Physica A 391 (2012) 866–871

Contents lists available at SciVerse ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Efficient and robust routing on scale-free networks

Cun-Lai Pu [a,b,c,*], Si-Yuan Zhou [a], Kai Wang [a], Yi-Feng Zhang [a], Wen-Jiang Pei [a]

[a] *School of Information Science and Engineering, Southeast University, Nanjing 210096, People's Republic of China*
[b] *Center for Complex Network Research, Department of Physics, Northeastern University, Boston, MA 02115, USA*
[c] *Center for Cancer Systems Biology, Dana–Farber Cancer Institute, Boston, MA 02115, USA*

ARTICLE INFO

Article history:
Received 26 March 2011
Received in revised form 16 August 2011
Available online 1 September 2011

Keywords:
Network capacity
Routing
Scale-free networks

ABSTRACT

Information routing is one of the most important problems in large communication networks. In this paper we propose a novel routing strategy in which the optimal paths between all pairs of nodes are chosen according to a cost function that incorporates degrees of nodes in paths. Results on large scale-free networks demonstrate that our routing strategy is more efficient than the shortest path algorithm and the efficient routing strategy proposed by Yan et al. [Phys. Rev. E 73, 046108 (2006)]. Furthermore our routing strategy has strong robustness against cascading failure attacks on networks.

1. Introduction

The underlying structures of many interconnected networks, such as the Internet, transportation networks, and power grids, have been well studied in past decades [1–3]. The next critical topic concerned with these infrastructural networks is how to improve searching or routing efficiency and avoid traffic congestion [4–14]. Traditionally information or goods are transported along the shortest paths which have the least numbers of hops or minimal sums of link weights if the network is weighted [15,16]. This type of path usually passes through hub nodes that are highly connected, but are relatively few in real networks. When traffic flow is heavy, there will be transport delay or congestion in hub nodes, and this type of congestion will propagate to other nodes in the network [16]. This problem is the starting point of current research on traffic dynamics in networks. Two behaviors related to transport dynamics in networks are often measured in research: average delivery time and network capacity [4,5]. Although the two properties are different, they connect to each other, and usually improving one will decrease the other [4]. Most of the current work is devoted to improving network capacity, and making the average delivery time as low as possible.

The network capacity will be improved by increasing the node capacity for delivery [16], or making the structure more homogeneous such as removing the most loaded edges in networks [17,18]. However, these strategies usually cost more and are impractical. In fact, dynamics on networks can be homogenized even when the network itself is heterogeneous, which is an approach that has been previously successfully applied to the enhancement of synchronization in networks [19]. Therefore, most of the studies of traffic dynamics focus on finding better routing strategies. The efficiency of information routing strongly depends on how much the strategies utilize the information about network structure and traffic conditions. A single random walk is inefficient since it uses no information about the network [20]. Using part of the network, such as degree of neighboring nodes [6,13,21], queue length of neighboring nodes [7], geographical location of nodes [22], or local betweenness centrality [23], the routing strategy can greatly improve the network capacity. Using full information of the network, transport capacity can be improved further. Danila et al. provide a simple heuristic algorithm [4], which makes the traffic load distribution even, by recursively minimizing the maximum node betweenness in networks. Although this

* Corresponding author at: School of Information Science and Engineering, Southeast University, Nanjing 210096, People's Republic of China.
E-mail addresses: pucunlai@gmail.com, pucunlai@yahoo.cn (C.-L. Pu), wjpei@seu.edu.cn (W.-J. Pei).

0378-4371/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.physa.2011.08.044

algorithm can improve the network capacity, it is only practical for small networks due to the high computational cost. Yan et al. [5] obtain an efficient routing strategy which redistributes traffic load in central nodes to other non-central nodes. This strategy improves the network capacity more than 10 times. The network capacity can still be improved further since the load distribution in Yan's strategy is not completely homogeneous.

Robustness is a hot topic in the literature on complex networks [2]. Results for robustness show that in the Internet and the World Wide Web the giant component persists for high rates of random node removal. However, if the nodes are removed intentionally, the size of the fragments broken off increases rapidly [24,25]. In power grids and computer networks, a small fraction of the node removal can trigger a breakdown of the whole network which is called cascade failure attacks [26,27]. However, much of the current work on network robustness is devoted to the topological impact of attack and failures, and there is little research available on robustness of a routing strategy in networks. In this paper we propose a novel routing strategy based on a cost function which incorporates the degrees of nodes in the paths considered, and then we test the robustness of our routing strategy under cascade failure attacks. The outline of this paper is as follows: in Section 2 we introduce the traffic dynamics. In Section 3 we give our routing strategy. In Section 4 we test the efficiency of our routing strategy. Then in Section 5 we investigate the robustness of our routing strategy. Finally, Section 6 is devoted to the conclusion.

2. Traffic dynamics

Many large communication networks such as the Internet are scale-free [2], so we focus on scale-free networks and generate the underlying networks by using the famous Barabási–Albert (BA) model [28]. This model is defined with two steps:

(1) growth: at the beginning, there are m_0 isolated nodes. At each time step, a new node is added into the network, and this new node will connect to m ($m \leq m_0$) different nodes which are already present in the network.
(2) preferential attachment: node i will link to the new node with probability $P(k_i) = k_i / \sum_{j=1} k_j$, where k_i is the degree of node i.

This model will generate a scale-free network with $t + m_0$ nodes and mt edges in t steps. To study the traffic dynamics, we set each node in the scale-free network to be both host and router. An infinite queue with first in first out (FIFO) discipline is allocated to each node in the network. At each step, there are R information packets generated in the network. The source node and destination node of each packet are randomly allocated. Once a packet arrives at the destination node, the packet is removed from the system. If the current node where a packet arrives is not the destination, the packet will be delivered to a neighboring node according to some routing strategy used in the network. The maximum number of packets a node can deliver is set to be the constant $C = 5$.

3. Routing strategy

The major problem with traffic dynamics is traffic congestion when the traffic loads exceed the capacity of the transport systems. On the Internet, when all the packets follow the shortest paths to their destinations, they easily cause the overload of the heavily connected routers as a result of their limited capacity for delivering packets. Although traffic congestion is caused by the heterogeneous structure of the network, it is not necessary to make any changes to the network structure. We can homogenize the dynamics in the network and this is successfully applied to the enhancement of synchronization in networks [19]. When considering the delay of the delivery, the shortest paths in networks are not the best paths to deliver packets. Much research has been dedicated to finding the optimal paths for transport systems. Yan et al. provide an efficient routing (ER) strategy [5], in this ER strategy, for any path from node i to node j, as $i, \alpha_0, \alpha_1, \ldots, \alpha_n, j$, they denote:

$$\psi(\mu) = \sum_{m=0}^{n} k_m^{\mu}, \tag{1}$$

where μ is a tunable parameter, and k_m is the degree of node m. The efficient routing path between node i and node j corresponds to the route that makes the sum $\psi(\mu)$ a minimum. Simulation results demonstrate that if all nodes have the same delivery capacity in the network, then the efficient paths are the routes with the minimum sums of degrees of nodes in the paths. The idea of the ER strategy is that, in order to make the load distribute evenly, packets should avoid the heavily linked nodes since they are prone to congestion. The ER strategy increases the network capacity greatly. However, the ER strategy could still be improved. In Eq. (1) we can see that k_m is a critical element in the cost function of the ER strategy. According to this cost function the paths with high-degree nodes are ignored in scale-free networks, since the costs of these paths are significantly larger than that of the other paths. As a result high-degree nodes are not part of the optimal paths and they waste their delivery capability.

We propose an improved routing strategy called an active routing (AR) strategy. In the AR strategy, we set a cost function as follows:

$$\phi(\beta) = \sum_{m=0}^{n} (\log(\log(k_m)))^{\beta}, \tag{2}$$

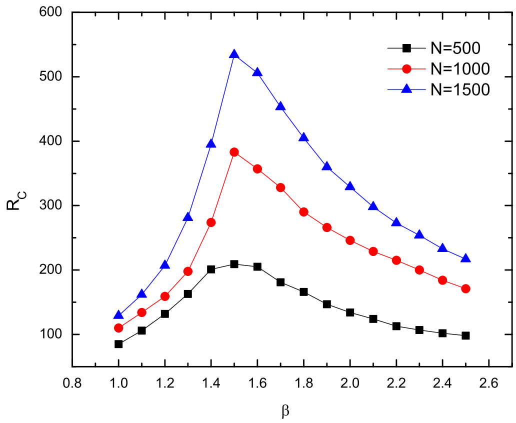

Fig. 1. The relationship between critical value R_C and β in the AR strategy on scale-free networks with different size. The average node degree is $\langle k\rangle = 6$. The power-law distribution parameter is $\gamma = 3$. The result is the average over 10 independent runs.

where β is a tunable parameter, $0 < \beta < \infty$, k_m is the degree of node m in the path. The optimal path we choose for delivering packets between each node pair is the one which makes the cost function (Eq. (2)) a minimum. When $\beta = 0$, we recover the shortest path (SP) strategy. In Eq. (2) we use $\log(\log(k_m))$ instead of k_m in the cost function which makes the costs of paths in networks more comparable. As a result, high-degree nodes also have an opportunity to be part of the optimal paths to deliver packets. From the global view of the whole network, the load of the AR strategy will distribute more evenly than that of the ER strategy. So far Eq. (2) is the best cost function we can obtain.

4. Performance of the AR strategy

Generally there are two properties that are often measured in the study of traffic dynamics on networks. One is the network capacity which indicates the maximum number of packets that a network can handle [4,5]. This property is reflected in a phase transition of the traffic dynamics. Arenas et al. defined an order parameter [29] to describe the phase transition as follows:

$$\eta(R) = \lim_{t\to\infty} \frac{1}{R}\frac{\langle \Delta W\rangle}{\Delta t}, \tag{3}$$

where $\Delta W = W(t+\Delta t)-W(t)$, $W(t)$ is the total number of packets in the network at time t, $\langle\cdots\rangle$ means average over time windows t. There is a critical value R_C, below which $\Delta W \approx 0$, $\eta \approx 0$, which means the number of generated packets and the removed packets are balanced. In this case, the traffic flow on the network is free and steady, but when $R > R_C$, $\eta > 0$, and this indicates the number of packets generated in the network is more than the amount the network can handle. Therefore, R_C is a good measure which reflects the network handling capacity. By measuring R_C, we can find out the optimal β in the AR strategy. In Fig. 1 when $\beta = 1.5$, R_C reaches the maximum value, and as the network size increases, the peak also increases. In the following discussion, we set $\beta = 1.5$ in our strategy. In this network, the maximum R_C values under different routing strategies are $R_C^{SP} = 33$, $R_C^{ER} = 435$, and $R_C^{AR} = 534$. The network capacity with the ER strategy is approximately 13 times larger than the one with the SP strategy, and the network capacity with the AR strategy is approximately 16 times larger than the one with the SP strategy. These results show that our routing strategy can improve the network capacity more than the other two strategies.

The other property is the average path length L_{ave} when the traffic is in the free flow state [5]. L_{ave} demonstrates how fast a network can transport packets, and it is calculated by averaging over the route lengths of all the optimal paths chosen from the corresponding routing strategy. In Fig. 2, we show the relationship between L_{ave} and network size N in different routing strategies. Although L_{ave}^{AR} in the AR strategy is much larger than the one in the SP strategy, it maintains the small-world property [30] for $L_{ave}^{AR} \sim \ln N$, and L_{ave}^{AR} is smaller than the one in the ER strategy.

To further study the advantage of the AR strategy, we show the load distribution in our strategy and compare it with the SP and ER strategies. The load is estimated by betweenness centrality which is well studied in social networks [31]. Similarly as in Ref. [5], the betweenness centrality in the AR strategy is as follows: for a given node degree k,

$$B(k) = \frac{1}{N(k)} \sum_{k_i=k}\sum_{s\neq t} \delta_{st}(i), \tag{4}$$

$$\delta_{st}(i) = \begin{cases} 1 & \text{if } v_s \to v_t \text{ pass through } v_i,\\ 0 & \text{otherwise,}\end{cases} \tag{5}$$

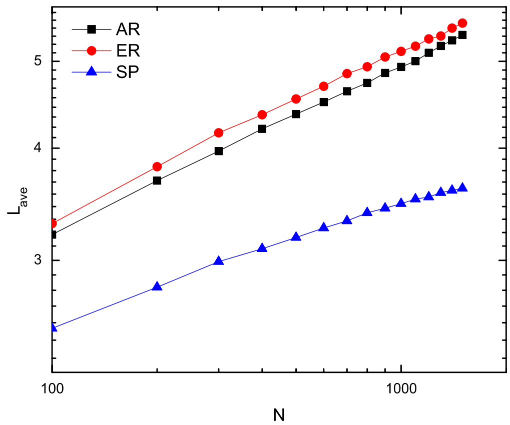

Fig. 2. The relationship between L_{ave} and network size N. The average node degree is $\langle k\rangle = 6$. The power-law parameter is $\gamma = 3$. The result is the average over 50 independent runs.

where $N(k)$ denotes the number of nodes with degree k. The relationship between load B and degree k is shown in Fig. 3. In the SP strategy $B(k) \sim k^{1.7}$. This means the large-degree nodes bear more load than the small-degree nodes. Considering the handling capacities of all the nodes are the same and fixed, large-degree nodes will congest first when the traffic is heavy, and this congestion will spread all over the network. In the ER strategy, the traffic load is redistributed from large-degree nodes to small-degree nodes. Since there are more small-degree nodes than large-degree nodes, the total handling ability of small-degree nodes is much larger than that of the large-degree nodes. As a result, the network can bear much more load in the ER strategy than in the SP strategy. The load distribution in the ER strategy is much more even than the one in the SP strategy, but it can still be improved if we increase the loads on high-degree nodes properly. In the AR strategy, the load distribution is more homogeneous, and this is why the network capacity in our strategy is higher than in the other two strategies.

5. Robustness of the AR strategy

Here we investigate robustness of routing strategies against cascade failure attacks on scale-free networks. The load $B_i(t)$ on node i at time t is the total number of the optimal paths passing through i at time t. The optimal paths are chosen according to the routing strategies studied. In the traffic model above we set the node capacity to be the same and fixed in all of the nodes, but here we assume the capacity C_i of node i to be proportional to its initial load $B_i(0)$ [26]:

$$C_i = cB_i(0), \quad i = 1, 2, \ldots, N. \tag{6}$$

$c \geq 1$ is the tolerance parameter of the network. For a given network, we assume every ordered pair of nodes exchanges one packet at each time step, and packets are transmitted along the optimal paths computed by the routing strategy studied. If a node with a small load is removed, there are no significant changes in the balance of loads, so cascade failures are unlikely to occur in the network. Here we focus on the cascade failures triggered by the removal of the node with the largest load in the network. Initially, we remove the node with the largest load. This removal changes the network topology and destroys some of the optimal paths. Then all the optimal paths are recomputed, and therefore the load distribution changes. If the loads of some nodes increase and become larger than the capacities of those nodes, these overloaded nodes will fail which may result in subsequent failures. The process will stop when there is no such node in which the load is larger than the capacity, and then we can measure some properties of the network. Here we investigate robustness of different routing strategies on scale-free networks generated by the static model [32]. We record the size of the giant component S_1 after the cascade failure, and denote the ratio between the sizes of the giant component after and before the cascade failure as H,

$$H = S_1/S_0. \tag{7}$$

Also we calculate the network efficiency [2], and this is the average of the reciprocal of the length of the optimal path which is as follows:

$$E = \frac{1}{N(N-1)} \sum_{i \neq j} \frac{1}{L_{ij}}, \tag{8}$$

where N is the size of the network, L_{ij} is the optimal path length between node i and node j.

Fig. 4 shows the simulation results. Both H and E increase rapidly with the c at the initial time, and then they tend to be stable. When c is large enough, $H \approx 1$, and E is fixed. In this case, there is no cascade failure. When c is small, $H < 1$, and E is smaller than the one when there is no attack on the network, for this case cascade failure occurs. Also we see from the picture the AR and ER strategies are more robust against cascade failure than the SP strategy. If c is fixed, H and E in the AR

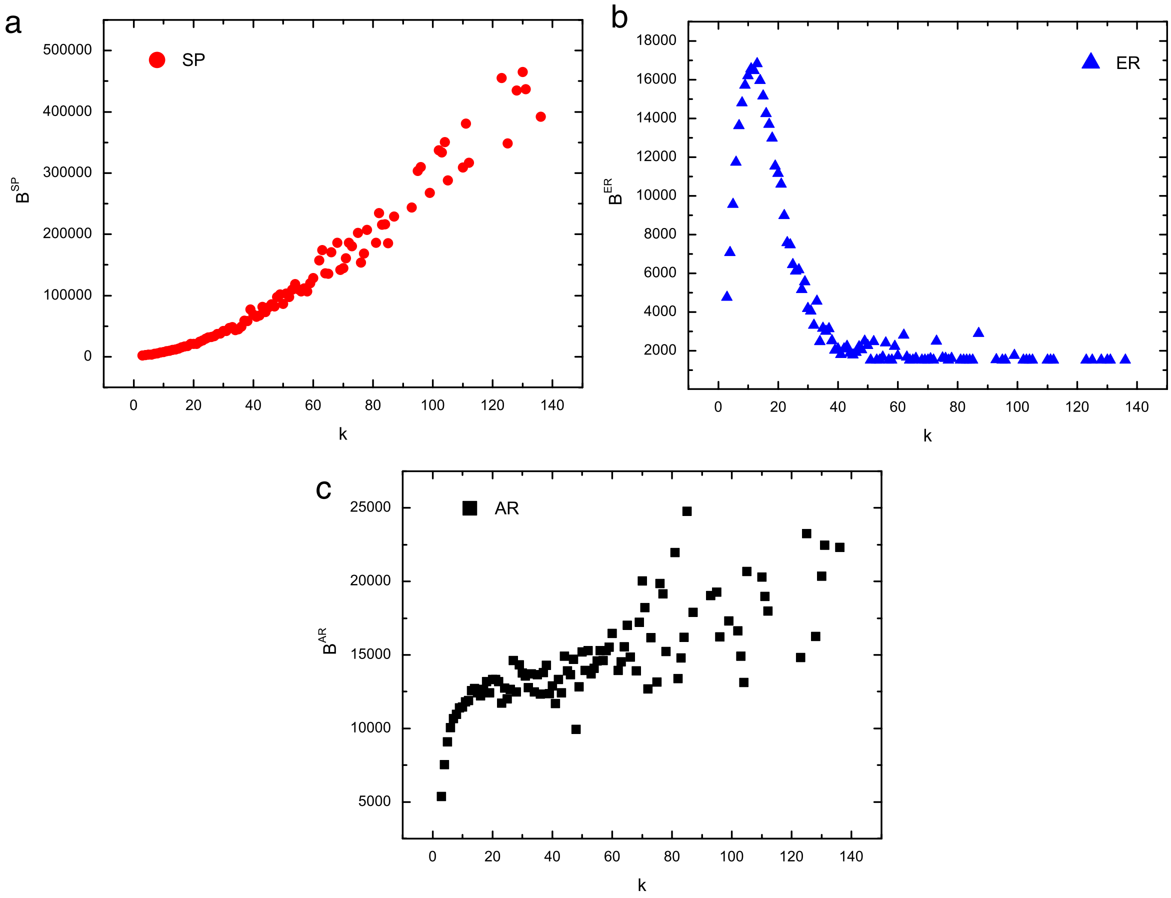

Fig. 3. The relationship between B and k for (a) SP, (b) ER, and (c) AR strategies on the scale-free network. The network size is $N = 1500$. The average node degree is $\langle k \rangle = 6$. The power-law parameter is $\gamma = 3$.

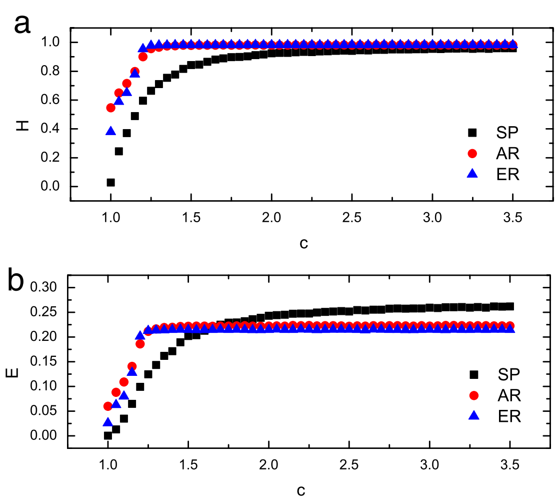

Fig. 4. (a) H vs. c and (b) E vs. c for three different strategies on the scale-free network. The network size is $N = 1000$. The average node degree is $\langle k \rangle = 6$. The power-law parameter is $\gamma = 3$.

and ER strategies are larger than in the SP strategy when there is a cascade failure in the network, and the AR strategy is slightly better than the ER strategy. From Fig. 4, for example, when $c = 1.25$, H and E in the AR strategy are about 48% larger than the ones in the SP strategy, and H and E in the ER strategy are about 43% larger than the ones in the SP strategy.

6. Conclusions

In summary we describe a new routing strategy based on our cost function to choose the optimal paths for delivering information packets in networks. In our routing strategy, the traffic load is distributed relatively evenly when compared to the case of the SP and ER strategies. As a result, the network capacity is higher in our strategy. Also, the robustness of our routing strategy against cascade failure is much better than of the SP strategy, and slightly better than that of the ER strategy. The routing table for our routing strategy is obtained by a small modification of the Dijkstra's algorithm. In the future, we will improve our routing strategy by introducing real-time information of load on nodes in the routing strategy. Also we will test the robustness of our routing strategy under other types of attacks.

Acknowledgments

The authors thank Andrew Michaelson, Dr. Chao-Ming Song, Dr. Yang Liu and Dr. Da-Shun Wang for their valuable suggestions. Also the authors would like to thank all the reviewers for their constructive comments that helped us improve this manuscript. This work was supported by the Natural Science Foundation of China under Grant No. 60672095, No. 60972165, the National High-Technology Project of China under Grant No. 2007AA11Z210, the Doctoral Fund of Ministry of Education of China under Grant No. 20070286004, the Foundation of High-Technology Project in Jiangsu Province, the Natural Science Foundation of Jiangsu Province under Grant No. BK2008281, the Special Scientific Foundation for the Eleventh-Five-Year Plan of China, the National torch plan, and the Excellent Young Teachers Program of Southeast University.

References

[1] A.L. Barabási, Science 325 (2009) 412.
[2] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, D.U. Hwang, Phys. Rep. 424 (2006) 175.
[3] C.M. Song, S. Havlin, H.A. Makse, Nature 433 (2005) 392.
[4] B. Danila, Y. Yu, J.A. Marsh, K.E. Bassler, Phys. Rev. E 74 (2006) 046106.
[5] G. Yan, T. Zhou, B. Hu, Z.Q. Fu, B.H. Wang, Phys. Rev. E 73 (2006) 046108.
[6] W.X. Wang, B.H. Wang, C.Y. Yin, Y.B. Xie, T. Zhou, Phys. Rev. E 73 (2006) 026111.
[7] W.X. Wang, C.Y. Yin, G. Yan, B.H. Wang, Phys. Rev. E 74 (2006) 016101.
[8] Z.X. Wu, W.X. Wang, K.H. Yeung, New J. Phys. 10 (2008) 023025.
[9] R. Yang, W.X. Wang, Y.C. Lai, G.R. Chen, Phys. Rev. E 79 (2009) 026112.
[10] A.P.S. de Moura, A.E. Motter, C. Grebogi, Phys. Rev. E 68 (2003) 036106.
[11] H. Zhu, Z.X. Huang, Phys. Rev. E 70 (2004) 036117.
[12] M. Rosvall, A. Grnlund, P. Minnhagen, K. Sneppen, Phys. Rev. E 72 (2005) 046117.
[13] B.J. Kim, C.N. Yoon, S.K. Han, H. Jeong, Phys. Rev. E 65 (2002) 027103.
[14] B. Tadi, S. Thurner, Physica A 346 (2005) 183.
[15] T. Zhou, Physica A 387 (2008) 3025.
[16] L. Zhao, Y.C. Lai, K. Park, N. Ye, Phys. Rev. E 71 (2005) 026125.
[17] Z. Liu, M.B. Hu, R. Jiang, W.X. Wang, Q.S. Wu, Phys. Rev. E 76 (2007) 037101.
[18] G.Q. Zhang, D. Wang, G.J. Li, Phys. Rev. E 76 (2007) 017101.
[19] A.E. Motter, C.S. Zhou, J. Kurths, Phys. Rev. E 71 (2005) 016116.
[20] J.D. Noh, Phys. Rev. Lett. 92 (2004) 11.
[21] L.A. Adamic, R.M. Lukose, A.R. Puniyani, B.A. Huberman, Phys. Rev. E 64 (2001) 046135.
[22] J.M. Kleinberg, Nature 406 (2000) 845.
[23] H.P. Thadakamalla, R. Albert, S.R.T. Kumara, Phys. Rev. E 72 (2005) 066128.
[24] R. Albert, H. Jeong, A.-L. Barabsi, Nature 406 (2000) 378.
[25] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins, J. Wiener, Comput. Netw. 33 (2000) 309.
[26] A.E. Motter, Y. Lai, Phys. Rev. E 66 (2002) 065102(R).
[27] A.E. Motter, Phys. Rev. Lett. 93 (2004) 098701.
[28] A.L. Barabási, R. Albert, Science 286 (1999) 509.
[29] A. Arenas, A. Diaz-Guilera, R. Guimera, Phys. Rev. Lett. 86 (2001) 3196.
[30] D.J. Watts, S.H. Strogatz, Nature 393 (1998) 440.
[31] M.E.J. Newman, Soc. Netw. 27 (2005) 39.
[32] K.I. Goh, B. Kahng, D. Kim, Phys. Rev. Lett. 87 (2001) 278701.