Abrupt transition to complete congestion on complex networks and control

Wen-Xu Wang,[1] Zhi-Xi Wu,[2] Rui Jiang,[3] Guanrong Chen,[2] and Ying-Cheng Lai[1,4]
[1]*Department of Electrical Engineering, Arizona State University, Tempe, Arizona 85287, USA*
[2]*Department of Electronic Engineering, City University of Hong Kong, Hong Kong SAR, China*
[3]*School of Engineering Science, University of Science and Technology of China, Hefei 230026, China*
[4]*Department of Physics, Arizona State University, Tempe, Arizona 85287, USA*

(Received 6 March 2009; accepted 30 June 2009; published online 24 July 2009)

Previous works on traffic-flow dynamics on complex networks have mostly focused on continuous phase transition from a free-flow state to a locally congested state as a parameter, such as the packet-generating rate, is increased through a critical value. Above the transition point congestion occurs on a small subset of nodes. Utilizing a conventional traffic-flow model based on the packet birth-death process and more importantly, taking into account the fact that in realistic networks nodes have only finite buffers, we find an *abrupt* transition from free flow to *complete* congestion. Slightly below the transition point, the network can support the maximum amount of traffic for some optimal value of the routing parameter. We develop a mean-field theory to explain the surprising transition phenomenon and provide numerical support. Furthermore, we propose a control strategy based on the idea of random packet dropping to prevent/break complete congestion. Our finding provides insights into realistic communication networks where complete congestion can occur directly from a free-flow state without any apparent precursor, and our control strategy can be effective to restore traffic flow once complete congestion has occurred. © *2009 American Institute of Physics*. [DOI: 10.1063/1.3184539]

Understanding the dynamics of information flow on large complex networks is a problem of considerable recent interest. An important issue is the transition from free flow to congestion, as a good understanding of the emergence of congestion is a key to devising efficient strategies to prevent and control such events. A number of previous studies have reported the existence of continuous phase transition from a free-flow state to a locally congested state. The associated congested states are typically local in that jamming occurs only on a small subset of nodes in the network. In realistic situations congestions on large scales and global congestions can emerge directly from a free-flow state, which has been reported, e.g., on the Internet. It is thus desirable to understand the occurrence of global traffic congestion on complex networks. So far, little attention has been given to the abrupt transition to global congestion in network traffic. By using a local traffic-routing model incorporating both finite packet generation and finite buffer capabilities for nodes, we find not only abrupt transition but also emergence of complete congestion state. Based on numerical evidence and a mean-field argument, we develop a theoretical understanding of this phenomenon. In addition, we articulate a control strategy to prevent/break the complete congestion and demonstrate its working on scale-free networks.

I. INTRODUCTION

Communication networks such as the Internet and various wireless telephone networks are examples of key infrastructure networks in a modern society. Due to the rapid growths of these networks, congestions of information traffic are expected to occur more and more frequently. It is a problem of broad interest to understand the dynamics of traffic flow so as to develop control strategies to alleviate/eliminate traffic congestions.[1] There have been many empirical results in this area, including the observation of phase transition from a free-flow to a congested state,[2] self-similarity of traffic flows,[3] scaling relations between the flow and fluctuations,[4,5] and congestion cascading.[6,7] Reproducing these observations by means of modeling represents a useful approach to understanding and explaining the fundamental dynamics of information traffic on various networks.[8]

Recent developments in complex networks[9–12] have inspired interest in traffic-flow dynamics on these networks. There is evidence that many large communication networks exhibit highly heterogeneous degree distributions,[9,10] generating efforts to explore the effects of different complex-network topologies on traffic dynamics.[11,12] For example, packet-hopping models with different routing strategies, a standard class of models in computer science, have been studied on networks of regular and complex topologies.[13–20] In some models, the number of packets is fixed, for which analytical theories can be developed.[19,20] However, a more realistic situation is where the number of packets changes constantly with time due to the birth and the death of packets associated with traffic flow on the network,[15–18] where death is referred to as the removal of packets after arriving at their destinations. In a packet-hopping model incorporating the birth-death process, if the packet-generation rate exceeds a critical value, a continuous phase transition from a free-flow state to congestion can occur,[15,18] regardless of the routing protocol and of the network structure. After the transition, the degree of congestion tends to become more severe as the birth rate is increased further. Based on these results, many existing works have focused on improving the network resilience to congestion by making perturbatively small adjustments to the structure of the underlying network[21–25] or by designing better routing strategies.[26–37] There is also a modeling effort to reproduce the various dynamical behaviors associated with realistic information traffic.[38–42]

Many previous works have focused on the onset of local congestion where only a few nodes are jammed. This can occur slightly beyond the onset of congestion due to the continuous nature of the phase transition from a free-flow state to congestion.[15,18,24,25,31–33,35,42] However, in a realistic communication network in the absence of effective control, congestion may occur on a much large scale *directly* from a free-flow state without any apparent precursor. A typical example is the congestion cascading behavior occurred on the Internet.[6,7] It is thus of interest to study how traffic congestion can arise on a *global* scale on complex networks from a free-flow state. So far, abrupt congestion has received relatively little attention. In Refs. 29 and 30, it was found that traffic-aware schemes can induce rich behaviors in a phase diagram, including a discontinuous phase transition between a free-flow phase and a congested phase. In this paper, we report a more severe congestion phenomenon, namely, an abrupt transition from a free-flow state to a complete congestion state as a consequence of a cascading process. The key ingredient responsible for triggering the emergence of complete congestion is the limitation of buffer capacities of nodes. We investigate the emergence of globally abrupt congestion by incorporating the finite buffer capacity into a paradigmatic birth-death packet hopping model with a local routing strategy.

In the free-flow regime, packets are forwarded to their destinations in finite times. When the packet-generation rate exceeds a critical value, *all* nodes are gradually fully occupied by packets and no packets can hop among nodes ultimately. In this case, the average time for a packet to reach its destination becomes infinite. The striking feature is that no intermediate state can exist between the two regimes. The transition from a free-flow state to complete congestion is thus abrupt, which we find is the consequence of a severe type of congestion cascading process, where once small numbers of nodes become congested, the congestion will spread from the nodes to all other nodes in the network. We also find that the maximum amount of traffic flow is achieved infinitesimally below the phase-transition point. We consider different values of the routing parameter for both homogenous and heterogeneous packet-handling abilities of nodes and obtain the optimal value of the routing parameter with respect to the phase-transition point. We then develop an analytic estimate for the optimal parameter value. Based on these results, we propose a probabilistic packet-dropping strategy to prevent a complete congestion by imposing that congested nodes have an adjustable probability to drop all packets stored in their buffers. It is thus desirable to find the optimal dropping probability that can induce the highest transmission efficiency under different conditions. This control strategy is practically implementable on real communication and information networked traffic systems.

In Sec. II, we describe our traffic-flow model. In Sec. III, we study traffic dynamics based on homogeneous delivering abilities for nodes. In Sec. IV, we investigate the flow dynamics on networks with heterogeneous node delivering abilities. In Sec. V, a control strategy is articulated and analyzed for preventing/breaking the complete congestion. Conclusions and discussions are presented in Sec. VI.

II. MODEL

Our traffic model is as follows. Any node is characterized by a delivering capability d and a finite buffer capacity B, where d denotes the maximum number of packets that can be delivered from the node at each time step and B is the maximum number of packets that can be stored in the node ready for delivery. Here, B is an adjustable parameter and identical for all nodes. With respect to the choice of the parameter d, we consider two cases. First, in an idealized setting the parameter d can be assumed to be identical for all nodes. In this case we set $d=1$ for all nodes. Second, in a realistic network, hubs can have a larger capability to process and deliver packets. It is then reasonable to assume that the delivering capability of node i is proportional to its degree k_i. For this case we shall set $d_i=k_i$.

The dynamical evolution of the traffic flow is governed by the following rules.

(1) *Packet birth.* At each time step, on each node one packet is generated with probability g. The total number of generated packets on the network at each time step is then Ng. Each generated packet is given a randomly selected destination node.

(2) *Packet routing.* At each time step, all nodes forward packets simultaneously. Node i performs a local search of destinations within its neighbors for the packets stored in its buffer. If a packet's destination is one of i neighbors, there are two possible cases: (a) if the destination's buffer is not full, the packet is delivered directly to the destination and then removed from the original node and (b) if the destination's buffer is full so that it cannot accept more packets, the packet has to stay at node i and wait for the next delivering opportunity. If the packet's destination is not in node i neighbors, we adopt a local routing strategy proposed in Ref. 31, i.e., node i selects one of its neighbors, say, node j, according to the probability $k_j^{\beta}/\Sigma_{l\in\Gamma_i}k_l^{\beta}$ for possible delivery of the packet, where β is an adjustable routing parameter and Γ_i denotes the set of neighboring nodes of node i. Even when node j buffer is not full, delivery may still fail, as the total number of delivered packets to j may exceed the available space in j buffer. (Here, the available space is the extra space in j buffer, having accommodated packets from the last time step.) In this case, we randomly select some packets to fill in j buffer, while those packets that cannot be forwarded stay at their original nodes.

(3) *Packet ordering.* At each time step, every node can deliver at most d packets, and in the buffer of each node, the last-in–first-out (LIFO) queuing rule is applied. (We

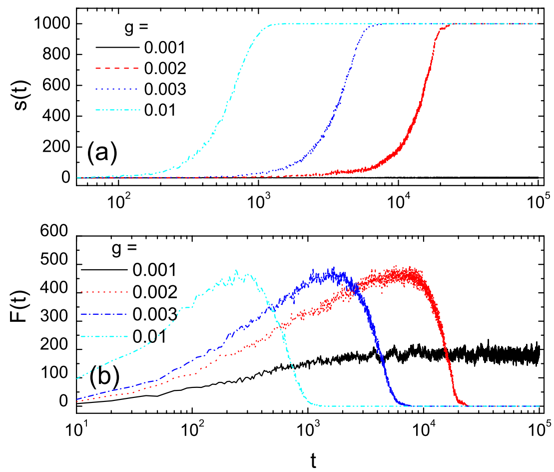

FIG. 1. (Color online) Time evolution of (a) congestion size $s(t)$ and (b) traffic flux $F(t)$ under different packet generation rate g. Simulation parameters are $B=5$, $\beta=0$, and $N=1000$.

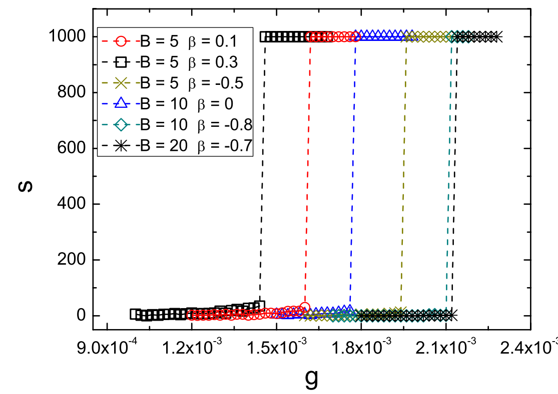

FIG. 2. (Color online) Steady-state congestion size s as a function of the packet-generation rate g for different values of the buffer capacity B and the routing parameter β. The size is calculated by averaging over 50 000 time steps after the system enters a steady state. In all cases, there is an abrupt transition to massive congestion where all nodes in the network are congested.

have checked that using the first-in–first-out rule leads to the same results as LIFO.)

The main difference between our model and previous models[15,18,24,25,31–33,35,42] is that we impose buffer limitation, which changes characteristically the corresponding routing strategy. This consideration is quite realistic as it takes into account the fact that in manmade communication networks, the storage capabilities of hosts and routers for data packets are limited by cost, and this limitation will naturally influence the routing behavior when congestion occurs. As we will see, the buffer limitation can lead to a sudden occurrence of complete congestion on the network.

There are two purposes to introduce the routing parameter β. First, from the perspective of optimized routing, the optimal value of β can be sorted out in the presence of buffer limitation to enhance the efficiency of information transmission and to avoid congestion. Second, different values of β can be used to examine the generality of the abrupt transition to global congestion with respect to different routing behaviors. The results will be shown below.

III. TRAFFIC FLOW UNDER HOMOGENEOUS NODE DELIVERING CAPABILITY

We first consider our traffic model with identical delivering capability for all nodes on scale-free networks, i.e., $d=1$. All networks simulated in this paper are generated using the standard algorithm[43] with size $N=1000$ and average degree $\langle k\rangle=6$. In a free-flow state, the number of generated packets is approximately the number of removed packets. In a jammed state, a subset of nodes can no longer deliver all received packets, leading to a continuous accumulation of packets. To gain insights into the characteristics of congestion under different packet-generation rates, we examine the congestion size s and the traffic flux F, defined as the number of congested nodes and the total number of packets moving from one node to another on the network, respectively. Representative time evolutions of s and F for different values of the generation rates g are shown in Fig. 1. We see that for relatively small values of g (e.g., for $g=0.001$), $s(t)$ is zero so that the traffic is free of any congestion. In this case, $F(t)$ increases from zero first and then reaches a steady state but with fluctuations about its average value. For larger values of g, $s(t)$ tends to increase rapidly with time to N, the network size, indicating the occurrence of a complete congestion state where all nodes in the network are congested. Once such a congested state sets in, no packets can be forwarded and removed from the network, as can be seen by the corresponding evolution of the traffic flux $F(t)$, which, after arriving at a peak, decreases quickly to zero. This observation indicates that the complete congestion state can be persistent.

The steady-state congestion size s can be regarded effectively as an *order* parameter to characterize the phase transition from free flow to congestion. Figure 2 shows s versus the packet-generation rate g for different values of the buffer capacity B and routing parameter β. We observe that $s=0$ for small values of g. The striking phenomenon is that as g is increased through a critical value, s attains the value of N *abruptly*. That is, for g immediately above g_c, a complete congestion occurs, and there exists no intermediate state where only a subset of nodes is congested. One can also use other quantities as the order parameter to characterize the transition, for example, the size of the giant component of noncongested nodes. Because the transition occurs between the free-flow state with nearly no congestion and a completely congested state, the giant component of noncongested nodes can be either the scale of network size or vanish for the two phases, respectively. The nature of the transition is thus invariant with respect to two different order parameters. The dependencies of the traffic flux on g for different parameters are shown in Fig. 3. We observe that in a free-flow state, F is an increasing function of g. At the critical point g_c, F reaches a maximum. As g is increased through g_c, F decreases to zero rather suddenly. This abrupt transition to complete congestion is characteristically different[44,45] from various continuous phase transitions reported in the literature.[15,18,24,25,28,31–33,35,37,42]

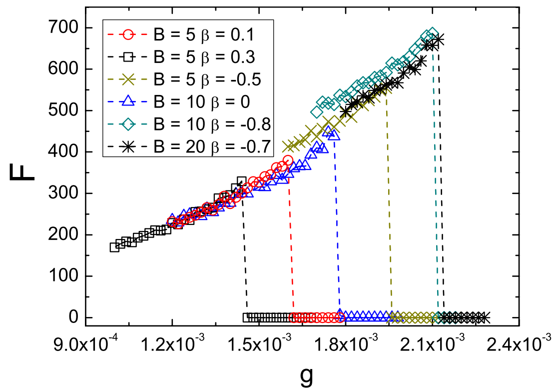

FIG. 3. (Color online) Traffic flux F as a function of g for different values of the buffer-capacity parameter B and of the routing parameter β, where F is calculated by averaging over 50 000 time steps after the system settles into a steady state.

Qualitatively, the abrupt transition from free flow to complete congestion is caused by the buffer limitation and the resulting routing behavior. Due to the buffer limitation, a node is not able to accept more packets when congested. The packets that are supposed to be delivered to the congested node from its neighbors will be accumulated at these neighboring nodes, leading to more congestions. There is a cascading process by which congestion can spread to the entire network in a short time. Besides, as reported in Refs. 29 and 30, abrupt transition can be induced by traffic-aware schemes as well. Our computations and analysis have indicated that the transition to globally complete congestion results from buffer limitation.

The results in Figs. 2 and 3 indicate that the maximally possible free-traffic flux is achieved for $g \lesssim g_c$. A larger value of g_c indicates that the underlying network is more resilient to traffic congestion. It is thus insightful to investigate the behavior of g_c with respect to variations in the buffer size B and in the routing parameter β. Figure 4 shows g_c as a function of β for different values of B. For each B value, there exists a maximum value for g_c in the middle range of β shown. The value of β for which g_c reaches a maximum is then the optimal routing parameter. For the optimal value of β, not only can the highest packet-generation rate for free flow be supported, but the traffic flux is also maximized. As B is increased, so is the maximum value of g_c. This is so because a node with a smaller buffer will be more susceptible to congestion. The optimal value of β, however, decreases with B, as indicated by the arrows in Fig. 4.

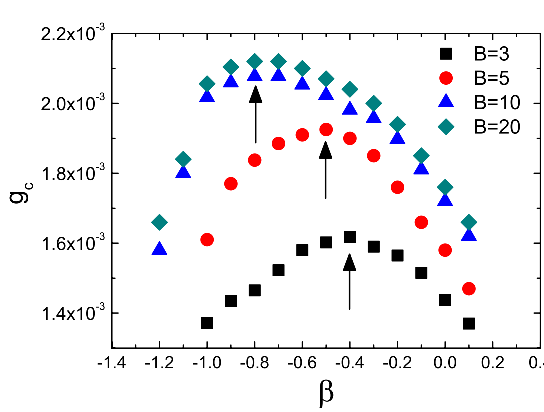

FIG. 4. (Color online) Critical generation rate g_c vs the routing parameter β for different values of the buffer capacity B. Each data point is obtained by averaging over ten network realizations and ten runs of traffic dynamics starting from random initial packet distribution for each network realization. The maximum values of g_c are marked by arrows.

To gain a quantitative understanding of the phenomenon in Fig. 4, we consider the number of packets received by a node of degree k at time t, denoted by $n_k^r(t)$. Let the number of delivered packets be $n_k^d(t)$. At time t, for $B>d$ and uncorrelated networks (no degree-degree correlation among nodes), our routing protocol gives

$$n_k^r(t) = k \sum_{k'=k_{\min}}^{k_{\max}} \frac{k' P(k')}{\langle k \rangle} n_{k'}^d(t) \frac{k^\beta}{k' \sum_{k''=k_{\min}}^{k_{\max}} \frac{k'' P(k'')}{\langle k \rangle} k''^\beta}, \tag{1}$$

where k' is the degree of a neighboring node of a node of degree k, k'' is the degree of a neighboring node of a node of degree k', $P(k')$ is the degree distribution of the network, and $k_{\min}$ and $k_{\max}$ are the minimum and the maximum degrees of the network, respectively. When the network attains the maximum packet-handling capability, the number of packets received at each node is equal to the number of nodes delivered so as to fully utilize the delivering capabilities of all nodes. Since in a free-flow state, the numbers of received and delivered packets are approximately independent of time, we can replace both $n_k^r(t)$ and $n_k^d(t)$ in Eq. (1) with d. This leads to

$$1 = \frac{k^{1+\beta}}{\langle k^{1+\beta} \rangle}, \tag{2}$$

where the following identities have been used: $\Sigma_{k=k_{\min}}^{k_{\max}} P(k)=1$ and $\Sigma_{k''=k_{\min}}^{k_{\max}} k''^{1+\beta} P(k'')=\langle k^{1+\beta} \rangle$. For a heterogeneous network, in order for Eq. (2) to be satisfied by nodes with varying degrees, the choice of β should be $\beta=-1$, which is the optimal value of the routing parameter.

The theoretical estimate $\beta_{\text{opt}}=-1$ is, however, an underestimate as compared to the numerical values. To explain the discrepancy, we study the number of packets at nodes of different degrees. As shown in Figs. 5(a) and 5(b), for $\beta=-1$, the average values of the maximum numbers of packets and the average numbers of packets for a small subset of large degree nodes are degree independent. There are fluctuations about the average values for nodes with smaller degrees. Since the buffer capacities are assumed to be the same for all nodes, larger fluctuations in the number of packets indicate a higher probability for the corresponding node to be congested. As a result, nodes of relatively small degrees are more susceptible to congestion. When a node gets congested, it can trigger subsequent congestions on its neighboring nodes because the packets delivered from the neighboring nodes cannot be accepted and have to queue. As a result, a complete congestion is more likely to occur due to congestions at low-degree nodes. To avoid this situation, the optimal routing parameter should be larger than -1 so that more packets are forwarded to higher-degree nodes. The phenomenon of fluctuation-induced congestion at lower-degree nodes can become more severe as the buffer capacity is de-

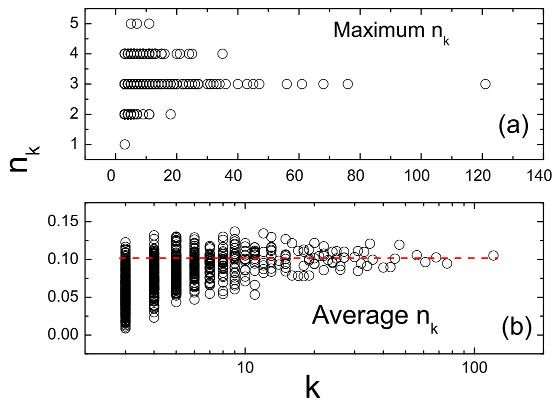

FIG. 5. (Color online) The maximum and average numbers of packets n_k in nodes as functions of their degrees k for $\beta=-1$. The scattered data at a given degree represent (a) the maximum number of packets and (b) the average number of packets on different nodes of the same degree. The average number of packets is calculated by averaging over 50 000 time steps after the system reaches a steady state. Other parameters are $g=0.0002$ and $N=1000$. The dashed line in (b) is for reference.

creased. From the perspective of control, for smaller buffer capacity, the value of β should be increased so that packets can bypass the lower-degree nodes. Traffic fluctuations at low-degree nodes are thus responsible for the deviation of the optimal routing parameter away from the theoretical estimate of -1.

IV. TRAFFIC FLOW UNDER HETEROGENEOUS NODE DELIVERING CAPACITY

A meaningful way to impose heterogeneity in node delivering capacity is to set $d_i=k_i$. Our computations have revealed, however, abrupt transitions from free-flow state to complete congestion similar to those observed under homogeneous node delivering capacities. The dependencies of the transition point g_c on the routing parameter β for different buffer capacities B are shown in Fig. 6. We again observe the existence of an optimal value of β for each B, which is larger

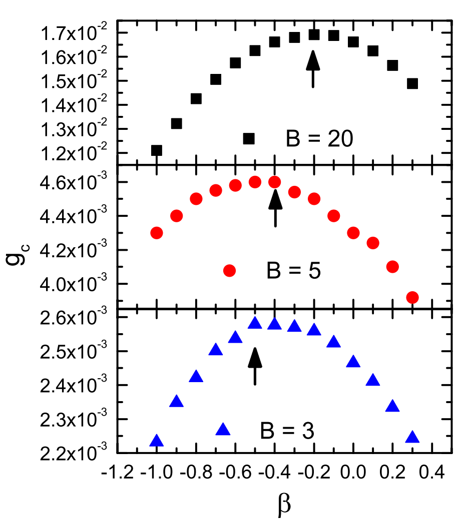

FIG. 6. (Color online) For heterogeneous node delivering capacities defined by $d_k=k$, the dependence of the critical generation rate g_c on the routing parameter β for different values of B. The maximum values of g_c are marked by arrows.

than -1. However, the dependence of the optimal value of β on B differs from that in the homogeneous case. There is in fact a positive correlation between the two quantities in the sense that a higher value of B leads to a larger optimal value for β. To understand this behavior, we again consider Eq. (1). For $B>d$, $n_k^r(t)$ is replaced by $d_k=k$. We then have

$$k = k \sum_{k'=k_{\min}}^{k_{\max}} \frac{k'^2 P(k')}{\langle k\rangle} \cdot \frac{k^\beta \langle k\rangle}{k' \sum_{k''=k_{\min}}^{k_{\max}} k''^{1+\beta} P(k'')}, \tag{3}$$

which yields

$$1 = \frac{k^\beta \langle k\rangle}{\langle k^{1+\beta}\rangle}. \tag{4}$$

Due to the heterogeneous degree distribution of the network, to satisfy Eq. (4) for all possible degrees k requires $\beta=0$. There are still, however, discrepancies between this estimate and the numerical values.

A heuristic explanation is provided in the following. For $\beta=0$, packets perform a random walk. If the buffers of nodes are unlimited, at each time, the number of packets received by a node from its neighbors is approximately proportional to the number of its neighbors for uncorrelated networks. Since the delivering capacities of nodes are proportional to node degrees as well, all nodes' delivering capacities are fully utilized. In other words, the delivering capacities match the receiving loads for all nodes so that in this case, it is unlikely that packets can accumulate in node buffers to trigger congestion cascades. Random walk is thus the optimal routing behavior for heterogeneous delivering capacity with unlimited buffer capacity. However, since the buffer capacities for all nodes are limited, higher-degree nodes are not able to accept more packets than those allowed by their buffer capacities. The excessive packets to be delivered from a neighboring node to a high-degree node are thus placed in a queue, leading to congestions at the neighboring node. As a consequence, the optimal value of β needs to be negative to allow redundant packets to bypass high-degree nodes. In addition, we observe that high-degree nodes with smaller values of B are more susceptible to congestion, which also argues for a lower value of β to avoid congestions. There is then a positive correlation between the optimal value of β and B.

V. A CONTROL STRATEGY

Control strategies have been previously proposed for mitigating traffic burden and enhancing network capacity in routing traffic, such as modifying the structure of underlying networks[21–25] and designing more efficient routing algorithms.[26,28,31–37] When a massive congestion occurs on a global scale where every node in the network is congested, these strategies will not be effective, since no nodes can receive additional packets. It is necessary to actually remove packets from the buffers of some nodes so that traffic may start to flow again. When a free-flow state is resumed, the removed packets can be redelivered from their sources to destinations. Motivated by this consideration, we propose a random packet-dropping process. In particular, let P_d be the

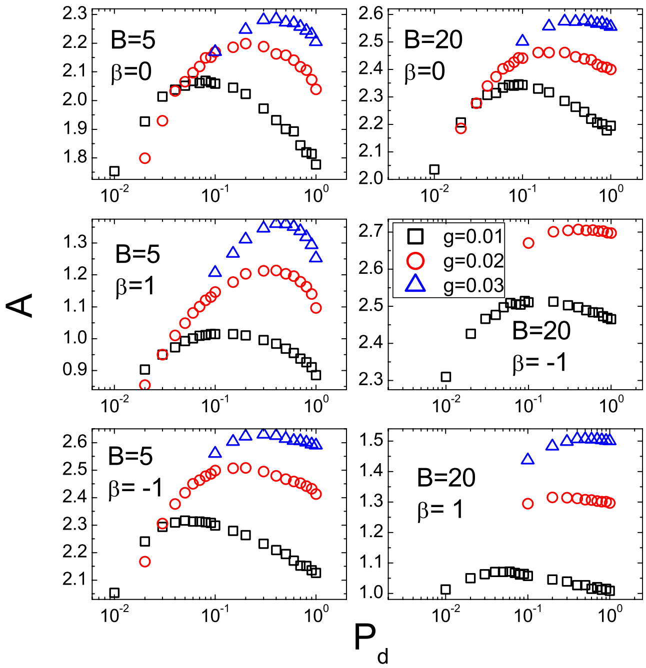

FIG. 7. (Color online) Arrival rate A as a function of random packet-dropping probability P_d for different values of g, B, and β, where A is obtained by averaging over 50 000 time steps after the system settles into a steady state. Each data point is also averaged by ten network realizations.

probability with which congested nodes drop all packets stored in their buffers. The quantity $1-P_d$ is then the probability that a congested node is not affected. The result of the random packet-dropping strategy can be quantified by the packet arrival rate A, defined as the average number of packets reaching their destinations at each time step. If the control fails to restore the traffic flow so that complete congestion persists, the value of A will be zero.

Figure 7 shows the arrival rate A as a function of the dropping probability P_d under different conditions. The dropping process is performed starting from the first time step. At each time step, each congested node drops all packets in its buffer according to probability P_d. We see that the strategy is indeed quite effective to avoid complete congestion, as reflected by the nonzero values of A in steady states. The optimal value of P_d that maximizes the value of A increases as the packet-generation rate g is increased. In addition, the value of A also increases with g. To gain insight into the occurrence of the optimal value of P_d, we consider the evolution of the total number $n_T(t)$ of packets on the network as follows:

$$\frac{dn_T(t)}{dt} = g[N - s(t)] - A(t) - s(t)P_dB, \tag{5}$$

where s is the congestion size. Since packets cannot be generated on congested nodes, $g[N-s(t)]$ is the total number of generated packets at time step t and $s(t)P_dB$ is the number of dropped packets. In a steady state, n_T, s, and A are independent of time and, hence, we have $dn_T(t)/dt=0$. Equation (5) then becomes

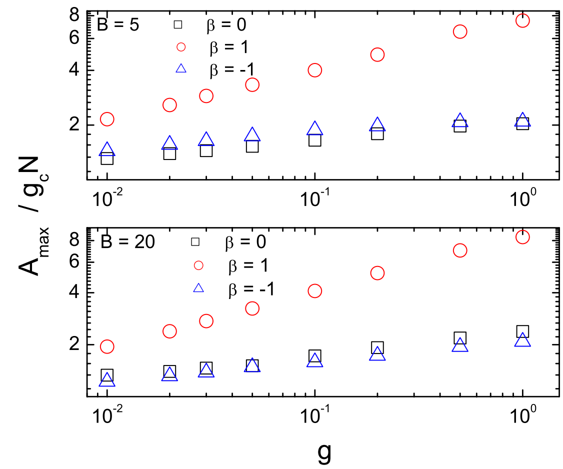

FIG. 8. (Color online) Ratio between the maximum arrival rate A_{max} and the arrival rate g_cN in the absence of control. The averaging process is the same as in Fig. 7.

$$A = gN - s(g + P_dB). \tag{6}$$

As the dropping probability P_d is increased, the congestion size s decreases. Equation (6) indicates that for fixed values of g and B, to maintain a certain value for A, increasing P_d will cause a decrease in s and vice versa. More intuitively, for a small value of P_d, the dropping strategy has a little effect on alleviating complete congestion so that only a small number of packets can reach their destinations and the value of A is close to zero. For a high value of P_d, many packets are dropped before arriving at their destinations. In this case, although complete congestion no longer occurs, the arrival rate is low. Therefore, we expect the value of A to be maximized for some value of P_d, neither too small nor too large.

In a free-flow state, due to the balance between the number of generated packets and that of removed packets, the number of packets reaching their destinations at each time step is equal to the product of the generation rate and the network size. Hence, the maximum value of A in the absence of control is g_cN. When the strategy of random-packet dropping is executed, the value of A can be maximized. Let the maximum value be A_{max}. Figure 8 shows the ratio $A_{max}/(g_cN)$ versus g for different values of B and β. We see that ratio can be enhanced for larger values of g. The fact that all values of the ratio are larger than unity indicates the effectiveness of the control strategy.

We remark that our packet-dropping strategy is applicable to traffic systems where complete congestion occurs due to cascading. Dropping packets from congested nodes can effectively inhibit the cascading process. For traffic systems with continuous phase transition where the node buffer capacity is tacitly assumed to be infinite, packets will continuously pile up at congested nodes and congestion cannot spread over the network in a cascading fashion. For such systems, our strategy will not be effective.

VI. CONCLUSIONS

We have studied the traffic dynamics on complex networks using a traffic-routing model based on local topological information under the realistic consideration that both the node delivering and buffer capacities are finite. Utilizing the congestion size as an order parameter, we have found an abrupt transition from a free-flow state to a complete congestion state. Infinitesimally below the transition point, packet generation rate and the amount of traffic flow on the network are maximized. We have also investigated the dependence of the transition point on the routing parameter and found the existence of optimal values of the routing parameter for both situations where the node-delivering capabilities are homogeneous and heterogeneous. When the routing parameter assumes the optimal value, network traffic is highly efficient in the sense that both the packet generation rate and the amount of traffic flow can be maximized. There exists a positive correlation between the optimal value of the routing parameter and the node-buffer capacity. We have utilized the mean-field approximation to explain this "resonance" phenomenon. To break the complete congestion, we have proposed and tested a random packet-dropping strategy. Our computations reveal the existence of an optimal value for the dropping probability for which complete congestion can be effectively eliminated.

ACKNOWLEDGMENTS

W.-X.W. and Y.-C.L. were supported by AFOSR under Grant No. FA9550-07-1-0045. G.C. was supported by the NSFC-HKRGC under Grant No. N_CityU 107/07. W.-X.W. and R.J. were also supported by the NNSFC under Grant Nos. 10635040 and 10532060 and the Specialized Research Fund for the Doctoral Program of Higher Education (Grant No. SRFDP 20020358009).

[1]M. Welzl, *Network Congestion Control: Managing Internet Traffic* (Wiley, New York, 2005).
[2]M. Takayasu, H. Takayasu, and T. Sato, Physica A **233**, 824 (1996).
[3]W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson, IEEE/ACM Trans. Netw. **2**, 1 (1994).
[4]M. Argollo de Menezes and A.-L. Barabási, Phys. Rev. Lett. **92**, 028701 (2004).
[5]M. Argollo de Menezes and A.-L. Barabási, Phys. Rev. Lett. **93**, 068701 (2004).
[6]V. Jacobson, Comput. Commun. Rev. **18**, 314 (1988).
[7]R. Guimerá, A. Arenas, A. Díaz-Guilera, and F. Giralt, Phys. Rev. E **66**, 026704 (2002).
[8]B. Tadić, in *Modeling Traffic of Information Packets on Graphs with Complex Topology*, Lecture Notes in Computer Science, Vol. 2657, Part I (Springer, Berlin, 2003).
[9]R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).
[10]S. N. Dorogovtsev and J. F. F. Mendes, Adv. Phys. **51**, 1079 (2002).
[11]M. E. J. Newman, SIAM Rev. **45**, 167 (2003).
[12]S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D. U. Hwang, Phys. Rep. **424**, 175 (2006).
[13]R. V. Solé and S. Valverde, Physica A **289**, 595 (2001).
[14]S. Valverde and R. V. Solé, Physica A **312**, 636 (2002).
[15]A. Arenas, A. Díaz-Guilera, and R. Guimerà, Phys. Rev. Lett. **86**, 3196 (2001).
[16]P. Holme and B. J. Kim, Phys. Rev. E **65**, 066109 (2002).
[17]B. Tadić, S. Thurner, and G. J. Rodgers, Phys. Rev. E **69**, 036102 (2004).
[18]L. Zhao, Y.-C. Lai, K. Park, and N. Ye, Phys. Rev. E **71**, 026125 (2005).
[19]A. P. S. de Moura, Phys. Rev. E **71**, 066114 (2005).
[20]R. Germano and A. P. S. de Moura, Phys. Rev. E **74**, 036117 (2006).
[21]R. Guimerà, A. Díaz-Guilera, F. Vega-Redondo, A. Cabrales, and A. Arenas, Phys. Rev. Lett. **89**, 248701 (2002).
[22]V. Cholvi, V. Laderas, L. López, and A. Fernández, Phys. Rev. E **71**, 035103(R) (2005).
[23]B. K. Singh and N. Gupte, Phys. Rev. E **71**, 055103(R) (2005).
[24]Z. Liu, M.-B. Hu, R. Jiang, W.-X. Wang, and Q.-S. Wu, Phys. Rev. E **76**, 037101 (2007).
[25]G.-Q. Zhang, D. Wang, and G.-J. Li, Phys. Rev. E **76**, 017101 (2007).
[26]B. Tadić and S. Thurner, Physica A **332**, 566 (2004).
[27]S. Valverde and R. V. Solé, Eur. Phys. J. B **38**, 245 (2004).
[28]P. Echenique, J. Gómez-Gardeñes, and Y. Moreno, Phys. Rev. E **70**, 056105 (2004).
[29]P. Echenique, J. Gómez-Gardeñes, and Y. Moreno, Europhys. Lett. **71**, 325 (2005).
[30]D. De Martino, L. Dall'Asta, G. Bianconi, and M. Marsili, Phys. Rev. E **79**, 015101(R) (2009).
[31]W.-X. Wang, B.-H. Wang, C.-Y. Yin, Y.-B. Xie, and T. Zhou, Phys. Rev. E **73**, 026111 (2006).
[32]W.-X. Wang, C.-Y. Yin, G. Yan, and B.-H. Wang, Phys. Rev. E **74**, 016101 (2006).
[33]Z. Y. Chen and X. F. Wang, Phys. Rev. E **73**, 036107 (2006).
[34]B. Danila, Y. Yu, J. A. Marsh, and K. E. Bassler, Phys. Rev. E **74**, 046106 (2006).
[35]C.-Y. Yin, B.-H. Wang, W.-X. Wang, G. Yan, and H.-J. Yang, Eur. Phys. J. B **49**, 205 (2006).
[36]S. Sreenivasan, R. Cohen, E. López, Z. Toroczkai, and H. E. Stanley, Phys. Rev. E **75**, 036105 (2007).
[37]H. Zhang, Z. Liu, M. Tang, and P. M. Hui, Phys. Lett. A **364**, 177 (2007).
[38]M. Barthélemy, B. Gondran, and E. Guichard, Phys. Rev. E **66**, 056110 (2002).
[39]E. J. Lee, K.-I. Goh, B. Kahng, and D. Kim, Phys. Rev. E **71**, 056108 (2005).
[40]J. Duch and A. Arenas, Phys. Rev. Lett. **96**, 218702 (2006).
[41]M.-B. Hu, W.-X. Wang, R. Jiang, Q.-S. Wu, and Y.-H. Wu, Phys. Rev. E **75**, 036102 (2007).
[42]M.-B. Hu, W.-X. Wang, R. Jiang, Q.-S. Wu, and Y.-H. Wu, Europhys. Lett. **79**, 14003 (2007).
[43]A.-L. Barabási and R. Albert, Science **286**, 509 (1999).
[44]For example, in Refs. 15, 18, 24, 28, 31, and 42, an order parameter η, which measures the packet-accumulation rate in the network, is used to quantify the phase transition from a free-flow state to a congestion state. In a free-flow state, the number of created and removed (after arriving at destinations) packets is balanced so that no packet can accumulate in the network, leading to $\eta=0$. In a congested state, at each time step the packet-generation rate is larger than the packet-removal rate. As a result, packets tend to accumulate on some congested nodes, leading to an increase in the value of η continuously from zero. The phase transition is thus continuous. This order parameter, however, is inappropriate for characterizing the abrupt transition phenomenon because it is zero for both the free-flow and the completely congested states. In particular, in a free-flow state η is naturally zero, but in a completely congested state, no packet can be created and, hence, no packet can move on the network, leading again to $\eta=0$.
[45]Our order parameter s can also be used to characterize a continuous phase transition for situations where the node buffer sizes are assumed to be infinite so that complete congestion does not occur. In such a case, at the onset of the phase transition, only one or a few nodes are congested and packets continuously accumulate at these nodes. The reason for the emergence of congestion at these nodes is that their delivering capability is not sufficient to handle the received packets at each time step. As a result, the queue lengths at these nodes increase with time which, however, will not affect the states of the nodes in the network that are not yet congested. This is so because the number of packets delivered from congested nodes to free nodes is determined exclusively by the delivering capabilities of the congested nodes. Since the delivering capabilities of all nodes are fixed, the number of packets delivered from congested nodes to free nodes is independent of the increase in the queue length at the congested nodes. That is, for a fixed generation rate, after congestion occurs, nodes free of congestion will remain free and only the packet accumulations at the congested nodes contribute to the increase in η from zero. After the transition, there still exist many congestion-free nodes and packets can still arrive at their destinations by passing through these nodes. This analysis is consistent with, for example, the results presented in Ref. 33 that when the packet-generation rate is not too large, the congested nodes are only a small fraction of all the nodes in the network. Since the traffic flows at majority of the nodes are locally free, the number of congested node will not increase with time, preventing complete congestion from occurring.