

A Weibo local network growth model constructed from the perspective of following-followed

Fu-Zhong Nian(年福忠)† and Ran-Qing Yao(姚然庆)
Citation: Chin. Phys. B, 2024, 33 (**12**): 128702. DOI: 10.1088/1674-1056/ad84c4
Journal homepage: http://cpb.iphy.ac.cn; http://iopscience.iop.org/cpb

What follows is a list of articles you may be interested in

Effects of individual heterogeneity on social contagions

Fu-Zhong Nian(年福忠) and Yu Yang(杨宇)
Chin. Phys. B, 2024, 33 (**5**): 058705. DOI: 10.1088/1674-1056/ad225e

Dynamics of information diffusion and disease transmission in time-varying multiplex networks with asymmetric activity levels

Xiao-Xiao Xie(谢笑笑), Liang-An Huo(霍良安), Ya-Fang Dong(董雅芳), and Ying-Ying Cheng(程英英)
Chin. Phys. B, 2024, 33 (**3**): 038704. DOI: 10.1088/1674-1056/ad1176

Studying the co-evolution of information diffusion, vaccination behavior and disease transmission in multilayer networks with local and global effects

Liang'an Huo(霍良安) and Bingjie Wu(武兵杰)
Chin. Phys. B, 2024, 33 (**3**): 038702. DOI: 10.1088/1674-1056/ad0114

Epidemic threshold influenced by non-pharmaceutical interventions in residential university environments

Zechao Lu(卢泽超), Shengmei Zhao(赵生妹), Huazhong Shu(束华中), and Long-Yan Gong(巩龙延)
Chin. Phys. B, 2024, 33 (**2**): 028707. DOI: 10.1088/1674-1056/ace2b0

Stability and optimal control for delayed rumor-spreading model with nonlinear incidence over heterogeneous networks

Xupeng Luo(罗续鹏), Haijun Jiang(蒋海军), Shanshan Chen(陈珊珊), and Jiarong Li(李佳容)
Chin. Phys. B, 2023, 32 (**5**): 058702. DOI: 10.1088/1674-1056/acb490

A Weibo local network growth model constructed from the perspective of following–followed

Fu-Zhong Nian(年福忠)[†] and Ran-Qing Yao(姚然庆)

School of Computer and Communication, Lanzhou University of Technology, Lanzhou 730050, *China*

(Received 9 July 2024; revised manuscript received 5 September 2024; accepted manuscript online 9 October 2024)

In order to explore the evolution process of the Weibo local network, this study first defines four factors influencing the evolution of the Weibo network. On this basis, the BA scale-free network model was enhanced by incorporating these four factors and accounting for directionality, resulting in a Weibo local network evolution model based on user attributes and behavioral similarity. The model's validity was validated by comparing simulation results with real data. The findings indicate that the Weibo local network exhibits both small-world characteristics and distinctive features. The results show that the Weibo local network exhibits both small-world characteristics and distinctive properties. The in-degree distribution follows a mixed pattern of exponential and power-law distributions, the degree-degree shows isomatching, and both the in-degree centrality and eigenvector centrality values are relatively low. This research contributes to our understanding of user behaviour in the Weibo network, and provides a structural basis for exploring the impact of Weibo network structure on information dissemination.

Keywords: Weibo network, user similarity, attribute characteristics, network evolution, topology analysis

PACS: 87.23.Ge, 87.23.Kg, 05.90.+m, 89.75.–k **DOI:** 10.1088/1674-1056/ad84c4

1. Introduction

The small-world network model, as described in Ref. [1], illustrates the phenomenon of high clustering coefficients and short average path lengths observed in the real world. Barabasi and Albert proposed a scale-free network model based on growth and preferential attachment, detailed in Ref. [2]. Subsequently, numerous extension models of BA scale-free networks were proposed.[3–8] One such a model, a local world network model[3] was proposed, which allows for the adjustment of local world scale parameters, thereby enabling the degree distribution to vary between index and power law, parameters, thereby enabling the degree distribution to vary between index and power law. Yijun Ran *et al*. discussed link prediction techniques in network information mining, with a focus on long-path networks. They introduce a new SESPL index that significantly improves the accuracy of link prediction in long-path networks and outperforms existing similarity-based predictors in tests across 548 real-world networks. This research provides fresh insights into the mechanisms of network evolution and highlights the potential of SESPL as an independent feature within traditional similarity-based models.[9] In Ref. [10], Han and Qiao proposed a spatially weighted network model based on expected traffic, incorporating the effects of geographic factors and node intrinsic fitness. This model aims to maximize the total expected traffic within the network. By varying two parameters, control fitness and geographic constraints, the model can modify its topological structure and demonstrate a range of statistical properties observed in real-world networks.[10] Bianconi and Barabási proposed a fitness model[5] whereby the final selection and connection probability is proportional to the product of node fitness and degree. Considering the normal distribution of fitness and the exponential growth of the number of nodes, a stochastic network model[6] was introduced, which can produce a double power degree distribution. Social networks have been widely studied within the framework of complex networks.[11–15] In recent years, with the rapid development of online social networks (OSNs), there has been a great deal of researches conducted on the structure and function of these networks,[16–22] the social structure of the Facebook network has been studied, including the classification coefficient and community structure based on different user characteristics.[19] In order to reveal the formation mechanism of OSNs, many models have been proposed. In Ref. [23], the authors roposed a simple OSNs model and concluded that it is important to consider different user categories. Based on the local world evolutionary algorithm, an updated evolutionary model[24] is proposed, which can describe the same features as the OSNs features in real life. Weibo is a type of OSN with complex network characteristics, that attracts a large number of users to participate. Previous studies have mostly focused on the topology characteristics and information dissemination.[25–34] Zhang *et al*. investigated user participation and social influence during information dissemination on Twitter. They developed a measurement method incorporating three metrics to evaluate users' efforts in spreading information and found that a small fraction of users with exceptional participation can exert significant influence, while

[†]Corresponding author. E-mail: gdnfz@lut.edu.cn

© 2024 Chinese Physical Society and IOP Publishing Ltd

the majority of users act as intermediaries in the propagation process. This study underscores that existing topological measures may not fully capture individual influence and offers new insights into the dynamics of information spreading.[35] There is little research conducted on the real development process and internal mechanism of the local Weibo user relationship network. In order to describe the evolution law of local Weibo network more accurate, this paper proposes a network evolution model based on user attribute and behaviour feature similarity. The simulation results show that the model is capable of better fitting various topological characteristics of actual networks.

2. Construction of network structure

2.1. A Weibo network model with two relationships

Abstracting users in Weibo into nodes in the network and representing relationships between users as edges forms a complex network. Following others is represented by the in-degree of a node in the network (according to the direction of information flow), being followed is represented by the out-degree of a node, and mutual following is represented by a two-way connection. As shown in Table 1.

Table 1. Two relationships.

Type	Relationship
A $\rightarrow$ B	user B follows user A
B $\leftrightarrow$ C	user B and user C follow each other

2.2. Model hypothesis

The construction of a network evolution model necessitates the consideration of four types of changes in the network: node addition, node removal, the addition of edges, and the removal of edges. The following assumptions are proposed in this article.

Node addition New nodes continue to join the network, which is represented by new users signing up for Weibo accounts in the real Weibo network. Therefore, when constructing the Weibo network evolution model, it is assumed that a new node joins the Weibo network at each iterative point in time, or in other words, a new user registers for Weibo at each point in time.

Node exit Existing nodes in the network leave the network, as evidenced by the user's account cancellation resulting in their no longer being logged in to the Weibo network. In general, once a person starts using Weibo, they are unlikely to cancel their account; thus, the phenomenon of node exit is not considered when building the Weibo network model.

Addition of connected edges There are two main scenarios for increasing connected edges. One approach is to add edges between new nodes and existing nodes. The other is to add edges between existing nodes in the network. In the Weibo network, this typically reflects users following others.

Removal of edges Removing existing edges in the network. In the Weibo network, the likelihood of a user unfollowing another user is low. From a psychological perspective, fans are loyal audience groups with a high degree of stability. Therefore, the probability of users unfollowing another user is very small, and disconnected edges are not considered when establishing the Weibo network evolution model.

Both mutual and unidirectional concerns are present in Weibo networks. Based on the above model assumptions, a directed network is constructed, which solely considers the phenomenon of nodes joining and adding edges.

2.3. Factors affecting the formation of Weibo network

2.3.1. User attribute similarity

For a user u in Weibo, the user information contains a variety of attribute information, which is marked as follows:

$$p = [n_1, n_2, n_3, n_4, n_5, \ldots, n_m],$$

where $n_1, \ldots, n_m$ represent the attribute information of user u.

For two Weibo users u and v, the user attribute similarity is calculated according to the attribute information between them. The calculation formula is as follows:

$$\begin{aligned}\mathrm{sim}\,(p_u, p_v) \;=\; & \omega_1 \mathrm{sim}\,(u_{n1}, v_{n1}) + \omega_2 \mathrm{sim}\,(u_{n2}, v_{n2}) + \cdots \\ & + \omega_m \mathrm{sim}\,(u_{nm}, v_{nm}),\end{aligned} \tag{1}$$

where u_{n1} represents the first attribute information of user u, ω_1 represents the weight of the first attribute information, and so on.

2.3.2. The similarity of social relationships

For one user u on Weibo, his followers and fans are recorded as sets U and V, respectively:

$$U = \{u_1, u_2, \ldots, u_n\},$$
$$V = \{v_1, v_2, \ldots, v_m\}.$$

The social relationship similarity between users u and v is calculated according to the fan and follow information between the two. The calculation formula is as follows:

$$\mathrm{sim}_{\mathrm{soc}}\,(u, v) = \varphi_1 \mathrm{sim}_{\mathrm{fe}}\,(u, v) + \varphi_2 \mathrm{sim}_{\mathrm{fan}}\,(u, v), \tag{2}$$

where φ_1 represents the weight of the following similarity, and φ_2 represents the weight of the follower similarity. $\mathrm{sim}_{\mathrm{fe}}\,(u, v)$ and $\mathrm{sim}_{\mathrm{fan}}\,(u, v)$ represent the following similarity and the follower similarity, respectively,

$$\mathrm{sim}_{\mathrm{fe}}\,(u, v) = \frac{|U_u \bigcap U_v|}{\log |U_u|^2 \times \log |U_v|^2}, \tag{3}$$

$$\mathrm{sim}_{\mathrm{fan}}\,(u, v) = \frac{|V_u \bigcap V_v|}{\log |V_u|^2 \times \log |V_v|^2}, \tag{4}$$

where U_u is the follow set of user u, and V_u is the fan set of user u.

2.3.3. Interaction intimacy between users

For users u and v, the greater number of likes, retweets, and comments between u and v, the more likely a relationship is to be established between u and v. The interaction density between u and v, denoted as $\mathrm{ina}(u,v)$, represents the degree of interaction closeness between the two users, reflected in a value within the range $[0,1]$. A higher value indicates a higher degree of interaction density. The formula to calculate the interaction density between users u and v is as follows:

$$\mathrm{ina}(u,v) = 1 - \mathrm{e}^{-r}, \tag{5}$$

where,

$$\begin{aligned} r = {} & \vartheta_1 \frac{R_n(u\to v)+R_n(v\to u)}{2} \\ & + \vartheta_2 \frac{C_n(u\to v)+C_n(v\to u)}{2} \\ & + \vartheta_3 \frac{L_n(u\to v)+L_n(v\to u)}{2}, \end{aligned}$$

$R_n(u\to v)$ represents the number of Weibo retweets of user u to user v, $C_n(u\to v)$ represents the number of Weibo comments of user u to user v, $L_n(u\to v)$ represents the number of Weibo likes of user v by user u.

If we define

$$\begin{aligned} R &= R_n(u\to v)+R_n(v\to u), \\ C &= C_n(u\to v)+C_n(v\to u), \\ L &= L_n(u\to v)+L_n(v\to u), \end{aligned}$$

then,

$$\vartheta_1 = \frac{R}{R+C+L}, \quad \vartheta_2 = \frac{C}{R+C+L}, \quad \vartheta_3 = \frac{L}{R+C+L}.$$

2.3.4. Initial attraction of nodes

New nodes joining the network receive different degrees of attention. For example, celebrities joining the Weibo network, will inevitably receive widespread attention, while ordinary people join Weibo network are unlikely to attract significant attention. To describe this phenomenon, the initial attraction of nodes is defined. The initial attraction of a node is drawn from the power-law distribution, and the calculation formula is as follows:

$$A_i = \left(\frac{x}{x_{\min}}\right)^{-\delta}, \tag{6}$$

where A_i is the initial attraction of the node, $0 < x < 1$, $0 < x_{\min} < 1$. x, $x_{\min}$, and δ are parameters used to adjust the power-law distribution.

Due to the characteristics of power-law distribution, this distribution will reflect the phenomenon that the initial attraction of most nodes is low, while the initial attraction of a few nodes is very high.

2.4. Network topology model considering four user attribute information

For a user u on Weibo, their user information contains a variety of attribute information. Here, we only consider four kinds of attribute information, recorded as

$$p_u = [n_1, n_2, n_3, n_4],$$

where n_1 represents the domain information of u; n_2 represents the interest tag of u; n_3 represents the location information of u; n_4 represents the age information of u. Then, for two Weibo users (u,v), the formula for calculating user attribute similarity is

$$\begin{aligned} \mathrm{sim}(p_u,p_v) = {} & \omega_1 \mathrm{sim}(u_{n_1}, v_{n_1}) + \omega_2 \mathrm{sim}(u_{n_2}, v_{n_2}) \\ & + \omega_3 \mathrm{sim}(u_{n_3}, v_{n_3}) + \omega_4 \mathrm{sim}(u_{n_4}, v_{n_4}), \end{aligned} \tag{7}$$

where $\mathrm{sim}(u_{n_1}, v_{n_1})$ represents the domain similarity between u and v, ω_1 represents its weight; $\mathrm{sim}(u_{n_2}, v_{n_2})$ represents the similarity of interest between u and v, ω_2 represents its weight; $\mathrm{sim}(u_{n_3}, v_{n_3})$ represents the similarity of location information between u and v, ω_3 represents its weight; $\mathrm{sim}(u_{n_4}, v_{n_4})$ represents the age similarity between u and v, ω_4 represents its weight.

In order to calculate domain similarity, the domain tags officially provided by Sina Weibo are classified by relevance as shown in Table 2.

Table 2. Domain label classification.

Category	Tags
Technology	internet technology, science popularization, digital, science fiction
Culture and entertainment	music, film, TV drama, anime, drama, crosstalk
Sports	sports, esports, athletics
Art	photography, dance
Fashion and beauty	fashion, design aesthetics, cosmetics
Health and psychology	health, psychology
Education	education, reading materials
Entertainment	Korean star, variety show
Transportation	aviation, automobile
Culture and arts	traditional Chinese culture, calligraphy, Chinese painting
History	history
Lifestyle	cuisine, travel
Animal	adorable pets
Emotional	emotions
Legal	law
Collection	collectibles
Real estate	property
Maternity and parenting	maternity and parenting
Public welfare	public welfare
Combat sports	combat sports, martial arts

If classification is considered as set K, and label is considered as set T. Then, let

$$K = \{k_1, k_2, \ldots, k_{20}\}, \quad T = \{t_1, t_2, \ldots, t_{20}\},$$

then,

$$\mathrm{sim}\,(u_{n_1}, v_{n_1}) = \begin{cases} 1, & k_i = k_j \ \ \& \ \ t_i = t_j, \\ 0.5, & k_i = k_j \ \ \& \ \ t_i \neq t_j, \\ 0, & k_i \neq k_j \ \ \& \ \ t_i \neq t_j, \end{cases} \tag{8}$$

where $\mathrm{sim}(u_{n_1}, v_{n_1})$ represents the domain similarity between u and v. k_i and k_j represent the domains of the users u and v, t_i and t_j correspond to the labels in k_i and k_j respectively.

The $\mathrm{sim}\,(u_{n_2}, v_{n_2})$ represents the similarity of interest between u and v. If the interest of user u is considered as set I_u, then let

$$I_u = \{i_1, i_2, \ldots, i_n\},$$

then,

$$\mathrm{sim}\,(u_{n2}, v_{n2}) = \frac{|I_u \bigcap I_v|}{|I_u \bigcup I_v|}. \tag{9}$$

The $\mathrm{sim}\left(u_{n_3}, v_{n_3}\right)$ represents the similarity of location information between u and v. In Weibo, the location information is a "province/city" structure, where "province" is represented as set P and "city" is represented as set C, then let

$$P = \{p_1, p_2, \ldots, p_{34}\}, \quad C = \{c_1, c_2, \ldots, c_{339}\},$$

then,

$$\mathrm{sim}\left(u_{n_3}, v_{n_3}\right) = \begin{cases} 1, & p_i = p_j \ \ \& \ \ c_i = c_j, \\ 1/3, & p_i = p_j \ \ \& \ \ c_i \neq c_j, \\ 0, & p_i \neq p_j \ \ \& \ \ c_i \neq c_j, \end{cases} \tag{10}$$

where $\mathrm{sim}\,(u_{n_3}, v_{n_3})$ represents the similarity of location information between u and v. p_i and p_j denote the provinces of the users u and v, c_i and c_j correspond to the cities of p_i and p_j, respectively.

The $\mathrm{sim}\,(u_{n4}, v_{n4})$ represents the age similarity between u and v. The formula for calculating the similarity of age is as follows:

$$\mathrm{sim}\left(u_{n4}, v_{n4}\right) = 1 - \frac{\left|a_u - a_v\right|}{|d|}, \tag{11}$$

where d indicates the maximum age difference, and a_u represents the age of the user u.

In order to construct a network model that adhers to the evolutionary characteristics of the Weibo network, the evolution idea of the BA scale-free network is applied, and the many above-mentioned factors that affect the formation of the Weibo network are taken into account. The growth model is as follows:

The first step, Initial Network: The initial network has I_0 nodes, To ensure network connectivity, these nodes are connected end to end, and the network size is set to N according to the research needs.

The second step, Growth Mechanism: Each time a new node i and $n+m$ new edges are added to the original network, and when each new node i joins the network, bars are added to the edge from the existing node to the new node. m bars are added to the edge from the existing node to the new node i. n edges from the new node i to the existing nodes in the network. The direction of the arrow represents the direction of information dissemination. If the m arrows point to the new node i, it means that the i node pays attention to these nodes and becomes their fan. when these nodes release the dynamic, the node i will receive the message.

(i) The probability of the new node i being preferentially connected to the original node j in the network, when adding m edges where the existing nodes in the network point to the new node i, can be expressed as

$$p_{j\to i} = \left(\frac{k_j}{\sum\limits_{t\in V} k_t}\right)^{\alpha} + \mathrm{sim}\,(p_i, p_j), \tag{12}$$

where k_j is the out-degree of the original node j; $\sum_{t\in V} k_t$ represents the sum of the out-degrees of all nodes in the network; α is an adjustable parameter used to reflect nonlinear preferential attachment; $\mathrm{sim}\,(p_i, p_j)$ is the attribute similarity between the newly added node i and the existing node j.

(ii) When adding n edges from the new node i to the existing nodes in the network, the probability of the existing node j in the network being connected to the new node i can be expressed as

$$p_{i\to j} = A_i + \mathrm{sim}\,(p_i, p_j), \tag{13}$$

where A_i is the initial attractiveness of the newly added node i; $\mathrm{sim}\,(p_i, p_j)$ denotes the attribute similarity between the newly added node i and the existing node j.

The third step, Development Mechanism: For the nodes i and j that exist in the network, add two-way edges between the two nodes with probability p_{ij}

$$\begin{aligned} p_{ij} \ = \ & \beta_1 \times \mathrm{sim}_{\mathrm{soc}}\,(i,j) + \beta_2 \times \mathrm{ina}\,(i,j) \\ & + \beta_3 \times \mathrm{sim}\,(p_i, p_j), \end{aligned} \tag{14}$$

where $\mathrm{sim}_{\mathrm{soc}}\,(i,j)$ represents the similarity of social relationships between users i and j; $\mathrm{ina}\,(i,j)$ denotes the interaction density between users i and j; $\mathrm{sim}\,(p_i, p_j)$ represents the attribute similarity between users i and j; β_1, β_2, and β_3 are weight parameters.

Repeat steps 2 and 3 until the network grows to the desired size. That is, the total number of nodes in the network is N.

3. Experiment and result analysis

3.1. Weibo data

In order to verify the rationality of the evolution model, we obtained the relationship data of the real Weibo network for analysis and verification.

The data set was obtained from Sina Weibo (www.weibo.com). Based on Sina Weibo API, the snowball sampling method was employed to collect user data and relationships. A specific user was selected as the starting point to obtain personal information, messages, and relationships, and then the dataset of their friends and followers was crawled. this process was repeated until the desired number of users has been reached. Profile information includes the user ID, number of friends, followers, and fans. The relationship between Weibo users is the format of the user's follow object ID and user ID. Finally, the original data set was obtained. We selected four users as the sampling starting point and obtained four original datasets: A, B, C, and D, as shown in Table 3.

Table 3. Original dataset.

Dataset	Users	Relationships
A	108895	229999
B	115841	240189
C	125479	250140
D	130655	250001

Due to the large amount of data collected and the limitations of Weibo data collection, the original user relationship data in Table 3 cannot build a complete and closed network. In order to facilitate the analysis of the topological characteristics of the Weibo network, the original data set was further processed. In this step, we refer to the processing method in literature.[36] Four connected networks were constructed. Isolated nodes are removed, forming a huge network. This process resulted in four networks with 1078, 1086, 1021, and 1056 nodes respectively, which are recorded as Weibo networks a, b, c, and d, respectively. The basic topological properties of the four networks are shown in Table 4, where V represents the number of nodes, E is the number of edges, K is the average degree, C is the clustering coefficient, L is the average path length, M is the density. D is the diameter.

Table 4. Basic topological properties of Weibo networks a, b, c, d.

Network	a	b	c	d
V	1078	1086	1021	1056
E	15401	11508	8300	8690
K	14.3	10.6	8.1	8.2
C	0.062	0.059	0.072	0.127
L	3.21	3.57	3.86	4.12
M	0.0132	0.0098	0.0080	0.0078
D	9	11	10	11

Taking Weibo network a as an example, the average path length is 3.21 and the diameter is 9. This means that, on average, information between any two users passes through 3.21 people, and a maximum of 9 people can connect any two users. Weibo network a exhibits a shorter average path length, consistent with the six-degree segmentation theory, and displays small-world characteristics. The relatively short average path length of Weibo indicates that information disseminates more rapidly on this platform compared to traditional media and social networks.

3.2. Simulation

This section analyzes and compares the statistical characteristics of four real networks and four simulated networks. The number of nodes in the simulated networks is set to match the real networks in Table 4. The remaining parameters of the simulated networks are specified as follows: $N_a = 1078$, $N_b = 1086$, $N_c = 1021$, $N_d = 1056$, $\omega_1 = 2$, $\omega_2 = 2$, $\omega_3 = 1$, $\omega_4 = 1$, $\varphi_1 = 1$, $\varphi_2 = 1$, $\delta = 2.5$, $\alpha = 1.2$, $\beta_1 = 0.3$, $\beta_2 = 0.3$, $\beta_3 = 0.3$.

The above parameters are set for the following reasons. Because the research purpose of this article is to explore the evolution process of the local Weibo network, the goal is to propose a model that requires the simulated network generated by this model to be close to the network generated by real data in terms of network characteristics. In order to achieve this, the network characteristics of the real network were counted before the simulation experiment. Based on these characteristics, some experience values are assigned during simulation and adjustments are made continuously. In the process of continuous adjustment, different parameters will generate different networks, but our goal is to find suitable model parameters so that the generated network is close to the real Weibo network. Therefore, the final selection is the most appropriate parameter after continuous adjustment. Make the network generated by simulation close to the real network.

3.2.1. Global properties

The initial network lacks any distinct network characteristics. As the edges and nodes increase, the network structure gradually develops certain characteristics. The growth of edges in the network is driven by the following factors: new nodes preferentially connect to existing nodes; existing nodes form edges with newly added nodes based on their similarity and attractiveness; and internal nodes establish two-way edges based on the comprehensive influence of social relationship similarity, interaction closeness, and attribute similarity. Therefore, edge growth does not solely originate from newly added nodes; it also arises from interactions between existing nodes. This reflects the actual behavior of Weibo network users.

Table 5 shows the topological statistics of the simulated network generated by the model. The results demonstrate a high degree of consistency with the real network in terms of several topological characteristics. In the following sections,

we will compare the following topological properties of real and simulated networks: degree distribution, degree–degree correlation, average shortest path length distribution, and centrality.

Table 5. Topological statistics of simulated networks.

Network	a	b	c	d
V	1078	1086	1021	1056
E	13482	10240	8052	8357
K	12.5	9.43	7.96	8.0
C	0.017	0.014	0.012	0.012
L	3.21	3.46	3.72	3.77
M	0.0116	0.0087	0.0079	0.0076
D	6	6	7	7

3.2.2. Degree distribution

The most basic topological characteristics of the network can be expressed by its degree distribution $P(k)$. Degree distribution analysis describes the degree of interaction between users. As shown in Fig. 1, the degree distribution curve $P(k)$ follows a mixed distribution that adheres to both exponential and power laws, which is consistent with empirical data and previous theoretical research. The CCDF (cumulative distribution function) curve approximately follows a power law distribution. As shown in Fig. 2, In this case, the curve is steep at the beginning, indicating that there are a large number of nodes with lower degrees in the network. As the degree increases, the curve gradually becomes flatter, indicating that there are relatively fewer nodes with higher degrees. This pattern means that there are a few popular users who receive widespread attention in the Weibo network, while most users have a small number of fans or followers. This phenomenon aligns with real-world observations.

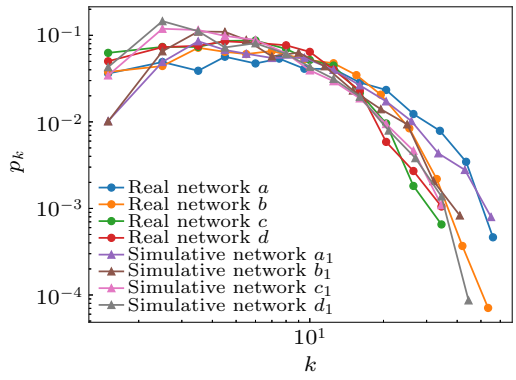


Fig. 1. Entrance distribution.

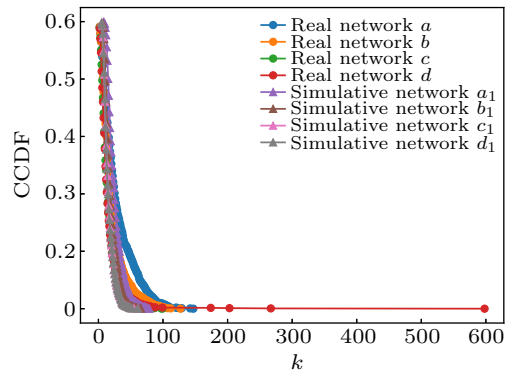


Fig. 2. CCDF curve distribution.

3.2.3. Degree–degree correlation

The degree–degree correlation reveals the characteristics of the second-order degree distribution of the network. Several methods can capture this correlation. One method for obtaining the degree–degree correlation is to examine the properties of the degree correlation function $K_{n,n}$. Defined as

$$K_{n,n}(k_i) = \frac{\sum_{j\in\Gamma(i)} k_{j,i}}{k}, \tag{15}$$

where $\Gamma(i)$ represents the set of neighbors of node i. If this function is increasing, the network exhibits assortativity. Conversely, if the function is decreasing, the network is disassortative. The degree k_i, $K_{n,n}$ can be plotted on the graph to describe the overall assortativity trend of the network. Real social networks (such as Facebook) are often associated with assortativity.[20] Figure 3 displays an increasing trend, indicating assortativity among Weibo users in real life, meaning that most users tend to establish relationships with others of equal or greater status. Additionally, the simulated network also exhibits assortativity, which is consistent with the real Weibo user network.

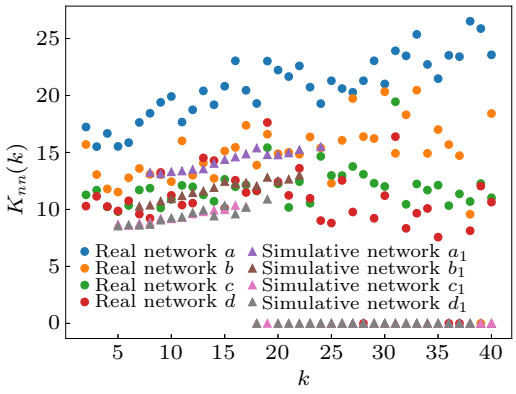


Fig. 3. Out-degree–degree correlation.

3.2.4. Average shortest path length distribution

The average shortest path length of a network is defined as the average shortest distance between any two nodes. Figure 4 shows the shortest path length distribution for both the real and simulated networks. In the real network, most distances are between 3 and 5, which demonstrates the small-world nature of the local Weibo network. This characteristic is also evident in the simulated network.

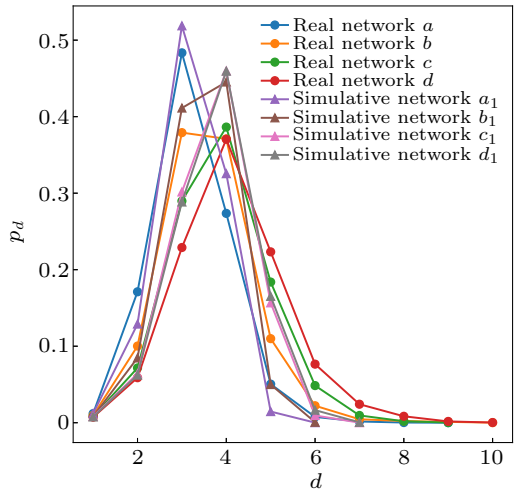


Fig. 4. Average shortest path length distribution.

3.2.5. Centrality

As shown in Figs. 5 and 6, the overall centrality values of both real and simulated networks are relatively low. This is because there are a small number of highly influential users on Weibo, whose posts, behavior, and communication ability have an important impact on the network as a whole. While nodes corresponding to these high-impact users have relatively high centrality values, their proportion in the overall nodes is relatively low, resulting in lower overall centrality values. The value ranges of the two centrality indicators of the simulated network are more concentrated, while those in the real network are more dispersed. This also reflects the complexity of social relationships. Weibo, as a complex social network platform, features a complex web of relationships between users. The value range of the centrality of the real network is relatively scattered, reflecting the differences between Weibo users and the diversity of the network. There are differences in the influence, popularity, and information dissemination ability of different users in the network, as well as the uncertainty of establishing contribute to the wide distribution of centrality values. On the other hand, the simulated network incorporates the specific attributes of Weibo network users, introduces a particular connection mechanism, and employs a specific network evolution model. Given its more controlled factors, the centrality values in the simulated network are more concentrated. Additionally, some nodes in the simulated network have higher centrality values compared to those in the real network. This discrepancy may be due to the fact that as more users join Weibo, early participants or users with specific advantages might accumulate more attention, interaction, and influence, resulting in higher centrality values for these nodes. Based on the evolution idea of the BA scale-free network, the 'rich-get-richer' effect is evident, which is consistent with both theory and observed phenomena. Although the centrality indices in the simulated network differ from those in the real network, the simulation captures the real network's characteristics relatively well, evident in the lower centrality values shared by both networks.

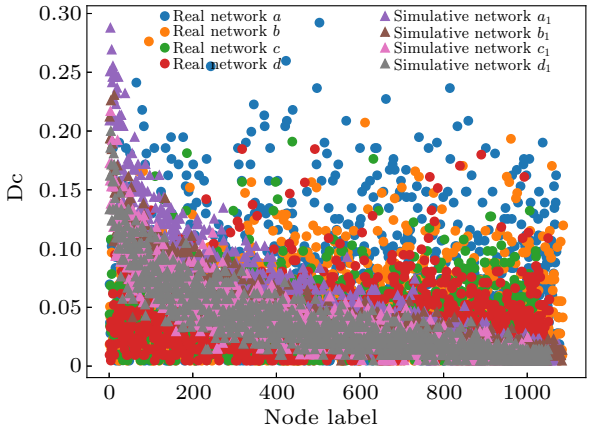


Fig. 5. Degree centrality.

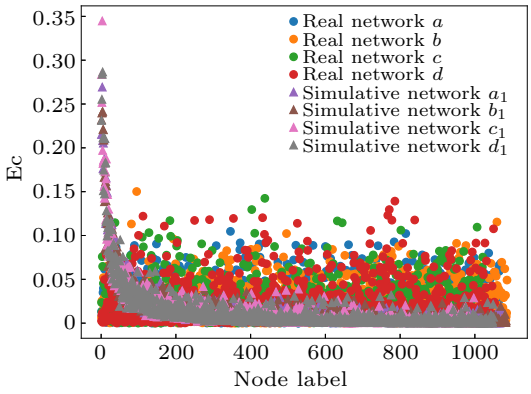


Fig. 6. Eigenvector centrality.

4. Conclusion

This paper proposes a local Weibo network evolution model based on the similarity of user attributes and behavioral characteristics. First, it abstractly defines four factors that affect the evolution of Weibo network. Building on this foundation, the BA scale-free network model was modified to incorporate these four factors and account for directionality, constructing a Weibo local network evolution model based on user attributes and behavioral similarity. By comparing the simulation results with empirical data, the validity of the model is confirmed. The results show that the local Weibo network not only has small-world characteristics, but also has some special properties. These include the in-degree distribution following a mixed distribution of exponential and power law, the degree–degree correlation demonstrating isomatching, and both the in-degree centrality value and the characteristic vector centrality value being relatively low. This research contributes to our understanding of user behaviour in the Weibo network, and provides a structural basis for exploring the impact of the Weibo network structure on information dissemination. However, it should be noted that this model has multiple parameters, reflecting a complex generation process. There are certain differences between the simulation network and the real network.

With some parts of the simulated network having higher centrality values than those of the real network, and the clustering coefficient being smaller in the simulated network than in the real network. Therefore, future research should aim to develop more concise and analyzable evolutionary network models.

Acknowledgement

Project supported by the National Natural Science Foundation of China (Grant Nos. 62266030 and 61863025).

References

[1] Watts D J and Strogatz S H 1998 *Nature* **393** 440
[2] Barabási A L and Albert R 1999 *Science* **286** 509
[3] Dorogovtsev S N, Mendes J F F and Samukhin A N 2000 *Phys. Rev. Lett.* **85** 4633
[4] Li X and Chen G 2023 *Physica A* **328** 274
[5] Bianconi G and Barabási A L 2001 *Europhys. Lett.* **54** 436
[6] Han D, Qian J and Ma Y 2011 *Europhys. Lett.* **94** 28006
[7] Zou Z Y, Liu P, Lei L and Gao J Z 2022 *Chin. Phys. B* **21** 028904
[8] Guo J L 2014 *Chin. Phys. B* **23** 070206
[9] Ran Y, Liu T, Jia T and Xu X K 2022 *Chin. Phys. B* **31** 068902
[10] Qian J H and Han D D 2009 *Physica A* **388** 4248
[11] Wasserman S and Faust K 1994 *Social network analysis: Methods and applications*
[12] Newman M E and Park J 2013 *Phys. Rev. E* **68** 036122
[13] Junju L and Jian W 2014 *IEEE 5th International Conference on Software Engineering and Service Science*, Beijing, China, 2014, pp. 983-986
[14] Guo J L, Guo Z H and Liu X J 2011 *Chin. Phys. B* **20** 118902
[15] Wu Y, Xiao J H, Wu Z Y, Yang J Z and Ma B J 2007 *Acta Phys. Sin.* **56** 2037 (in Chinese)
[16] Ahn Y Y, Han S, Kwak H, Moon S and Jeong H 2007 *In Proceedings of the 16th international conference on World Wide Web* (WWW '07). Association for Computing Machinery, New York, NY, USA, pp. 835–844
[17] Mislove A, Marcon M, Gummadi K P, Druschel P and Bhattacharjee B 2007 *Proceedings of the 7th ACM SIGCOMM conference on Internet measurement* (IMC '07). Association for Computing Machinery, New York, NY, USA, pp. 29–42
[18] Fu F, Liu L and Wang L 2008 *Physica A* **387** 675
[19] Traud A L, Mucha P J and Porter M A 2012 *Physica A* **391** 4165
[20] Xiong X, Cao W, Zhou X and Hu Y 2012 *Journal of Sichuan University* (Engineering Science Edition) **44** 140
[21] Kumar R, Novak J and Tomkins A 2006 *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining* pp. 611–617
[22] Xiong F, Liu Y, Si X and Ding F 2010 *Acta Phys. Sin.* **59** 6889 (in Chinese)
[23] Bu Z, Xia Z, Wang J and Zhang C 2013 *Physica A* **392** 2240
[24] Hu H B, Guo J L and Chen J 2012 *Chin. Phys. B* **21** 118902
[25] Java A, Song X, Finin T and Tseng B 2007 *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 workshop on Web mining and social network analysis* pp. 56–65
[26] Kwak H, Lee C, Park H and Moon S 2010 *Proceedings of the 19th international conference on World wide web* pp. 591–600
[27] Yan Q, Wu L and Zheng L 2013 *Physica A* **392** 1712
[28] Junju L and Jian W 2014 2014 *IEEE 5th International Conference on Software Engineering and Service Science* (IEEE), Beijing, China pp. 983–986
[29] Tang X, Miao Q, Quan Y, Tang J and Deng K 2015 *Knowledge-Based Systems* **89** 681
[30] Nian F, Shi Y and Dang Z 2020 *Int. J. Mod. Phys. B* **34** 2050323
[31] Nian F, Luo L and Liu X 2023 *IEEE Transactions on Computational Social Systems* **11** 1967
[32] Nian F, Ren J and Yu X 2023 *IEEE Transactions on Computational Social Systems* **11** 1277
[33] Yu X, Nian F, Yao Y and Luo L 2021 *IEEE Transactions on Computational Social Systems* **8** 1143
[34] Nian F and Diao H 2019 *IEEE Transactions on Network Science and Engineering* **7** 1394
[35] Zhang X, Han D D, Yang R and Zhang Z 2017 *PloS one* **12** e0183290
[36] Yuan W G and Liu Y 2015 *Physica A* **432** 167