then employs crawler tools via the "Micro Index" platform for data collection. The study collected 245 articles from 38 official accounts published during March 21-24, 2017, selecting 7 closely connected accounts to form a fully connected network where each node serves as both a "source" and "sourced" node. Using Access and Excel for data processing, the initial network based on node and article source relationships is shown in Figure 4 [Figure 4: see original paper].

The seven official account samples are Xinhua Net, "I Was Shocked At That Time," Beijing News, NetEase News, Southern Metropolis Daily, Shanghai Morning Post, and Travel Headlines. Directed edges represent article source directions between WeChat official accounts, with numbers indicating source frequency. For example, Xinhua Net has 5 articles sourced from "I Was Shocked At That Time" and 2 articles sourced from Beijing News, representing interaction frequency between nodes.

## 4 Data Results and Discussion Analysis

### 4.1 Node Influence Calculation Results

**4.1.1 Direct Influence** When calculating direct influence of network nodes in new media environments, Java programming was used to compute connected node quantity information entropy and node interaction frequency information entropy for each node in the constructed WeChat official account network, with direct influence being the product of these two values. The results show nodes ranked from lowest to highest direct influence are: NetEase News, Beijing News, Xinhua Net, Southern Metropolis Daily, "I Was Shocked At That Time," Shanghai Morning Post, and Travel Headlines. NetEase News has the highest direct influence at 0.66037 because it has 6 directly connected nodes while other nodes have only 3. Consequently, NetEase News's connected node information entropy and interaction frequency information entropy values are significantly higher at 0.903089 and 0.731234, respectively. The remaining six nodes have relatively similar direct influence values because each has the same number of direct connections (3), resulting in identical connected node information entropy values of 0.421758. However, due to different interaction frequencies, their interaction frequency information entropy values vary, creating certain differences in direct influence values that all fall between 0.16-0.19. Travel Headlines has the lowest direct influence at 0.165348. Detailed statistics are shown in Table 1 .

**4.1.2 Indirect Influence** In the constructed WeChat official account network, there are 24 connecting lines between any two nodes. Except for NetEase News which has 6 directly connected nodes, all other nodes have 3 direct connections. The interaction frequency information entropy values between adjacent nodes are shown in Table 2 . The largest value is 0.159176 for articles from "I Was Shocked At That Time" sourced from Xinhua Net, while the smallest is 0.081866 for articles from Travel Headlines sourced from Shanghai Morning Post—the only propagation path with connection node information entropy below 0.1. Using Java programming to input connected node quantities and adjacent node interaction frequency information entropy values, the indirect influence of six nodes in the WeChat official account network was calculated, with results shown in Table 3 . NetEase News has the highest indirect influence value at 0.474152, followed in descending order by "I Was Shocked At That Time," Beijing News, Xinhua Net, Shanghai Morning Post, Southern Metropolis Daily, and Travel Headlines. All six official accounts have indirect influence values between 0.23-0.27.

**4.1.3 Comprehensive Influence** Comprehensive influence calculation results for network nodes in new media environments are shown in Table 4 (with direct influence weight coefficient $\alpha = 0.6$ and indirect influence weight coefficient $\beta = 0.4$). Nodes ranked by comprehensive influence from highest to lowest are: NetEase News, Beijing News, Xinhua Net, "I Was Shocked At That Time," Southern Metropolis Daily, Shanghai Morning Post, and Travel Headlines. NetEase News's comprehensive influence reaches 0.585883, far exceeding the other six nodes, making it the core node in this WeChat official account network. The other six nodes have comprehensive influence values between 0.19-0.22. The network constructed based on WeChat official account source data is extensible; when node quantities increase and connections become more complex, the established network node influence measurement model can process larger datasets, playing an important role in core node classification and opinion leader identification during network public opinion information propagation.

### 4.2 Simulation Analysis

MATLAB is a high-level technical computing language and interactive environment for algorithm development, data visualization, data analysis, and numerical computation, with strong capabilities for matrix operations, function and data plotting, and algorithm implementation. This study uses MATLAB to visualize variation trends of direct, indirect, and comprehensive influence through cloud maps.

**4.2.1 Direct Influence Simulation (1) Connected node quantity information entropy.** In MATLAB, the mathematical expression for network node influence information entropy in new media environments was converted to C language for programming. Using scatter(x,y) to plot the variation trend of connected node quantity information entropy (y-axis) with node quantity (x-axis) yielded the simulation result shown in Figure 5 [Figure 5: see original paper]. The figure shows that connected node quantity information entropy increases with the number of nodes connected to a network node, but the growth rate varies significantly with network scale. When connected node quantity is below 700, information entropy grows rapidly; above 700, growth slows. Trend lines were fitted separately for these two ranges, as shown in Figures 6(a) and 6(b) [Figure 6: see original paper]. When connected node quantity is below 700, the trend follows a standard logarithmic curve: y = 0.4343 ln(x) - 1E-13 ($R^2$ = 1). When connected node quantity exceeds 700, the trend follows: y = 0.00001x + 2.489 ($R^2$ = 0.997). For original nodes with few connections, connected node information entropy grows rapidly at rate 0.4343; with many connections, growth slows to below 0.001, with entropy values approaching 3. The $R^2$ value of 0.997 indicates high linear correlation between information entropy magnitude and node connection quantity when connections exceed 700. Thus, in new media information propagation, connected node quantity information entropy shows logarithmic growth when node connections increase from 0, but transitions to linear growth when connections exceed 700.

**(2) Interaction frequency information entropy.** MATLAB simulation of node interaction frequency information entropy in new media environments is shown in Figure 7 [Figure 7: see original paper]. Figure 7(a) shows that interaction frequency information entropy values are negatively correlated with node quantity. When network node quantity is below 200, interaction frequency information entropy values range between 0-0.16 with relatively dispersed distribution. When node quantity exceeds 200, values concentrate mainly between 0-0.02 with relatively dense distribution, showing a trend toward greater concentration and decline with increasing node quantity. After filtering out points with interaction frequency information entropy below 0.02 and arranging remaining values in descending order, a relatively standard power-law distribution curve emerges, as shown in Figure 7(b), with the polynomial formula $y = 0.5356x^{-0} \cdot {}^{6}$ ($R^2$ = 0.895). Therefore, in new media information propagation, node interaction frequency information entropy values decline with increasing node quantity. When node quantity is below 200, interaction frequency information entropy values are relatively large but do not exceed 0.16; when node quantity exceeds 200, these values approach 0. However, due to weak variation trends and similar numerical values, connected node quantity information entropy has greater impact on direct influence calculation. Thus, for government and public opinion management departments, establishing connections with more key nodes and conducting public opinion monitoring is more effective for improving direct influence than increasing control frequency over a single opinion leader.

**(3) Direct influence simulation.** Direct influence of network nodes in new media environments is the product of connected node quantity information entropy and interaction frequency information entropy. MATLAB simulation results are shown in Figure 8 [Figure 8: see original paper]. The figure demonstrates that direct influence is positively correlated with connected node quantity, with values becoming increasingly dispersed. The statistical relationship between node direct influence and connection quantity is shown in Figure 8(b). Peak points of original nodes' direct influence show linear increasing trends with connection quantity, with the fitted trend line formula y = 0.2406x - 1.485 ($R^2 = 0.873$). Using MATLAB with C language to input the direct influence information entropy expression and scatter(x,y,z) to plot direct influence information entropy (z-axis) variation with node quantity (x-axis) and interaction frequency (y-axis) yields the 3D simulation shown in Figure 9 [Figure 9: see original paper]. Node direct influence increases with both connected node quantity and interaction frequency, though most nodes show uniformly distributed and relatively low direct influence information entropy values. Therefore, in new media information propagation, original nodes with fewer connections and lower interaction frequency have smaller direct influence, while those with more connections and higher interaction frequency have larger direct influence. In public opinion management, governments and relevant departments should simultaneously increase connection frequency with key nodes and the number of managed nodes to optimally enhance control effectiveness. The threshold trend line also indicates that network node management has certain limits—beyond the threshold range, management of these nodes becomes ineffective.

**4.2.2 Indirect Influence Simulation** Based on direct influence calculation results, MATLAB simulation of indirect influence variation trends yields the results shown in Figure 10 [Figure 10: see original paper]. The statistical relationship between an original node's indirect influence information entropy value and connection quantity is shown in Figure 10(b). When an original node's connected node quantity is below 500, indirect influence information entropy values grow relatively quickly following a logarithmic distribution: y = 0.5702 ln(x) + 6E-12 ($R^2 = 0.892$). When connected node quantity exceeds 500, the growth rate slows, approaching 120. The 3D simulation of an original node's indirect influence variation with node quantity and interaction frequency is shown in Figure 11 [Figure 11: see original paper]. Node indirect influence increases with connected node quantity and interaction frequency. However, when calculating comprehensive influence, if an original node's connected node quantity exceeds 700, its indirect influence has a lower impact index on comprehensive influence. Therefore, in new media information propagation, governments and relevant departments can guide or control information indirect propagation by increasing the number of managed nodes and management frequency to improve indirect influence. However, when managed node quantity and management frequency exceed certain ranges, their effect on improving indirect influence becomes minimal, and management focus should shift to enhancing direct influence.

**4.2.3 Comprehensive Influence Simulation** In MATLAB, setting direct influence weight $\alpha = 0.6$ and indirect influence weight $\beta = 0.4$ yields the comprehensive influence simulation shown in Figure 12 [Figure 12: see original paper]. This figure is similar to the direct influence simulation but with different numerical variation amplitude. The statistical relationship between an original node's comprehensive influence information entropy value and connection quantity is shown in Figure 12(b). Node comprehensive influence is positively correlated with connected node quantity, with increasingly dispersed value distribution. The fitted trend line formula is y = 9.391x^0.2634 - 12.48 ($R^2$ = 0.7218). In new media environments, network node comprehensive influence grows rapidly when connection quantity is small, but its growth rate slows as connections increase.

The simulation of comprehensive influence variation with connected node quantity and interaction frequency is shown in Figure 13 [Figure 13: see original paper]. As original node connection quantity and interaction frequency increase, comprehensive influence also increases, though interaction frequency changes have weaker impact than connection quantity changes. In MATLAB, treating direct and indirect influence as two variables yields the scatter simulation showing comprehensive influence variation with both variables. Under the combined effect of direct and indirect influence, comprehensive influence shows an upward trend. When an original node's indirect influence is small, direct and indirect influence have similar effect magnitudes. However, when indirect influence exceeds 100, it has minimal impact on comprehensive influence, and direct influence becomes the primary factor affecting network node comprehensive influence in new media environments. Therefore, when node quantity and interaction frequency exceed certain ranges in a network, public opinion managers should prioritize direct influence results for node influence identification and classification.

## 5 Research Conclusions

At the theoretical level, this study constructs a network node influence model for new media environments based on information entropy theory, determining mathematical expressions for connected node quantity information entropy and interaction frequency information entropy, and establishing mathematical representations for direct influence information entropy, indirect influence information entropy, and comprehensive influence information entropy. MATLAB software was used for model simulation and systematic analysis of results. Data show that as connected node quantity and interaction frequency increase in new media environments, both direct and indirect influence increase, though at different rates with direct influence growing faster. Comprehensive influence also increases with connection quantity and interaction frequency, growing rapidly when node quantity is small but slowing as node quantity continues to increase beyond a certain range. When indirect influence information entropy reaches 100, it no longer significantly affects comprehensive influence, and direct influence becomes the primary factor. This research provides new theoretical foundations and model analysis methods for studying network node influence in new media environments.

At the practical level, using WeChat official accounts as a case study, this paper draws cloud maps to predict influence variation trends as node quantity increases, providing strong support for governments and relevant departments to better monitor network information dissemination in new media environments. Governments and managers can use the proposed network node influence model to measure node influence magnitude, monitor changes in node quantity and interaction frequency in real-time, focus management on high-influence nodes, and thereby effectively monitor and manage information propagation to prevent the spread of harmful information in new media environments.

This study has certain limitations, having only selected source data from 7 WeChat official accounts to measure network node influence in new media environments, resulting in limited sample coverage. Future research will select larger sample datasets, such as Weibo forwarding and comment data, to validate the model and conduct early warning analysis for information propagation in new media environments, comparing similarities and differences in information dissemination influence across different new media environments. Additionally, further exploration of information entropy theory applications in information propagation and other aspects will be pursued.

## References

[1] China Internet Network Information Center. The 39th Statistical Report on China's Internet Development Status [EB/OL]. [2017-01-22]. http://www.cnnic.net.cn/hlwfzyj/hlwxzbg/hlwtjbg/201701/t20170122_{66437}.htm.

[2] Kim K, Baek YM, Kim N. Online news diffusion dynamics and public opinion formation: a case study of the controversy over judges' personal opinion expression on SNS in Korea [J]. The social science journal, 2015, 52(2): 205-216.

[3] Miorandi D, Pellegrini FD. K-shell decomposition for dynamic optimization in mobile, ad hoc and wireless networks [C]//Proceedings of the international symposium on modeling and optimization in mobile, ad hoc and wireless networks. Paris: IEEE, 2010: 488-496.

[4] Lyu L, Zhang Y, Yeung CH. Leaders in social networks, the delicious case [J]. PLoS ONE, 2011, 6(6): e21202.

[5] Deanne K, Katharine H, Honert R, et al. Nuclear power in Australia: a comparative analysis of public opinion regarding climate change and the Fukushima disaster [J]. Energy policy, 2014, 65: 644-653.

[6] Arnaboldi V, Conti M, Lagala M. Ego network structure in online social networks and its impact on information diffusion [J]. Computer communications, 2016, 76: 26-41.

[7] Cao Xueyan, Duan Feifei, Fang Kuan, et al. Research on identification and classification of key nodes in emergency public opinion from the perspective of network forums [J]. Library and Information Service, 2014, 58(4): 65-71.

[8] Kang Wei. Measurement and analysis of social network structure for public opinion dissemination in emergencies—an empirical study based on the "11・16 School Bus Accident" [J]. China Soft Science, 2012(7): 169-178.

[9] Jiang Kan, Tang Zhufa, Sui Hao. Identification of key nodes in network public opinion based on information diffusion quality in Weibo [J]. Information Science, 2016, 34(7): 64-69.

[10] Wang Yuefen, Hang Wei, Liang Dingjie. Identification and application of key nodes in social networks for Weibo public opinion [J]. Information Studies: Theory & Application, 2016(3): 6-11.

[11] Shannon CE. A mathematical theory of communication [J]. Bell system technical journal, 1948, 27: 379-423. doi:10.1002/j.1538-7305.1948.tb01338.x.

[12] Sangam R, Om H. The k-modes algorithm with entropy-based similarity coefficient [J]. Computer science, 2015, 50: 93-98.

[13] Li Y, Sha Y, Shan J. The research of weighted community partition based on SimHash [J]. Computer science, 2013, 17: 797-801.

[14] Wei Zhihui, He Yue. Identification of Weibo opinion leaders based on information entropy and unascertained measurement model—a case study of the "Gansu Qingyang School Bus Emergency" [J]. Information Science, 2014, 32(10): 38-43.

[15] Chen Yuan, Li Yunhui, Zhang Min. Empirical research on measuring SNS user information propagation contribution based on node degree—taking Tencent Weibo as an example [J]. Journal of Intelligence, 2014, 33(10): 159-164.

[16] Manku GS, Jain A. Detecting near-duplicates for Web crawling [M]. New York: ACM, 2007.

[17] Charikar MS. Similarity estimation techniques from rounding algorithms [M]. New York: ACM, 2002.

[18] Wu Peng, Wang Hengshan. Evaluation of important nodes in social information supernetworks based on eigenvector centrality [J]. Information Studies:

Theory & Application, 2014, 37(5): 107-113.

[19] Wang Xiwei, Xing Yunfei, Zhao Dan, et al. Research on network public opinion information propagation paths and patterns in mobile environments [J]. Information Studies: Theory & Application, 2016, 39(9): 107-113.

[20] Kim K, Baek YM, Kim N. Online news diffusion dynamics and public opinion formation: a case study of the controversy over judges' personal opinion expression on SNS in Korea [J]. The social science journal, 2015, 52(2): 205-216.

[21] Wang Xiwei, Xing Yunfei, Zhao Dan, et al. Research on network public opinion information propagation based on social network analysis in mobile environments—taking the "haze" topic on Sina Weibo as an example [J]. Library and Information Service, 2015, 59(7): 14-22.

[22] Peng S, Yang A, Cao L. Social influence modeling using information theory in mobile social networks [J]. Information science, 2017, 379: 146-159.

[23] iMedia Report. 2017 China New Media Industry Panorama Report [EB/OL]. [2017-03-01]. http://www.iimedia.cn/50347.html.

[24] Chinese Social Sciences Net. China New Media Development Report [EB/OL]. [2017-09-01]. http://www.cssn.cn/zx/bwyc/201706/t20170626_{3560419}1.shtml.

[25] Fang Jing, Lu Wei. Empirical study on factors influencing information propagation heat of WeChat official accounts [J]. Journal of Intelligence, 2016, 35(2): 157-162.

## Author Contributions

Xing Yunfei: Paper writing and revision;
Wang Xiwei: Research proposition and framework design, final paper revision;
Han Xuewen: Data collection and processing;
Zhang Changliang: English literature collection and abstract translation.

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