

Research Article

Identifying Super-Spreader Nodes in Complex Networks

Yu-Hsiang Fu,¹ Chung-Yuan Huang,² and Chuen-Tsai Sun¹

¹*Department of Computer Science, National Chiao Tung University, 1001 Ta Hsueh Road, Hsinchu 300, Taiwan*

²*Department of Computer Science and Information Engineering, School of Electrical and Computer Engineering, College of Engineering, Chang Gung University, 259 Wen Hwa 1st Road, Taoyuan 333, Taiwan*

Correspondence should be addressed to Chung-Yuan Huang; gscott@mail.cgu.edu.tw

Received 26 May 2014; Accepted 25 September 2014

Academic Editor: He Huang

Copyright © 2015 Yu-Hsiang Fu 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.

Identifying the most influential individuals spreading information or infectious diseases can assist or hinder information dissemination, product exposure, and contagious disease detection. Hub nodes, high betweenness nodes, high closeness nodes, and high k -shell nodes have been identified as good initial spreaders, but efforts to use node diversity within network structures to measure spreading ability are few. Here we describe a two-step framework that combines global diversity and local features to identify the most influential network nodes. Results from susceptible-infected-recovered epidemic simulations indicate that our proposed method performs well and stably in single initial spreader scenarios associated with various complex network datasets.

1. Introduction

Network-spreading studies range from information diffusion via online social media sites, to viral marketing, to epidemic disease identification and control, among many others [1–9]. Key spreader identification strategies are being established and tested to accelerate information dissemination, increase product exposure, detect contagious disease outbreaks, and execute early intervention strategies [10]. Topological structure is a core concept in this identification process [1, 2, 11–15].

Centrality measures for identifying influential social network nodes are broadly categorized as local or global [3, 7, 16]. Degree centrality (the number of nodes that a focal node is connected to) measures node involvement in a network. However, most network node researchers fail to consider global topological structures. Betweenness (which assesses the degree to which a node lies on the shortest path between two other nodes) and closeness (the inverse sum of the shortest distances from a focal node to all other nodes) are the two most widely used measures for overcoming these limitations. Influence is tied to advantageous network positions,

including high degree, high closeness, and high betweenness. In simple network structures, these advantages tend to vary; in complex networks, significant disjunctures can emerge among position characteristics, so that a spreader's location may be simultaneously advantageous and disadvantageous.

Results from k -shell decomposition analyses indicate that network nodes in core layers are capable of spreading throughout much broader areas compared to those in peripheral layers [1, 2]. Although spreading capability differs among nodes, those with similar k -shell values are perceived as having equal importance. To rank spreaders, a method called mixed degree decomposition adds otherwise ignored degree nodes to the decomposition process [3, 6, 8, 17]. Still, researchers have tended to overlook the importance of network topology and node diversity, despite their positive correlations with events such as community economic development [18].

We used the concept of entropy to develop a robust and reliable method for measuring the spreading capability of nodes and identifying super-spreader nodes in complex networks. It can be used to analyze numbers of global network topological layers and local neighborhood nodes

affected by specific individual nodes. Our assumption is that k -shell decomposition [1, 2] can be used for global analysis, with high global diversity/high local centrality nodes capable of penetrating multiple global layers and influencing large numbers of neighbors in local layers of complex networks.

To measure node influence, we propose a two-step framework for acquiring global and local node information within complex networks. Global node information is initially obtained using algorithms (e.g., a community detection algorithm for complex networks [5, 19, 20] or a k -shell decomposition algorithm for core/periphery network layers), after which entropy is used to evaluate network node global diversity. Next, local node information is acquired using various types of local centrality. Last, global diversity and local features are combined to determine node influence. In our experiments, spreading ability equaled the total number of recovered nodes over time. We used a susceptible-infective-recovered (SIR) epidemic simulation with various social network datasets [21–25] to compare the spreading capabilities of our proposed measure and social network local/global centralities [2, 26, 27].

2. Background

To represent a complex network, let an undirected graph $G = (V, E)$, where V is the network node set and E the edge set. $n = |V|$ indicates the number of network nodes and $m = |E|$ the number of edges. Network structure is represented as an adjacency matrix $A = \{a_{ij}\}$ and $a_{ij} \in R^n$, where $a_{ij} = 1$ if a link exists between nodes i and j , otherwise $a_{ij} = 0$.

Degree (or local) centrality is a simple yet effective method for measuring node influence in a complex network. Let $C_d(i)$ denote node i degree centrality. Higher values indicate larger numbers of connections between a node and its neighbors. $NB_h(i)$ denotes the set of node i neighbors at a h -hop distance. Node i degree centrality is therefore defined as

$$C_d(i) = |NB_h(i)| = \sum_{j=1}^n a_{ij}, \quad (1)$$

where $|NB_h(i)|$ is the number of node i neighbors at a h -hop distance; in most cases, $h = 1$ [7].

Betweenness centrality or dependency measures the proportion of shortest paths going through a node in a complex network. $C_b(i)$ denotes node i betweenness centrality. Higher values indicate that a complex network node is located along an important communication path. Accordingly, node i betweenness centrality is defined as

$$C_b(i) = \sum_{s \neq t \neq v \in V} \frac{Q_{st}(i)}{Q_{st}}, \quad (2)$$

where $Q_{st}(i)$ is the number of shortest paths from node s to node t through node i and Q_{st} is the total number of shortest paths from node s to node t [3, 7, 16].

Closeness (or global) centrality measures the average length of the shortest paths from one node to other nodes.

Let $C_l(i)$ denote node i closeness centrality. Higher values indicate node location in the center of a complex network, with a shorter average distance from that node to other nodes. Node i closeness centrality is thus defined as

$$C_l(i) = \frac{1}{l_i}, \quad l_i = \frac{1}{n} \cdot \sum_{j=1}^n d_{ij}, \quad (3)$$

where l_i is the average length of the shortest paths from node i to the other nodes and d_{ij} is the distance from node i to node j [16].

k -shell decomposition [1, 2] iteratively assigns k -shell layer values to all nodes in a complex network. During the first step, let $k = 1$ and remove all nodes where $C_d(n) = k = 1$. Following removal, some remaining network node degrees may be $k = 1$. Nodes are continuously pruned until there are no $k = 1$ nodes. All removed nodes are assigned a k -shell value of 1. The next step is similar: let $k = 2$, prune nodes, and assign a k -shell value of 2 to all removed nodes. Repeat the procedure until all network nodes are removed and assigned k -shell indexes. This method reveals the significant features of a complex network—for example, all Internet nodes can be classified as nuclei, peer-connected components, or isolated components [1].

The SIR epidemic model [2, 26, 27] is used in many fields to study the spreading processes of information, rumors, biological diseases, and other phenomena. The model consists of three states: susceptible (S), infective (I), and recovered (R). S nodes are susceptible to information or diseases, I nodes are capable of infecting neighbors, and R nodes are immune and cannot be reinfected. Initially, almost all network nodes are in the S set, with a small number of infected nodes acting as spreaders. During each time step, I nodes infect their neighbors at a preestablished infection rate, after which they become recovered nodes at a recovery rate of γ . The total number of nodes in an SIR model is $S(t) + I(t) + R(t) = n$, with $S(t)$ denoting the number of susceptible nodes at time t , $I(t)$ the number of infected nodes at time t , $R(t)$ the number of recovered nodes at time t , and $\rho(t) = R(t)/N$ the proportion of immune nodes.

3. The Proposed Measure

Our two-step method for obtaining global and local node information in a complex network is illustrated in the following steps. In step 1, global algorithms (e.g., community detection, graph clustering, and k -shell decomposition) are used to analyze the global features of nodes, and results are used to compute their global diversity. In step 2, degree centrality is used to measure local node features. Global diversity and local features are then combined to determine the influence of complex network nodes.

In step 1, k -shell decomposition was used as an example for obtaining global node information in a complex network, with Shannon's entropy [28] used to calculate node k -shell values and to determine how many network layers are affected by a node. According to (4), maximum entropy indicates a case in which a node is capable of connecting with all layers of a complex network, and minimum entropy

(0) indicates a case in which all node connections are in the same network layer. The k -shell entropy of node i , which ensures that its neighbors' k -shell values are significantly more diverse, is defined as

$$E_i(X_i) = - \sum_{j=1}^{ks_{\max}} p_i(x_j) \cdot \log_2 p_i(x_j), \quad (4)$$

$$p_i(x_j) = \frac{|x_j|}{\sum_{j=1}^{ks_{\max}} x_j}, \quad (5)$$

$$\hat{E}_i(X_i) = \frac{E_i(X_i)}{\log_2 ks_{\max}}, \quad (6)$$

where $X_i = \{1, 2, \dots, ks_{\max}\}$ are the k -shell values of the neighbors of node i , $p_i(x_j)$ the probability of the x_j -core layer of neighbors, $|x_j|$ the number of nodes in the x_j -core layer of the complex network, and $\hat{E}_i(X_i)$ the normalized k -core entropy required for the case under consideration.

In step 2, the node's degree centrality is used to analyze the value of local features in the complex network; the degree centralities of neighbors are also considered. High influence values indicate high degree centralities of a node and its neighbors, meaning that the node is capable of reaching the widest possible local range. The local feature of node i is defined as

$$L_i(i) = \log_2 \left(\sum_{j \in \text{NB}_{h=1}(i)} C_d(j) \right), \quad (7)$$

where $C_d(j)$ is the degree centrality of neighbor j and $\text{NB}_{h=1}(i)$ is the node i neighbor set at a h -hop distance. $L_i(i)$ can be extended to become a "neighbor's neighbor" version, meaning that all node i neighbors with a 2-hop distance are considered.

Finally, E_i and L_i are combined to denote IF_i , the final influence of node i , defined as

$$IF_i = E_i \cdot L_i. \quad (8)$$

4. Results and Discussion

Basic complex network properties and results from a network GCC structure analysis are shown in Table 1. We used three network dataset classifications: scientific collaboration, traditional social, and "other." Measures were degree, betweenness, and closeness centralities; k -shell decomposition; neighbor's core (also known as *coreness*) [29]; PageRank [30]; and our proposed method. Spreading experiment and SIR epidemic model parameters were 1,000 simulations for each dataset, 50 time steps per simulation, and with the top-1 node for each measure serving as the initial spreader. β infection rates are shown in Table 1. According to at least one study, a large infection rate makes no difference in terms of spreading measures [2]. To assign a suitable infection rate for each network dataset, rates were determined by comparing the theoretical epidemic threshold β_{thd} with the number used in referenced studies [29]. Recovery rate was always $\gamma = 1$,

meaning that every node in set I entered set R immediately after infecting its neighbors.

Experimental results and details are shown in Figure 1 and Table 2. We found that the leading group could be defined as the spreading result of measures that are larger than the maximum result minus an inaccuracy factor of 1%:

$$LG = \{m \mid p_m(t) \geq (p_{\max}(t) - \text{err} * p_{\max}(t)), \\ m \in M \text{ and } \text{err} \in [0, 1]\}, \quad (9)$$

where M is the set of measures used in the experiment, $p_{\max}(t)$ is the maximum result at time t , err is the inaccuracy rate (0.01), and time step $t = 50$.

The number of recovered nodes $\rho(t)$ was used to measure and rank the spreading capability of various measures. The leading group can help determine measure stability for identifying the influence of nodes in different networks. Measures inside the leading group had approximately the same spreading capability. The average rank shown in Table 2 was used to interpret the expected rank in different networks: a measure with a lower average rank was viewed as having better discrimination in terms of identifying good spreaders.

According to the inside leading group number (a measure stability indicator), our proposed method performed well in terms of identifying the most influential network nodes and thus is capable of identifying nodes that serve as good spreaders with global diversity in a complex network. In addition to being within the leading group, the method also had a better ranking compared to other measures within that group. The identified influence spreaders were capable of reaching large numbers of network nodes through their diverse global connections, of affecting network layers, and of exerting a maximum spreading effect. Our results also indicate that the degree centrality of a node and its neighbors can be used to maintain the number of contact nodes in the local layer of a complex network. However, important differences were noted among measures. For example, the closeness measure performed well in the top-1 position of the ca-HepTh and Email-Enron networks (Figure 1, Table 2), but not in the ca-GrQc, jazz_musician, or NetScience networks. Since the characteristic the measure wanted to capture may not have been sufficiently strong in those networks, the most influential spreaders could not be identified.

Although our proposed method underscores the robustness and stability of identifying the most influential nodes, we acknowledge two limitations. First, in cases of global node diversity and lower node degree centrality, the spreading capability of nodes is constrained and dependent on the degree centrality of their neighbors. The influence of a node is limited to the local layer of a complex network when the degree centrality of its neighbors is lower. The spreading range is also limited when a node's connected neighbors are located in the network's peripheral layer. However, the spreading range of nodes may be wide when the node's neighbors are located near the hub and within the core network layers and when information and ideas can still be spread to infect a large number of nodes throughout the network.

TABLE 1: Properties of the real-world networks used in this project. We considered only the largest connected network components when the original network was disconnected.

Network type	Network	Description	N	E	$\langle c \rangle$	k_{\max}	$\langle k \rangle$	$k_{S_{\max}}$	$\langle ks \rangle$	H	r	β_{thd}	β
Collaboration	ca-AstroPh	Coauthorship in astro-ph of arxiv.org.	17903	196972	0.63	504	22.00	56	13.11	2.99	0.20	0.02	0.02
	ca-CondMat	Coauthorship in cond-mat category.	21363	91286	0.64	279	8.55	25	5.12	2.63	0.13	0.04	0.05
	ca-GrQc	Coauthorship in gr-qc category.	4158	13422	0.56	81	6.46	43	4.58	2.79	0.64	0.06	0.15
	ca-HepPh	Coauthorship in hep-ph category.	11204	117619	0.62	491	21.00	238	15.93	6.23	0.63	0.01	0.05
	ca-HepTh	Coauthorship in hep-th category.	8638	24806	0.48	65	5.74	31	3.41	2.26	0.24	0.08	0.12
Social	Jazz-Musicians	Collaborations among 1920's jazz musicians.	198	2742	0.62	100	27.70	29	17.27	1.40	0.02	0.03	0.04
	Email-Contacts	Email contacts in the Computer Science Department of the University College, London.	12625	20362	0.11	576	3.23	23	1.65	34.25	-0.39	0.01	0.05
	Email-Enron	Enron email dataset.	33696	180811	0.51	1383	10.73	43	5.73	13.27	-0.12	0.01	0.05
	CelegansNeural	Neural network of the C. Elegans nematode.	297	2148	0.29	134	14.46	10	7.98	1.80	-0.16	0.04	0.06
Other	Dolphins	Frequent associations between 62 Dolphins.	62	159	0.26	12	5.13	4	3.16	1.33	-0.04	0.15	0.15
	LesMis	Les miserables network.	77	254	0.57	36	6.60	9	4.73	1.83	-0.17	0.08	0.08
	NetScience	Network science collaborations.	379	914	0.74	34	4.82	8	3.47	1.66	-0.08	0.12	0.20
	PolBlogs	Political blogs.	1222	16714	0.32	351	27.36	36	14.82	2.97	-0.22	0.01	0.02

$H = \langle k^2 \rangle / \langle k \rangle^2$, degree heterogeneity [32].
 $\beta_{\text{thd}} = \langle k \rangle / \langle k^2 \rangle$, theoretical epidemic threshold [33].

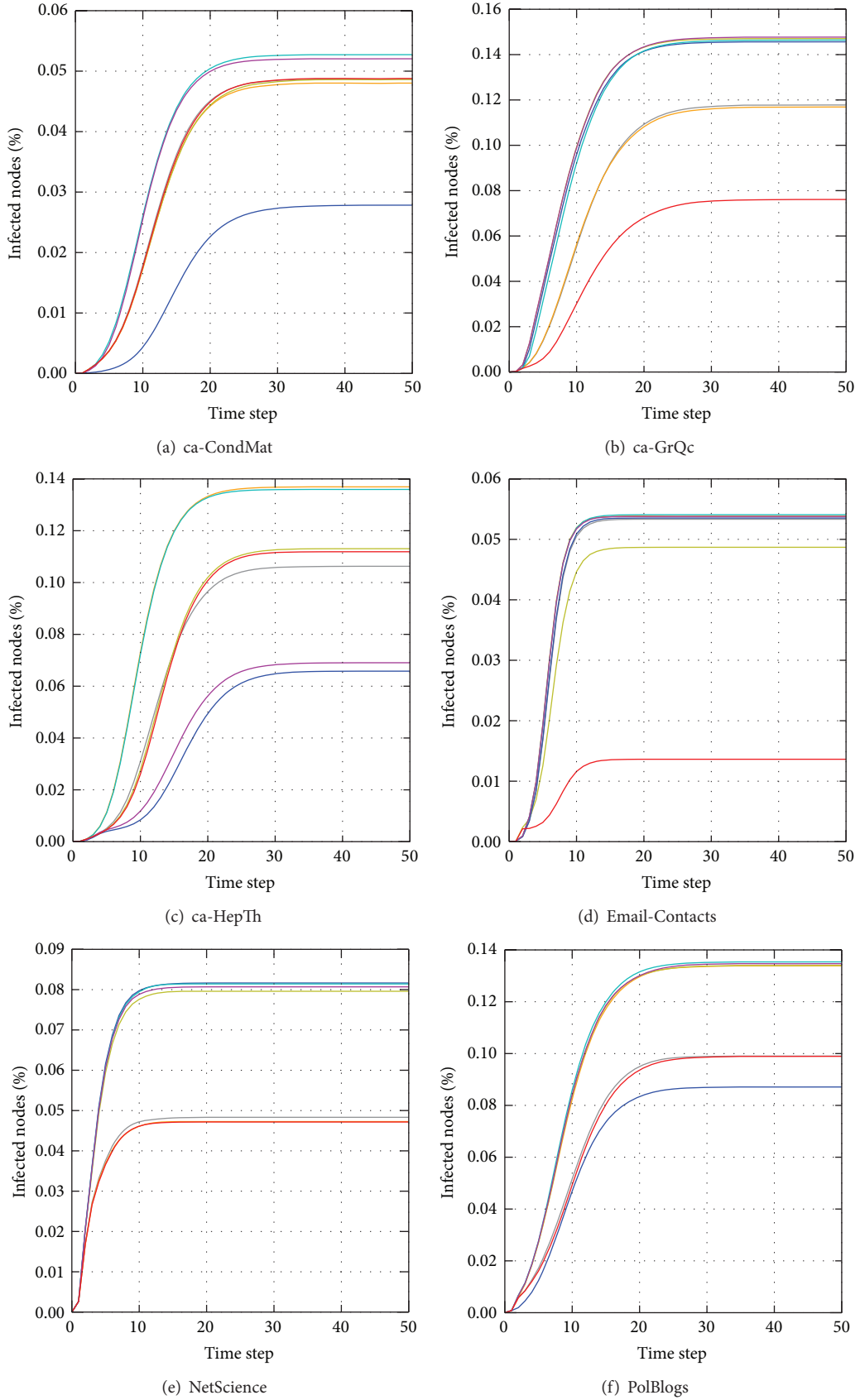


FIGURE 1: Spreading dynamic results for different networks. Measurements shown are betweenness (gray), closeness (orange), degree (yellow), k -core (blue), proposed method (cyan), neighbor-core (magenta) and PageRank (red).

TABLE 2: Comparison of simulation results from different measures (including our proposed method) in experiments using the real-world networks shown in Table 1.

Network name	$\rho(t)$ and $t = 50$						
Giant connected component (GCC)	Degree	Betweenness	Closeness	k -core	Neighbor-core	PageRank	Proposed
ca-AstroPh	0.1521 ₃	0.1499 ₅	0.1521 ₃	0.1205 ₇	0.1522 ₂	0.1523 ₁	0.1480 ₆
ca-CondMat*	0.0486 ₅	0.0487 ₄	0.0480 ₆	0.0278 ₇	0.0520 ₂	0.0488 ₃	0.0527 ₁
ca-GrQc*	0.1471 ₂	0.1178 ₅	0.1169 ₆	0.1456 ₄	0.1477 ₁	0.0761 ₇	0.1464 ₃
ca-HepPh	0.1953 ₁	0.1940 ₇	0.1952 ₂	0.1951 ₃	0.1950 ₅	0.1951 ₃	0.1943 ₆
ca-HepTh*	0.1131 ₃	0.1063 ₅	0.1369 ₁	0.0658 ₇	0.0690 ₆	0.1119 ₄	0.1359 ₂
Jazz-Musicians	0.3037 ₅	0.3050 ₄	0.3014 ₆	0.2191 ₇	0.3149 ₁	0.3055 ₃	0.3139 ₂
Email-Contacts*	0.0487 ₆	0.0533 ₅	0.0538 ₂	0.0535 ₄	0.0538 ₂	0.0136 ₇	0.0541 ₁
Email-Enron	0.1011 ₅	0.1009 ₆	0.1620 ₁	0.1618 ₄	0.1620 ₁	0.1003 ₇	0.1619 ₃
CelegansNeural	0.1939 ₃	0.1919 ₅	0.1911 ₆	0.0644 ₇	0.1926 ₄	0.2011 ₁	0.1975 ₂
Dolphins	0.1107 ₁	0.0754 ₆	0.0795 ₅	0.0702 ₇	0.1019 ₃	0.1089 ₂	0.0897 ₄
LesMis	0.0885 ₄	0.0890 ₃	0.0893 ₂	0.0792 ₆	0.0753 ₇	0.0878 ₅	0.0905 ₁
NetScience*	0.0796 ₄	0.0483 ₅	0.0473 ₆	0.0816 ₁	0.0807 ₃	0.0472 ₇	0.0814 ₂
PolBlogs*	0.1340 ₃	0.0990 ₅	0.1339 ₄	0.0871 ₇	0.1347 ₂	0.0989 ₆	0.1354 ₁
Inside leading group number:	9	7	9	5	11	7	12
Average rank:	3.4615	5.0000	3.8461	5.4615	3.0	4.6923	2.6153

* Network result shown in Figure 1.

Bold: measurement result is inside the leading network group.

Subscript: rank of network in the measurement.

Second, maximum k -shell values are lower and network sizes are considerably smaller in the absence of global diversity in a complex network. As shown in Table 2, nodes with high global diversity in the Dolphins network could not be identified. In that case, the spreading ability of nodes identified by our proposed method decreased to the degree centrality (ignoring the first term), and the influence of nodes was limited to local network layers. In the absence of global diversity, (8) becomes $IF_i \approx L_i$, which favors local network layers (i.e., degree centrality). The spreading ranges of nodes were also limited to local network layers when nodes were located in peripheral layers or inside local and dense clusters. However, broad spreading ranges were observed for nodes located in the network's core layers [2]. In addition, the $\hat{E}_i(X_i)$ normalized global diversity values produced by our proposed method are similar to the participation coefficients reported by Teitelbaum et al. [31], and the high global diversity values of nodes that we observed are similar to those of connector hubs and kinless hubs, both of which have distinct participation coefficients.

5. Conclusion

Our plans are to add considerable detail to our analysis, to introduce a sophisticated method for evaluating spreading ability, and to clarify how the proposed method is affected by network structure. For example, global algorithms such as community detection algorithms can be used to analyze and obtain global information on community network structures and to determine how factors such as position and node role [31] affect the degree to which spreaders distribute information or diseases throughout a complex network. We also plan to study strategies associated with multiple initial spreaders in networks. Since overlapping infected areas for selected spreaders must be minimized [2], a multiple initial spreader scenario may either accelerate or hinder spreading within a complex network.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The work was supported in part by a Grant from the Republic of China National Science Council (MOST-103-2221-E-182-052). The work was supported in part by the High Speed Intelligent Communication (HSIC) Research Center, Chang Gung University, Taiwan.

References

- [1] 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.
- [2] M. Kitsak, L. K. Gallos, S. Havlin et al., "Identification of influential spreaders in complex networks," *Nature Physics*, vol. 6, no. 11, pp. 888–893, 2010.
- [3] B. Hou, Y. Yao, and D. Liao, "Identifying all-around nodes for spreading dynamics in complex networks," *Physica A: Statistical Mechanics and its Applications*, vol. 391, no. 15, pp. 4012–4017, 2012.
- [4] D. Chen, L. Lü, M.-S. Shang, Y.-C. Zhang, and T. Zhou, "Identifying influential nodes in complex networks," *Physica A: Statistical Mechanics and its Applications*, vol. 391, no. 4, pp. 1777–1787, 2012.
- [5] X. Zhang, J. Zhu, Q. Wang, and H. Zhao, "Identifying influential nodes in complex networks with community structure," *Knowledge-Based Systems*, vol. 42, pp. 74–84, 2013.
- [6] A. Zeng and C.-J. Zhang, "Ranking spreaders by decomposing complex networks," *Physics Letters, Section A: General, Atomic and Solid State Physics*, vol. 377, no. 14, pp. 1031–1035, 2013.
- [7] S. Gao, J. Ma, Z. Chen, G. Wang, and C. Xing, "Ranking the spreading ability of nodes in complex networks based on local structure," *Physica A: Statistical Mechanics and its Applications*, vol. 403, pp. 130–147, 2014.
- [8] J.-G. Liu, Z.-M. Ren, and Q. Guo, "Ranking the spreading influence in complex networks," *Physica A: Statistical Mechanics and Its Applications*, vol. 392, no. 18, pp. 4154–4159, 2013.
- [9] B. Doerr, M. Fouz, and T. Friedrich, "Why rumors spread so quickly in social networks," *Communications of the ACM*, vol. 55, no. 6, pp. 70–75, 2012.
- [10] N. A. Christakis and J. H. Fowler, "Social network sensors for early detection of contagious outbreaks," *PLoS ONE*, vol. 5, no. 9, Article ID e12948, 2010.
- [11] J.-L. Wang and H.-N. Wu, "Local and global exponential output synchronization of complex delayed dynamical networks," *Nonlinear Dynamics*, vol. 67, no. 1, pp. 497–504, 2012.
- [12] J.-L. Wang and H.-N. Wu, "Adaptive output synchronization of complex delayed dynamical networks with output coupling," *Neurocomputing*, vol. 142, pp. 174–181, 2014.
- [13] J.-L. Wang and H.-N. Wu, "Synchronization criteria for impulsive complex dynamical networks with time-varying delay," *Nonlinear Dynamics*, vol. 70, no. 1, pp. 13–24, 2012.
- [14] J. Wang, Z. Yang, T. Huang, and M. Xiao, "Synchronization criteria in complex dynamical networks with nonsymmetric coupling and multiple time-varying delays," *Applicable Analysis*, vol. 91, no. 5, pp. 923–935, 2012.
- [15] J.-L. Wang, Z.-C. Yang, T. W. Huang, and M. Q. Xiao, "Local and global exponential synchronization of complex delayed dynamical networks with general topology," *Discrete and Continuous Dynamical Systems Series B*, vol. 16, no. 1, pp. 393–408, 2011.
- [16] M. E. Newman, *Networks: An Introduction*, Oxford University Press, Oxford, UK, 2010.
- [17] M. G. Kendall, "A new measure of rank correlation," *Biometrika*, vol. 30, no. 1-2, pp. 81–93, 1938.
- [18] N. Eagle, M. Macy, and R. Claxton, "Network diversity and economic development," *Science*, vol. 328, no. 5981, pp. 1029–1031, 2010.
- [19] M. Rosvall and C. T. Bergstrom, "An information-theoretic framework for resolving community structure in complex networks," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 18, pp. 7327–7331, 2007.
- [20] M. Girvan and M. E. Newman, "Community structure in social and biological networks," *Proceedings of the National Academy*

of Sciences of the United States of America, vol. 99, no. 12, pp. 7821–7826, 2002.

- [21] Stanford large network dataset collection, <http://snap.stanford.edu/>.
- [22] *Network Datasets Collected by Mark Newman*, <http://www-personal.umich.edu/~mejn/netdata/>.
- [23] Network datasets collected by Hernán Alejandro Makse, http://lisgil.engr.cuny.cuny.edu/~makse/soft_data.html.
- [24] *Jazz Musicians Network Dataset*, http://curtis.ml.cmu.edu/w/courses/index.php/Jazz_musicians_network.
- [25] P. Basaras, D. Katsaros, and L. Tassiulas, “Detecting influential spreaders in complex, dynamic networks,” *Computer*, vol. 46, no. 4, pp. 24–29, 2013.
- [26] R. Pastor-Satorras and A. Vespignani, “Epidemic dynamics and endemic states in complex networks,” *Physical Review E*, vol. 63, no. 6, Article ID 066117, 2001.
- [27] C.-Y. Huang, C.-L. Lee, T.-H. Wen, and C.-T. Sun, “A computer virus spreading model based on resource limitations and interaction costs,” *Journal of Systems and Software*, vol. 86, no. 3, pp. 801–808, 2013.
- [28] C. E. Shannon, “A mathematical theory of communication,” *The Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, 1948.
- [29] J. Bae and S. Kim, “Identifying and ranking influential spreaders in complex networks by neighborhood coreness,” *Physica A: Statistical Mechanics and its Applications*, vol. 395, pp. 549–559, 2014.
- [30] L. Page, S. Brin, R. Motwani, and T. Winograd, “The PageRank citation ranking: bringing order to the web,” Tech. Rep., Stanford InfoLab, 1999.
- [31] T. Teitelbaum, P. Balenzuela, P. Cano, and J. M. Buldú, “Community structures and role detection in music networks,” *Chaos*, vol. 18, no. 4, Article ID 043105, 2008.
- [32] H.-B. Hu and X.-F. Wang, “Unified index to quantifying heterogeneity of complex networks,” *Physica A: Statistical Mechanics and its Applications*, vol. 387, no. 14, pp. 3769–3780, 2008.
- [33] C. Castellano and R. Pastor-Satorras, “Thresholds for epidemic spreading in networks,” *Physical Review Letters*, vol. 105, no. 21, Article ID 218701, 2010.

