

Research Article

Analysis of the Dynamic Influence of Social Network Nodes

Hong-Jian Yin, Hai Yu, Yu-Li Zhao, Zhi-Liang Zhu, and Wei Zhang

Software College, Northeastern University, No. 195, Chuangxin Road, Shenyang, Liaoning, China

Correspondence should be addressed to Hai Yu; yuhai@126.com

Received 21 October 2016; Accepted 5 January 2017; Published 31 January 2017

Academic Editor: Fabrizio Riguzzi

Copyright © 2017 Hong-Jian Yin et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In recent years, with the development of the social network theories, how to find or mining the most significant node in social network for understanding or controlling the information dissemination has become a hot topic and a series of effective algorithms have been presented. In this paper, a new scheme to measure the dynamic influence of the nodes in a social network is proposed, in which the sum of trust values of the propagation nodes is used. Simulations have been carried out and the results show that our scheme is stable and accurate.

1. Introduction

In the past decades, the revolutionary developments of communication tools have made significant changes to peoples social relationships. In the 1960s, Milgram's small-world experiment showed that the average distance between any two people on Earth is six, and this phenomenon is referred to six degrees of separation [1, 2]. In 2011, the results of the analysis of the friend networks of 750 million active users in Facebook showed that the average distance between Facebook network nodes was only 4.74 degrees [3, 4]. In social network analysis, it is quite significant to find out or mining which node has the largest impact. Therefore, a lot of measurements have been proposed to calculate the importance of a node from different perspectives, including degree centrality, betweenness centrality, closeness centrality, k-shell centrality, eigenvector centrality, and the PageRank algorithm.

Degree centrality was proposed by Professor Linton, which reflects local properties of the network, and the main consideration is the node itself and the neighbors properties. Although the calculation of degree centrality is simple, it has some deficiencies [5–7]. Betweenness centrality, closeness centrality, and eigenvector centrality reflect the global property [8] of networks. Among them, betweenness centrality [9, 10] mainly considered the shortest path through the node. Closeness centrality [11] measures the difficulty to reach the other node. Eigenvector centrality [12, 13] mainly considered the status and prestige in the networks using the composition

of the reputation of other nodes to reflect the influence of the node for the entire network. K-shell centrality reflects the nodes location within networks to measure node communication capacities [14, 15]. In addition, the PageRank algorithm [16] is also used to measure the impact of network nodes.

Currently, the most measurements are based on statistical properties with the topology of the networks and do not take the impact of changes of mutual trust among the nodes during information dissemination into account. In this paper, a new scheme to measure the dynamic influence of the nodes in a social network is proposed. In this new scheme, the modification of node trust value during information propagation plays a significant role. Furthermore, the cumulative change of nodes trust value is also considered in the new scheme.

2. The Measure of Dynamic Influence

2.1. The Model of Information Dissemination. SI, SIS, SIR, and so forth [17, 18] were originally used to research the spread of disease [19–21]. In these networks, people and their relations are considered as node and edge, which can be represented by $G = (V, E)$, where V is a set of nodes and E is the set of connected edges. All nodes can be divided into three categories: class S (susceptible) refers to those who do not get sick, but, because of the lack of immunity, they are susceptible to infection after contact with a sick sense; class I (infective) refers to those who had infectious disease and it can spread

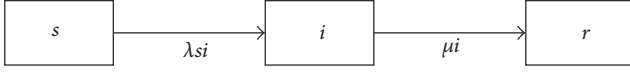


FIGURE 1: The process is shown for nodes from the susceptible to removal state.

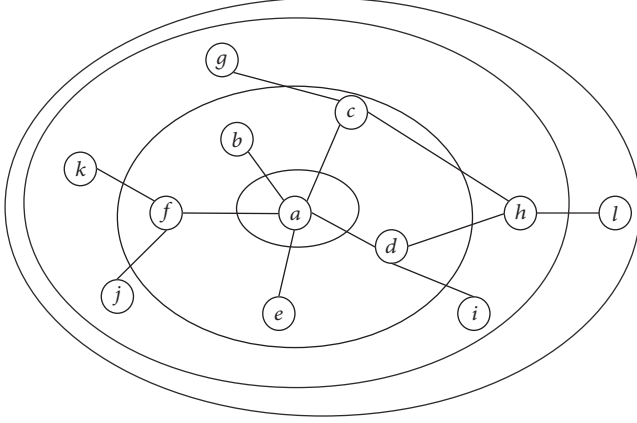


FIGURE 2: Layered network diagram.

to class S members; and class R (removal) refers to isolation or the person who has the illness and immunity.

In addition, suppose the number of nodes is constant λ , and each time the number of nodes to be removed is a constant proportion of the total number of μ . The average propagation period is $1/\mu$. The dissemination number is $\sigma = \lambda/\mu$. Figure 1 shows the procedure from susceptible state to removal state.

The SIR model is defined by following equation:

$$\begin{aligned} \frac{di}{dt} &= \lambda si - \mu i, \\ \frac{ds}{dt} &= -\lambda si, \\ \frac{dr}{dt} &= \mu i. \end{aligned} \quad (1)$$

If the SIR model is used to illustrate the information dissemination in network $G = (V, E)$, s is considered as a node that can receive information, i represents a node that has received information and has the ability to disseminate information, and r represents a node which has received information but does not have the ability to disseminate it.

2.2. The Measure of Node Dynamic Influence. In this section, to further investigate the relationship between the sender node and the receiver node in a complex network, the nodes are stratified according to their distance from the node of information source. The layered network is shown in Figure 2.

The dynamic influence index of a social networks is represented by I . Several rounds of information sources on the network node transmitting information are represented by k . In addition, the trust value t is the cumulative effect of

the information spreading, where t represents the node that is a push message acceptance of the push message from a node, namely, trust, $0 < t < 1$.

In Figure 2, node a is considered as the information source node, and node a disseminates information to a set of neighbors $\{b, c, d, e, f\}$ with a certain probability. Node a and node c are connected by edge P_{ac} ; t_{ac} is the trust value on edge P_{ac} . When variable k changes, the value of t is also changed. This feature can be considered as a dynamic influence.

Use V_c to represent the number of once valid pushing and V'_c to represent the number of once invalid pushing; $t_{k, P_{pq}}$ representing the trust value on edge P_{pq} of any two nodes p and q in the networks at k -rounds of dissemination, then the t value is

$$t_{k, P_{pq}} = \frac{V_{C_{k, P_{pq}}}}{V_{C_{k, P_{pq}}} + V_{C'_{k, P_{pq}}}} \quad (2)$$

The values of V_c at k -rounds and $(k + 1)$ -rounds of dissemination are

$$V_{C_{k+1, P_{pq}}} = \begin{cases} V_{C_{k, P_{pq}}} + 1, & p * d \geq 0.5 \\ V_{C_{k, P_{pq}}}, & p * d < 0.5. \end{cases} \quad (3)$$

Meanwhile, V'_c value is calculated by

$$V'_{C_{k+1, P_{pq}}} = \begin{cases} V'_{C_{k, P_{pq}}} + 1, & p * d < 0.5 \\ V'_{C_{k, P_{pq}}}, & p * d \geq 0.5. \end{cases} \quad (4)$$

If $p * d < 0.5$, the edge of a message is an invalid recommendation, while $p * d \geq 0.5$ means that there is a valid recommendation in (3) and (4), where p is the probability of the current node to propagate the message to its neighbor class S nodes and d is the number of neighbor node of the current node. It is important to note that d represents outdegree in the directed network, while, in the undirected network, it denotes the degree of the node.

Probability p is determined by the level of the node, the information lost during dissemination, and effects of cumulative history dissemination to the current dissemination. In Figure 2, according to the information attenuation principle, we can see that probability p is inversely proportional to variable l (when $l \geq 1$, the default value l for each node is 1). L indicates the distance of the current node to source node. Probability p is also inversely proportional to c (when $c \geq 0$, the default value c for each node is 0), which indicates the number of times that network information dissemination process. Probability p is proportional to trust t , which derived from edge connected the current node and its parent node.

The definition of probability p is as shown in

$$p = \begin{cases} \frac{1}{c * l^2}, & c \geq 1, l \geq 1, 0 \leq t < \alpha \\ \frac{1}{c * l^2 * (1 + \alpha - t)}, & c \geq 1, l \geq 1, 0.5 \leq \alpha < t < 1. \end{cases} \quad (5)$$

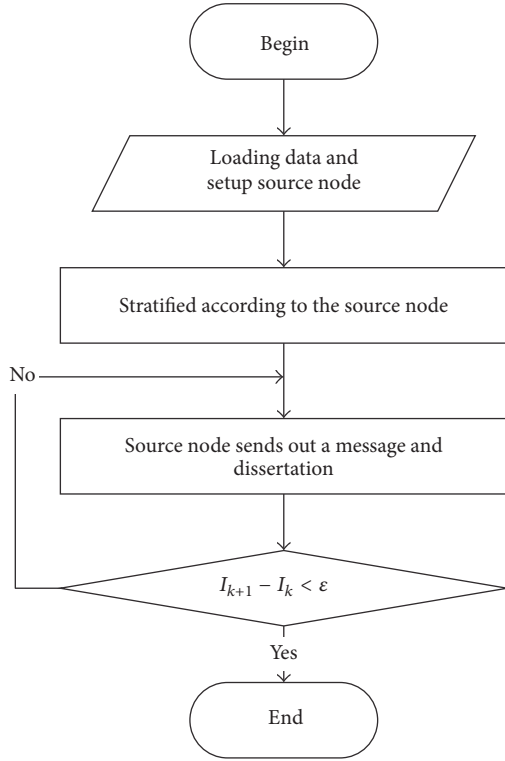


FIGURE 3: The flow chart of dynamic influence of node.

It is apparent that probability p decreases when layer decreased according to (3). Here, a new variable α is introduced to balance and limit the value range of p value.

After k -rounds of information dissemination, the number of edges is P_n , indicating the edge counts through the information route. In this case, the influence of source node a is defined as the following equation:

$$I_k = \sum_{(p,q)=0}^{P_n} t_{k,P_{p,q}}, \quad (6)$$

where $(p, q) = 0$ denotes the edge count; I_k is the influence of the source node a after k -rounds of the information dissemination.

2.3. The Detail of the Algorithm. The new algorithm aims to explore the relationship between the influence of the social network nodes and accumulation effects of information transmission. In this section, the stabilization of I_k is used to eventually measure the influence of the node. A flow chart of the detailed algorithm is shown in Figure 3.

The spreading node selection algorithm consists of two parts: (1) when one node pushing a message, most of its neighbors have t value to trust it. (2) Choose one neighbor node to receive the information.

The trade-offs of t value determine the value of probability p : a larger t indicates a larger p . In the algorithm, the algorithm selects the node who has the largest t value, which is $t = \max\{t_1, t_2, \dots, t_n\}$.

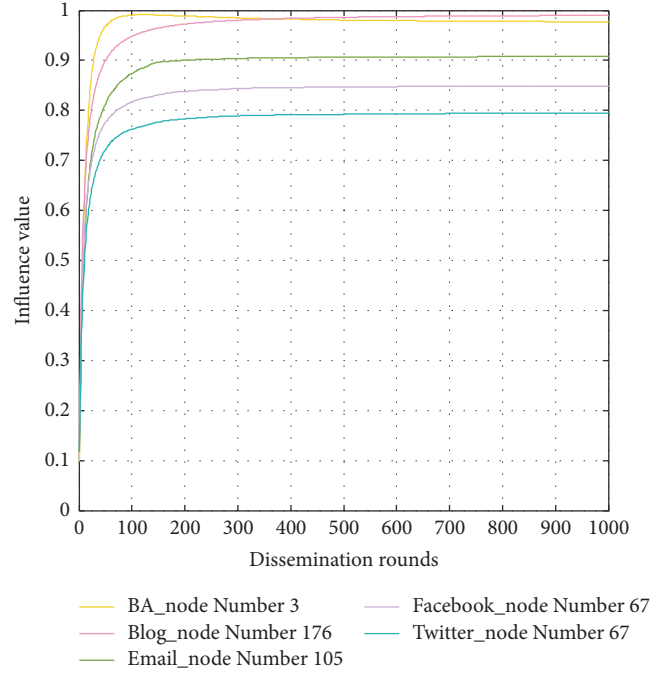


FIGURE 4: The changes of the dynamic influence of highest ranked node.

The push target node selection strategy is when a node has the ability to push messages selection in r neighbor nodes, push message in d neighbors, and select the maximum value of neighbor node t on the path of a push.

3. Simulation Results

The initial values of the coefficients in our simulation are set to $p = 0.5$, $l = 1$, $c = 0$, $\alpha = 0.5$, $t = 0.5$. All the data are from the public database [22, 23] shown in Table 1.

The dataset contains directed network, undirected networks, theoretical network, the real social networks, and the like. BA scale-free network [24] is a theoretical network that was proposed by Barabási and Albert to produce power-law distribution mechanism. It needs to specify that because the original online social network dataset contains isolated node, in this experiment, remove the original raw data collection network in the isolated node. In this experiment, the isolated nodes are removed in the original raw data collection network, and the maximum connectivity subgraph [25] is used in our simulations.

In the simulations, the data of $(|E|/|N|)$ are 7.2594, 33.9749, 4.8111, 32.3878, and 9.989, respectively. The simulation results are listed in Table 2. In Table 2, D represents an average degree [26] of the network, L is the average path length [27] of the network, ϕ is the density of the network [28], ρ is the density [29] of the network, W is the weighted average of the network, and C is the average clustering coefficient [30, 31] of the network.

In Figure 4, we selected the highest ranking node from the five network datasets. The ranking is according to the dynamic influence of 1000-round dissemination. In Figure 5, similarly, we selected the lowest ranking node from the five

TABLE 1: Dataset source and dataset size.

Datasets	Node count	Edge count	Description
Facebook ¹	347	2519	Slavo Zitnik in social networking sites Facebook friends relationship
Twitter ²	239	8120	1,367,531 nodes in Twitter social networks friends relationship
Email ³	1133	5451	Mr. Tsang communication contact record
Blog ⁴	10312	333983	The relationship between bloggers in Weibo
BA	5000	49945	5000 nodes in a BA scale-free network

¹ Facebook: <http://snap.stanford.edu/data/egonets-Facebook.html>.

² Twitter: <http://snap.stanford.edu/data/egonets-Twitter.html>.

³ Email: <http://www.cs.bris.ac.uk/steve/networks/peacockpaper/>.

⁴ Blog: <http://www.blogcatalog.com/>.

TABLE 2: The results of the statistical parameters of five network datasets.

Datasets	D	L	ϕ	ρ	W	C
Facebook	15.129	3.752	11	0.046	15.129	0.508
Twitter	33.975	2.096	7	0.143	33.975	0.365
Email	11.559	1.992	3	0.01	28.79	0.526
Blog	64.776	2.382	5	0.006	0.238	0.463
BA	19.978	2.922	4	0.004	19.978	0.017

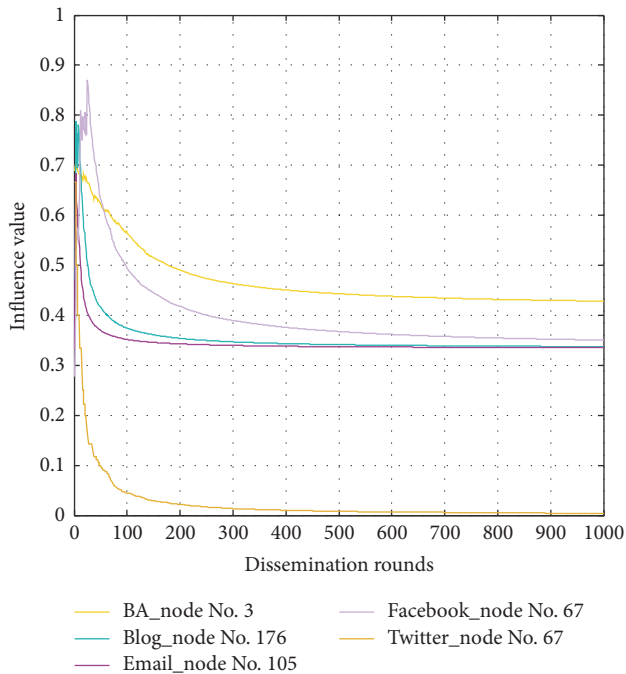


FIGURE 5: The changes of the dynamic influence of the lowest ranking.

network datasets; the ranking is according to the dynamic influence of 1000-round dissemination. The horizontal axis is the number of dissemination rounds; the vertical axis is the trust value that represents the dynamic influence, where the actual results of trust were normalized so that the results of different networks can be compared in the same coordinate system.

In the Facebook network, number 67 node has the highest influence, which is rising rapidly within 100 rounds of the dissemination phase. The influence of node number 67 increases slowly within 100 rounds to 200 rounds of the dissemination phase. Finally, stable dissemination occurs at approximately 400 rounds, and the trust value is 0.7911. Node number 157 has a very low influence and is slightly jittery within 50 rounds of the dissemination phase. The influence of node number 157 rapidly declines within 50 rounds to 200 rounds of the dissemination phase. Finally, stable dissemination occurs at approximately 784, with trust value 0.355.

According to the simulation results, we can see that although different networks have different statistical properties, they have a same pattern, and the node with the highest influence increases quickly when the amount of dissemination increased and became eventually stable at approximately 300 rounds. For the nodes with low influence, they decreased faster and became eventually stable at approximately 500 rounds.

In addition, the proposed algorithm (DI) is compared to degree centrality (DC), betweenness centrality (BC), closeness centrality (CC), eigenvector centrality (EC), and PageRank (PR) to verify the accuracy and validity of the algorithm.

In Table 3, the top four nodes are listed according to the different measure algorithms. The results are roughly the same. Finally, 10% of the nodes herein formerly had importance under each dataset used for analysis and comparison, as shown in Table 4.

In the Facebook network, I_{set} represents a collection of 10 elements under the five classical algorithms jointly that determined the top 10% nodes; I_{all} represents together with DC algorithm the top 10% to 10 nodes, namely, the intersection hits: $I_{hit} = 100\%$. Similarly, U_{set} represents the

TABLE 3: The top four nodes under several algorithms.

Datasets	Rank	DC	CC	BC	EC	PR	DI
Facebook	1	56	277	277	56	25	67
	2	67	25	175	67	56	56
	3	271	322	19	271	67	271
	4	322	67	23	26	322	26
Twitter	1	67	67	67	69	8	67
	2	47	47	144	67	2	175
	3	175	175	69	84	121105	47
	4	189	189	45	8	69	203
Email	1	105	333	333	105	105	105
	2	333	23	105	16	23	333
	3	16	105	23	196	333	42
	4	23	42	578	42	41	23
Blog	1	176	10271	4839	176	10298	176
	2	1226	10259	4375	233	10310	233
	3	446	8859	8859	283	9997	446
	4	645	10269	7806	371	10298	645
BA	1	3	3	3	3	3	3
	2	2	2	2	2	2	2
	3	4	4	4	4	4	4
	4	20	8	20	8	20	8

TABLE 4: Ranking hit rate in a variety of algorithms at the intersection and union.

Datasets	I_{set}	I_{all}	I_{hit}	U_{set}	U_{all}	U_{hit}
Facebook	10	10	100%	55	55	100%
Twitter	6	6	100%	62	65	95.38%
Email	38	38	100%	182	182	100%
Blog	157	135	85.99%	2320	2716	85.42%
BA	395	378	95.70%	610	645	94.57%

union top 10% nodes under the five classical algorithms and the number of elements in the set is 55; U_{all} represents the union together with DC algorithm in the top 10%, which is set as hits, $U_{hit} = 100\%$. Simulation results show that the proposed algorithm has good accuracy and effectiveness.

4. Conclusion

In this paper, a new judgment scheme on the dynamic influence of the social network nodes is proposed. Considering the effect of changes in the information dissemination process of trust values, a new measurement of node dynamic influence is proposed. It is an improvement of the traditional algorithms. Finally, we analyze the influence of nodes according to topology of the network or statistical properties and further compare it with several classical algorithms to verify the validity and accuracy of the algorithm.

Competing Interests

The authors declare that they have no competing interests.

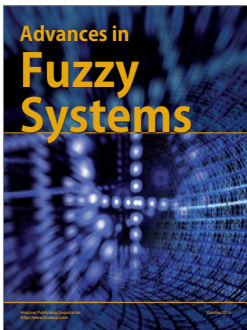
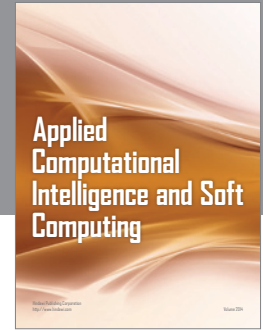
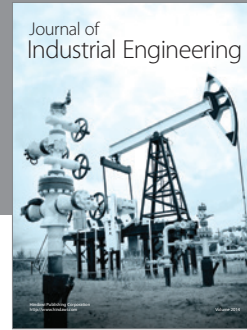
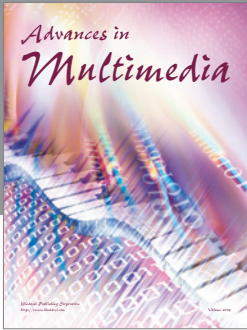
Acknowledgments

This research was supported by the National Natural Science Foundation of China (Grant nos. 61374178 and 61402092), the Online Education Research Fund of MOE Research Center for Online Education, China (Qtone Education, Grant no. 2016ZD306), and the Ph.D. Start-up Foundation of Liaoning Province, China (Grant no. 201501141).

References

- [1] S. Milgram, "The small world problem," *Psychology Today*, vol. 2, no. 1, pp. 60–67, 1967.
- [2] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, no. 6684, pp. 440–442, 1998.
- [3] L. Backstrom, P. Boldi, M. Rosa, J. Ugander, and S. Vigna, "Four degrees of separation," in *Proceedings of the 4th Annual ACM Web Science Conference*, pp. 33–42, 2012.
- [4] N. A. Christakis and J. H. Fowler, *Connected: The Surprising Power of Our Social Networks and How They Shape Our Lives—How Your Friends' Friends' Friends Affect Everything You Feel*,

- Think, and Do*, Little, Brown and Company, New York, NY, USA, 2011.
- [5] R. Albert, H. Jeong, and A.-L. Barabási, “Error and attack tolerance of complex networks,” *Nature*, vol. 406, no. 6794, pp. 378–382, 2000.
 - [6] R. Pastor-Satorras and A. Vespignani, “Epidemic spreading in scale-free networks,” *Physical Review Letters*, vol. 86, no. 14, pp. 3200–3203, 2001.
 - [7] R. Cohen, K. Erez, D. Ben-Avraham, and S. Havlin, “Breakdown of the internet under intentional attack,” *Physical Review Letters*, vol. 86, no. 16, pp. 3682–3685, 2001.
 - [8] L. Jian-Guo, R. Zhuo-Ming, G. Qiang, and W. Bing-Hong, *Node Importance Ranking of Complex Networks*, 2013.
 - [9] L. C. Freeman, “A Set of measures of centrality based on betweenness,” *Sociometry*, vol. 40, no. 1, pp. 35–41, 1977.
 - [10] N. E. Friedkin, “Theoretical foundations for centrality measures,” *American Journal of Sociology*, vol. 96, no. 6, pp. 1478–1504, 1991.
 - [11] G. Sabidussi, “The centrality index of a graph,” *Psychometrika*, vol. 31, no. 4, pp. 581–603, 1966.
 - [12] K. Stephenson and M. Zelen, “Rethinking centrality: methods and examples,” *Social Networks*, vol. 11, no. 1, pp. 1–37, 1989.
 - [13] S. P. Borgatti, “Centrality and network flow,” *Social Networks*, vol. 27, no. 1, pp. 55–71, 2005.
 - [14] L.-L. Ma, C. Ma, and H.-F. Zhang, “Identifying influential spreaders in complex networks based on gravity formula,” *Physica A*, vol. 451, pp. 205–212, 2016.
 - [15] S. Carmi, S. Havlin, S. Kirkpatrick, Y. Shavitt, and E. Shir, “A model of Internet topology using k-shell decomposition,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 27, pp. 11150–11154, 2007.
 - [16] L. Page, S. Brin, R. Motwani, and T. Winograd, “The pagerank citation ranking: bringing order to the web,” Stanford Digital Library Technologies Project, 1999.
 - [17] L. J. S. Allen, “Some discrete-time *SI*, *SIR*, and *SIS* epidemic models,” *Mathematical Biosciences*, vol. 124, no. 1, pp. 83–105, 1994.
 - [18] B. Shulgin, L. Stone, and Z. Agur, “Pulse vaccination strategy in the *SIR* epidemic model,” *Bulletin of Mathematical Biology*, vol. 60, no. 6, pp. 1123–1148, 1998.
 - [19] H. W. Hethcote, “The mathematics of infectious diseases,” *SIAM Review*, vol. 42, no. 4, pp. 599–653, 2000.
 - [20] M. J. Keeling and P. Rohani, *Modeling Infectious Diseases in Humans and Animals*, Princeton University Press, Princeton, NJ, USA, 2008.
 - [21] J. Heesterbeek, *Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation*, vol. 5, John Wiley & Sons, New York, NY, USA, 2000.
 - [22] N. Blagus, L. Šubelj, and M. Bajec, “Self-similar scaling of density in complex real-world networks,” *Physica A: Statistical Mechanics and its Applications*, vol. 391, no. 8, pp. 2794–2802, 2012.
 - [23] L. Tang and H. Liu, “Relational learning via latent social dimensions,” in *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD ’09)*, pp. 817–826, ACM, Paris, France, July 2009.
 - [24] A.-L. Barabási and R. Albert, “Emergence of scaling in random networks,” *Science*, vol. 286, no. 5439, pp. 509–512, 1999.
 - [25] P. Erdős and T. Gallai, “On maximal paths and circuits of graphs,” *Acta Mathematica Academiae Scientiarum Hungaricae*, vol. 10, no. 3–4, pp. 337–356, 1959.
 - [26] M. O. Jackson, *Social and economic networks*, Princeton University Press, Princeton, NJ, USA, 2008.
 - [27] A. Fronczak, P. Fronczak, and J. A. Hołyst, “Average path length in random networks,” *Physical Review E*, vol. 70, no. 5, Article ID 056110, 2004.
 - [28] A. Barreiras, “Diameter constrained network,” in *Proceedings of the World Congress on Engineering*, vol. 2, 2009.
 - [29] L. Kowalik, “Approximation scheme for lowest outdegree orientation and graph density measures,” in *Proceedings of the 17th International Conference on Algorithms and Computation (ISAAC ’06)*, pp. 557–566, Kolkata, India, December 2006.
 - [30] J. Saramäki, M. Kivelä, J.-P. Onnela, K. Kaski, and J. Kertész, “Generalizations of the clustering coefficient to weighted complex networks,” *Physical Review E*, vol. 75, no. 2, Article ID 027105, 2007.
 - [31] T. Schank and D. Wagner, “Approximating clustering coefficient and transitivity,” *Journal of Graph Algorithms and Applications*, vol. 9, no. 2, pp. 265–275, 2005.



Hindawi

Submit your manuscripts at
<https://www.hindawi.com>

