

Modeling and Statistical Properties Research on Online Real-Time Information Transmission Network

Guangming Deng^{1,2}, Zhen Jia^{1,2}

¹College of Science, Guilin University of Technology, Guilin, China

²Guangxi Key Laboratory of Spatial Information and Geomatics, Guilin, China

Email: dgm@glut.edu.cn

Received 22 February 2014; revised 26 March 2014; accepted 3 April 2014

Copyright © 2014 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

In this paper, the model of the online real-time information transmission network, such as wechat, micro-blog, and QQ network, is proposed and built, based on the connection properties between users of the online real-time information transmission network, and combined with the local world evolving characteristics in complex network, then the statistical topological properties of the network is obtained by numerical simulation. Furthermore, we simulated the process of information transmission on the network, according to the actual characteristics of the online real-time information transmission. Statistics show that the degree distribution presents the characteristics of scale free network, presenting power law distribution, while the average path length, the average clustering coefficient and the average size of the network also has a power-law relationship, moreover, the model parameters has no effect on power-law exponent. The spread of information on the network represents obvious fluctuation scaling, reflecting the characteristics that information transmission fluctuates over time.

Keywords

Network, Information Transmission, Real-Time Information, Fluctuation Scaling

1. Introduction

With the rapid development of Internet application technology, a variety of social networking services (SNS) site is also expanding rapidly, such as Tencent, xiaonei net, happy net and so on. SNS has attracted tens of millions of Internet users through online chat, wechat, micro-blog and the sharing of the community platform.

How to cite this paper: Deng, G.M. and Jia Z. (2014) Modeling and Statistical Properties Research on Online Real-Time Information Transmission Network. *Open Journal of Applied Sciences*, 4, 234-241.
<http://dx.doi.org/10.4236/ojapps.2014.45023>

People form a so-called “acquaintance acquaintance” of large-scale online social networks, based on their personal social circles intertwined together, transfer and sharing of information through wechat, micro-blog, QQ chat and other channels. In recent years, with the development and wide application of Instant Messaging (IM) system, many scholars have done the research on the topology evolution mechanism of IM network. In 2000, Barabasi and Albert (BA) extended the BA network, and built a GBA (General BA) model, topological evolution driven by local events [1], the study found that the degree distribution of the network still obey the power-law distribution; In 2002, the American scholar Reginald Smith used database to carry on the empirical analysis of BA networks, and statistics showed that the degree distribution of network follows a power-law distribution; In 2005, Chinese scholars Yao Yuanyuan extended the BA network and applied it to IM networks, then proposed VGBA model. The model is validated by the data [2], and the study shows that the degree distribution of the network also obeys power-law distribution. Many studies have shown that IM network has characteristics of scale-free networks [3] [4]. In addition, Barabasi found the relationship between the average and standard deviation of network nodes’ flow [5], and physicists call this relationship as fluctuation scaling (FS) [6]. Many studies have found that the dissemination process on complex networks exist fluctuations scaling characteristics [7]-[10].

Online real-time information transmission network (ORITN), as a carrier of information dissemination, has its own structural characteristics. Take Tencent QQ network for instance, each QQ users usually add new QQ users into their circle of friends, according to their own interests and hobbies or real social relations, in their own limited social circles, so as to establish the connection relationship between users.

Meanwhile, the transmission of real-time information in the QQ network also has its own characteristics. For instance, whether a QQ user to forward a message or not is depended on the user’s interests or the role of trust between the different friends, so it is subjective, resulting in the transmission of real-time information and shows strong randomness. Moreover, the way of information’s transmission in QQ network differs with the general on-line network [11]-[16]. The main difference is that the former is subjective initiative and non-contacted, while the latter is mechanical.

In this paper, take QQ network for instance, we proposed and built the model of the online real-time information transmission network, namely ORITN, based on the connection properties between users of network, and combined with the evolving characteristics in social network [17]-[20] and local world network [21], then the statistical topological properties of the network is obtained by numerical simulation. Furthermore, we simulated the process of information transmission on the network, and put the average entropy of the real-time information, which is received by the network nodes, as a time series. Through statistical analysis, we got the fluctuation scaling of the real-time information transmission on the network.

2. Network Modeling of ORITN

2.1. Connection Properties between Network Users

Take QQ network for instance, each QQ user represents a unique node, the friends relationship between users represents the undirected edges, an undirected graph, $G = \langle V, E \rangle$ is constructed for analysis. Wherein V is the set of nodes, namely QQ users, E is the set of edges, representing the friends’ relationship between users. If user i and user j mutually add each other to become a contact, then there is an edge between nodes i and node j , which is denoted by e_{ij} . On the contrary, there is no edge connected between them. The adjacency matrix of $G = \langle V, E \rangle$ is denoted by $A(G) = (a_{ij})$. Wherein $a_{ij} = 1$ if there is an edge between nodes i and node j , on the contrary, $a_{ij} = 0$.

The connection between nodes in ORITN network, to a great extent, is a reflection of real interpersonal relationship. In reality, each person’s range of social activities is limited, showing the characteristics of local-world. Take QQ network for instance, QQ user i ’s local network is equivalent to its social circle, just as “birds of a feather flock together”. On one hand, among its limited social circle, user i tend to add the active user j , thus the QQ network has the characteristics of local-world network [5]. Within the scope of the local-world, the more active the user is, the more easily connects with the new node. Wherein, QQ user j ’s activity level is determined by k_j , namely node j ’s degree. On the other hand, in the QQ network, when a new user joins the network and establish a connection with the old user, in addition to considering the activity level of each old node within the scope of the local-world, but also consider its cohesion in the whole network. The greater the cohesion of the old node is, the greater the probability of a new node is connected to it. In the ORITN network model, we introduce node-weighted to each node j , and use it to reflect the size of node’s cohesion, wherein, the node-weighted is

denoted as ω_j . Here

$$\omega_j = \frac{C_j}{\sum C_l}$$

in which C_j means node j 's clustering coefficient in the whole network, $\sum C_l$ means the sum of all nodes' clustering coefficient in the whole network.

Based on the analysis above, when node i add new contacts, we make the following assumptions about ORITN network model:

First, randomly select a certain number of nodes from the whole ORITN network to form a local network, wherein the number of the nodes is denoted by M , and the local network is denoted by Ω , while Ω corresponds to user i 's social circle. Secondly, among node i 's local network Ω , node i connects to node j in accordance with the principles of the value of priority of $P(j)$. Here

$$P(j) = \frac{\omega_j k_j}{\sum_{l \in \Omega} (\omega_l k_l)}$$

in which k_j means node j 's degree in the local network Ω , ω_j means node j 's cohesion in the whole ORITN network, namely node j 's node-weighted.

2.2. Topological Evolution of the ORITN Network Model

Based on the model assumptions above, the algorithm of ORITN model's topology evolution is as follows:

1) Growth mechanism: Initially, the initial network has n_0 nodes and m_0 edges, add a new node and its incidental m edges each time.

2) Local priority connection mechanism: Randomly select M nodes from the whole network to form a local network, which is denoted by Ω . The newly added node connects to m nodes in the local network Ω , based on the preferential attachment probability formula. Here

$$\Pi(j) = \frac{M}{n_0 + t} \cdot \frac{\omega_j k_j}{\sum_{l \in \Omega} \omega_l k_l}$$

in which k_j means node j 's degree in the local network Ω , ω_j means node j 's cohesion in the whole ORITN network.

After a t step evolution, an ORITN network is produced, in which the total number of nodes is $N = n_0 + t$, and the total number of edges is $E = m_0 + mt$.

3. Statistical Features of ORITN's Topology Structure

In this section, we use Matlab software to simulate the above-mentioned ORITN evolution model and analyze the statistical properties of the network topology.

First, apply the above evolutionary algorithm to generate an ORITN network model and set various parameters for numerical simulation. Then investigate its variation law of the degree distribution, the average clustering coefficient C and the average path length L of QQ network.

In this paper, simulation parameters are set as follows: initial nodes number $n_0 = 10$, initial edges number $m_0 = 10$, number of network evolution $t = 5000$. We get the figure of degree distribution of QQ network through computer simulation, and the result is shown in **Figure 1**.

As is shown in **Figure 1(a)**, most of the dispersed points coincide though parameter M takes different values, representing that the impact of M 's value on the degree distribution is small. Namely, the degree distribution of QQ network is not affected by the size of the local network. As is shown in **Figure 1(b)**, when parameter m takes different values, scattered points showing a clear classification, and there is a obvious similarity relation between each category, representing that the degree distribution of the network showing power-law distribution. It also shows that m , the incidental edge of the newly added node, only affects the proportion coefficient of distribution, without affecting the power exponent. The illustration in **Figure 1(a)** is the result of the linear fitting. It can be obtained from **Figure 1** that $P(k) \propto k^{-\gamma}$, $\gamma \approx 3$. Therefore, the degree distribution of the network

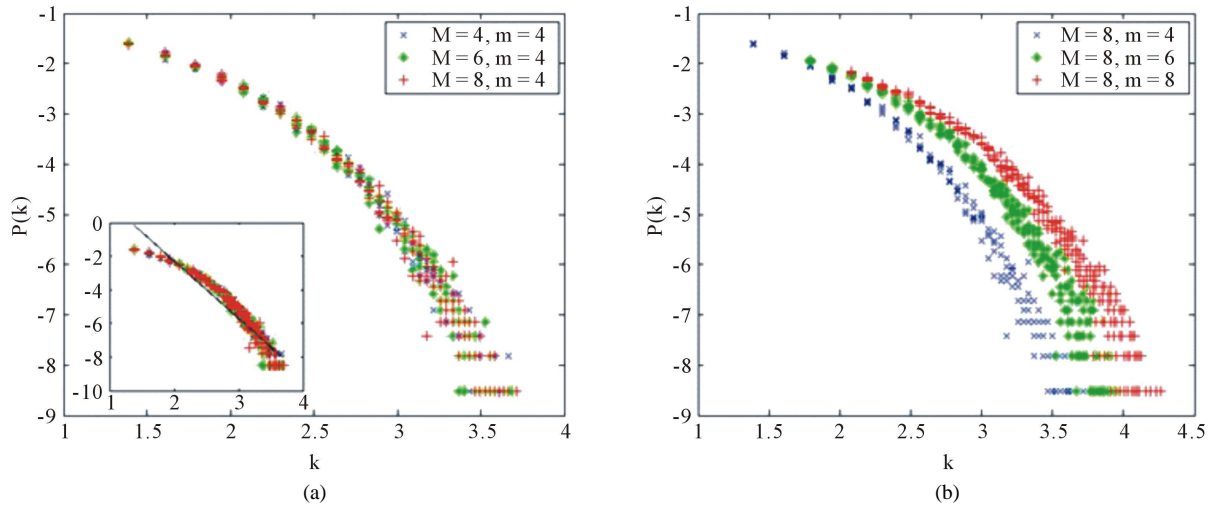


Figure 1. Probability versus degree. (a) Degree distribution when M takes different values; (b) Degree distribution when m takes different values.

follows a power-law distribution, $P(k) \propto k^{-3}$, indicating that the network has a power-law distribution characteristics.

Given $n_0 = 10$, $m_0 = 10$, $t = 5000$, we study the average clustering coefficient of QQ network with the change of network size through computer simulation, and the result is shown in **Figure 2**.

As is shown in **Figure 2(a)**, most of the dispersed points coincide though parameter M takes different values, representing that the impact of M 's value on the average clustering coefficient is small. Namely, the average clustering coefficient of QQ network is not affected by the size of the local network. As is shown in **Figure 2(b)**, when parameter m takes different values, scattered points showing a clear classification, and there is a obvious parallel relation between each category, representing the presence of power-law relationship between the average clustering coefficient C and the network size N . It also shows that m , the incidental edge of the newly added node, only affects the proportion coefficient of distribution, rather than the power exponent. The illustration in **Figure 2(a)** is the result of the linear fitting. It can be obtained from **Figure 2** that $\log(C) = -a \log(N) + b$, $a \approx 1$. The power-law relationship between the average clustering coefficient and the network size is $C \propto N^{-1}$.

Given $n_0 = 10$, $m_0 = 10$, $t = 5000$, we study the average path length of QQ network with the change of network size through computer simulation, and the result is shown in **Figure 3**.

As is shown in **Figure 3(a)**, all of the dispersed points completely superposed on the same straight line though parameter M take different values, representing that the average path length of QQ network is not affected by the size of the local network. As is shown in **Figure 3(b)**, when parameter m takes different values, scattered points showing a clear classification, and there is a obvious parallel relation between each category, representing the presence of power-law relationship between the average path length L and the network size N . It also shows that m , the incidental edge of the newly added node, only affects the proportion coefficient of distribution, rather than the power exponent. The illustration in **Figure 3(a)** is the result of the linear fitting. It can be obtained from **Figure 3** that $L = a_0 \cdot \log(N) + b_0$, $a_0 \approx 0.5$. Therefore, the relationship between the average path length and the network size is $\lim_{N \rightarrow \infty} \frac{L}{\log(N)} = a_0$, namely the average path length L and $\log(N)$ are of the same order,

indicating that the network has the characteristics of small world.

4. Characteristics of Real-Time Information Transmission on ORITN

In this section, the statistical properties of the real-time information's transmission on ORITN are further analyzed. Firstly, we propose the rule of information transmission according to its transmission characteristics, and simulate the process of real-time information's transmission on ORITN. Then statistics out the average times about the real-time information that nodes received through the transmission process, and take the average times as a time series. Afterwards, statistics out the standard deviation and average value of the time series, and find

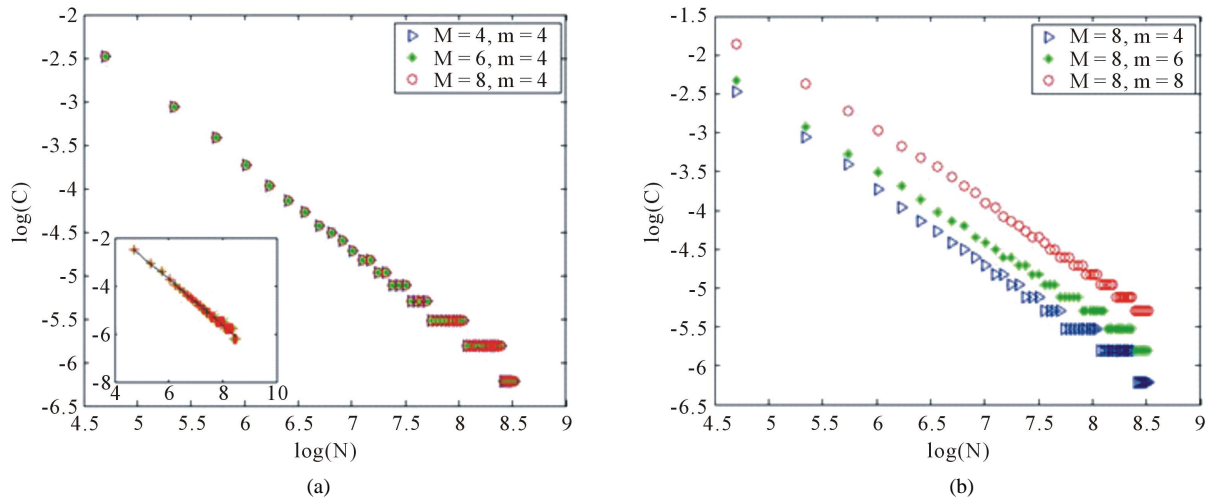


Figure 2. Average clustering coefficient versus network size. (a) Average clustering coefficient versus network size when M takes different values; (b) Average clustering coefficient versus network size when m takes different values.

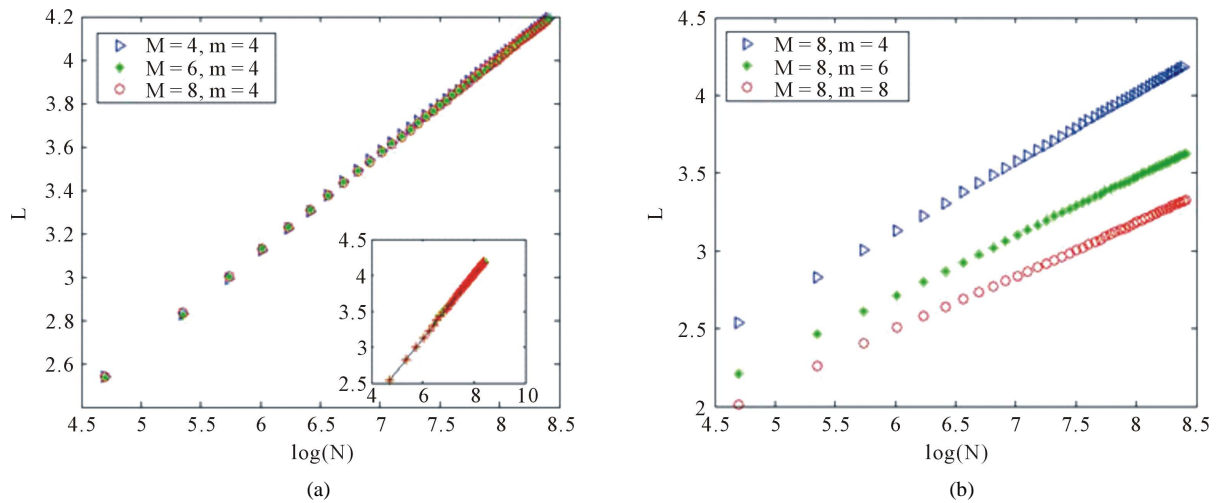


Figure 3. Average path length versus network size. (a) Average path length versus network size when M takes different values; (b) Average path length versus network size when m takes different values.

the function relation between them. Then further analyze whether the spread of information on the network represents the characteristics of fluctuation scaling.

4.1. Rule of Information Transmission

In reality, whether a user of ORITN transmits a real-time information is subjective, so it has a certain of randomness to transmit the real-time information on ORITN. Meanwhile, the information transmission is discontinuous, since the creation of real-time information is periodical and sudden. So in the process of information transmission, there may be a suspension, and then spreading it again. Based on the analysis above, we put forward the transmission rules as follow:

Step 1: Select a node i randomly from the network as a starting point for transmitting information, and create a new real-time information ψ . Node i transmit the information ψ to all its neighbor nodes, the total number of which is marked as K_i . Then each neighbor node j , who receive the information ψ , also transmit it to all its neighbor nodes instead of node i , at the probability of P . So the probability that node j doesn't transmit is $1 - P$, while the transmission probability is P . Proceed in accordance with such rules until the specific times of $[L]$. Wherein, L means the average path length of the network, while the top integral function $[L]$, is the function that

its value is the smallest integer greater than the independent variable or equal to it. For example, $[3.6] = 4$. After the suspension, each node in the network, marked as k , received the information ψ many times. The number of times is marked as a random variable $\xi_k(t=1)$, and the entire process is an one time step.

Step 2: Start the transmission at a new time step. Among all the nodes that has received the information ψ , select a node i_0 randomly as a starting point for information transmitting, then transmit the information according to the rule of step 1, until a entire time step is finished. After the T th time step, where in $T = 2, 3, \dots$, each node k received the information ψ many times. The total number of times is marked as the random variable $\xi_k(t=T)$.

4.2. Time Series Analysis and Fluctuation Scaling

It's obvious that the random variable $\xi_k(t)$, $t = 1, 2, 3, \dots, T$ is a time series, then we set $f(t) = \frac{1}{N} \sum_{k=1}^N \xi_k(t)$, in

which $f(t)$ is a time series and N is the scale of the network.

Then $f(t)$ represents the average times of the information that nodes in the network received, after the t th time step is accomplished, and the average value of $f(t)$ is $\langle f(t) \rangle = E(f(t))$. It's obvious that $\langle f(t) \rangle$ is a monotonic increasing function and $\lim_{t \rightarrow \infty} \langle f(t) \rangle = \infty$, and the standard deviation of $f(t)$ is

$$\sigma(t) = \left[E(f^2(t)) - \langle f(t) \rangle^2 \right]^{0.5}$$

Application of Matlab software, we conducted simulation experiments of real-time information transmission on the ORITN network of $N = 5010$. The total time step for each experiment is $T = 1000$. The average results of 50 times repeated simulation are shown in **Figure 4** and **Figure 5**.

As is shown in **Figure 4**, when the probability P takes different values, almost all of the scatted points are in the same line, indicating the presence of power-law relationship between the standard deviation $\sigma(t)$ and the average values $\langle f(t) \rangle$ of time series $f(t)$, and the value of the transmission probability P does not influence the exponent. The illustration in **Figure 4** is the result of the linear fitting, and it can be obtained from **Figure 4** that $\sigma(t) \propto \langle f(t) \rangle^\alpha$, $\alpha = 0.98 \approx 1$.

As is shown in **Figure 5** and its illustration, when the transmission probability P takes different values, the exponent values are very close to 1, wherein $0.94 \leq \alpha \leq 1$. It indicates that for any arbitrary time step t and the probability P , $\sigma(t) \propto \langle f(t) \rangle$ is constant established, in other words, information transmission on ORITN represents obvious fluctuation scaling.

5. Conclusion

In this paper, take QQ network for instance, by analyzing the relationship between network users and information

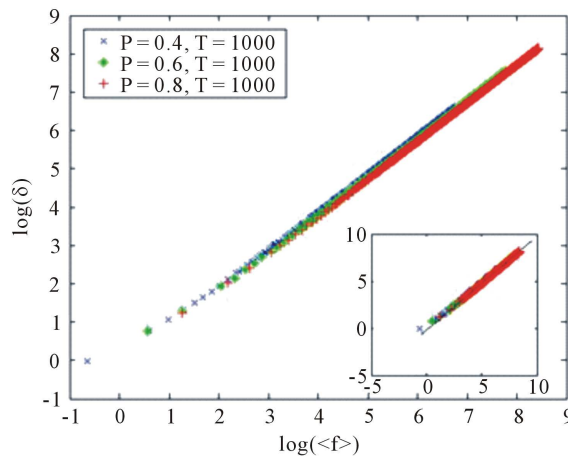


Figure 4. Time series' standard deviation versus average value.

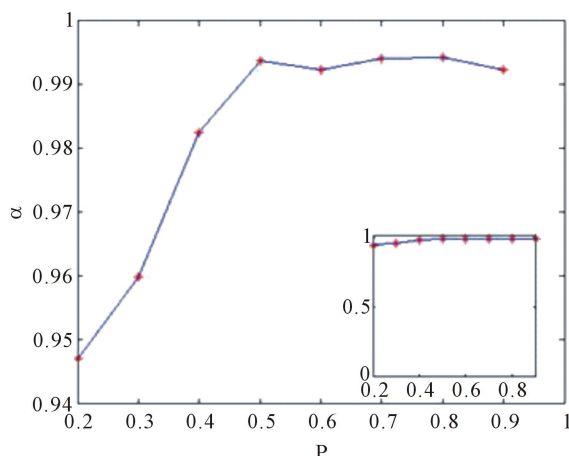


Figure 5. Exponent of time series versus transmission probability.

transmission characteristics, we proposed and built the model of the online real-time information transmission network, namely ORITN. Through the simulation to model algorithm, we found some important properties of ORITN from the statistical data. For instance, the degree distribution of the network follows power-law distribution, and $P(k) \propto k^{-3}$, wherein the exponent is 3. Both of the average clustering coefficient C and the average path length L have the power-law relationship with the network size N , wherein $C \propto N^{-1}$, $\lim_{N \rightarrow \infty} \frac{L}{\log(N)} = a_0$.

The power exponent has nothing to do with the parameters of the network. Meanwhile, by simulating real-time information transmission process, the statistical analysis of time series of $f(t)$ shows that the spread of information on the network represents obvious fluctuation scaling, reflecting the characteristics that information transmission fluctuates over time.

Acknowledgements

This work was jointly supported by the National Social Science Fund (No.13BTJ009), the National Natural Science Foundation (No.61164020), and the Guangxi Key Laboratory of Spatial Information and Geomatics (No.1207115-27).

References

- [1] Albert, R. and Barabasi, A.L. (2000) Topology of Evolving Networks: Local Events and Universality. *Physical Review*, **85**, 5234-5237. <http://dx.doi.org/10.1103/PhysRevLett.85.5234>
- [2] Yao, Y.Y. and Zhang, Z.H. (2006) Modeling the Instant Messaging Network as a Complex Network. *Microcomputer Information*, **22**, 76.
- [3] Albert, R., Jeong, H. and Barabasi, A.L.(2000) Error and Attack Tolerance in Complex Networks. *Nature*, **406**, 387-482.
- [4] Goh, K., Kahng, B. and Kim, D. (2002) Fluctuation-Driven Dynamics of the Internet Topology. *Physical Review Letters*, **88**, Article ID: 108701. <http://dx.doi.org/10.1103/PhysRevLett.88.108701>
- [5] Argollo de Menezes, M. and Barabasi, A.L. (2004) Fluctuations in Network Dynamics. *Physical Review Letters*, **92**, Article ID: 028701.
- [6] Eisler, Z. and Kertesz, J. (2006) Scaling Theory of Temporal Correlations and Size. *Physical Review E*, **73**, Article ID: 046109. <http://dx.doi.org/10.1103/PhysRevE.73.046109>
- [7] Eisler, Z. Bartos, I. and Kertész, J. (2008) Fluctuation Scaling in Complex Systems: Taylor's Law and Beyond. *Advances in Physics*, **57**, 89-142. <http://dx.doi.org/10.1080/00018730801893043>
- [8] Onnela, J.P., Saram, K.J., Hyvnen, J., Szabó, G., Lazer, D. and Kaski, K. (2005) Intensity and Coherence of Motifs in Weighted Complex Networks. *Physical Review E*, **71**, Article ID: 065103. <http://dx.doi.org/10.1103/PhysRevE.71.065103>

- [9] Duch, J. and Arenas, A. (2005) Community Detection in Complex Networks Using Extremal Optimization. *Physical Review*, **72**, Article ID: 027104.
- [10] Han, D.-D., Liu, J.-G. and Ma, Y.-G. (2008) Fluctuation of the Download Network. *Chinese Physics Letters*, **25**, 765-768. <http://dx.doi.org/10.1088/0256-307X/25/2/112>
- [11] Deng, Q.X., Jia, Z., Xie, M.S. and Chen, Y.F. (2013) Study of Directed Networks-Based Email Virus Propagation Model and Its Concussion Attractor. *Acta Physica Sinica*, **2**, Article ID: 020203.
- [12] Ebel, H., Mielsch, L.I. and Bornholdt, S. (2002) Scale-Free Topology of Email Networks. *Physical Review E*, **66**, Article ID: 035103.
- [13] Yan, G., Zhou, T.W, Wang, J., *et al.* (2005) Epidemic Spread in Weighted Scale-Free Networks. *Chinese Physics Letters*, **22**, 510-513. <http://dx.doi.org/10.1088/0256-307X/22/2/068>
- [14] Zanette, D.H. (2001) Critical Behavior of Propagation on Small-World Networks. *Physical Review E*, **64**, Article ID: 050901. <http://dx.doi.org/10.1103/PhysRevE.64.050901>
- [15] Zanette, D.H. (2002) Dynamics of Rumor Propagation on Small-World Networks. *Physical Review E*, **65**, Article ID: 041908. <http://dx.doi.org/10.1103/PhysRevE.65.041908>
- [16] Zhou, J., Liu, Z. and Li, B. (2007) Influence of Network Structure on Rumor Propagation. *Physics Letters A*, **368**, 458-463. <http://dx.doi.org/10.1016/j.physleta.2007.01.094>
- [17] Liu, Z., Lai, Y.C. and Ye, N. (2003) Propagation and Immunization of Infection on General Networks with Both Homogeneous and Heterogeneous Components. *Physical Review E*, **67**, Article ID: 031911. <http://dx.doi.org/10.1103/PhysRevE.67.031911>
- [18] Jin, E.M, Michelle, G. and Newman, M.E.J. (2001) Structure of Growing Social Networks. *Physics Reviews*, **64**, Article ID: 046132.
- [19] Feldstein, S. (1982) Impression Formation in Dyads: The Temporal Dimension. In: Davis, M., Ed., *Interactional Rhythms*, Human Sciences Press, New York.
- [20] Avrahami, D. and Scott E.H. (2006) Communication Characteristics of Instant Messaging: Effects and Predictions of Interpersonal Relationships. *Proceedings of the 20th Anniversary Conference on Computer Supported Cooperative Work*, Banff, Alberta, 4-8 November 2006, 505-514.
- [21] Li, X. and Chen, G.R. (2003) A Local-World Evolving Network Model. *Physica A*, **328**, 274-296. <http://dx.doi.org/10.1103/PhysRevE.67.031911>