PHYSICAL REVIEW E **79**, 015101(R) (2009)

Congestion phenomena on complex networks

Daniele De Martino,[1] Luca Dall'Asta,[2] Ginestra Bianconi,[2] and Matteo Marsili[2]
[1]*International School for Advanced Studies SISSA and INFN, via Beirut 2-4, 34014 Trieste, Italy*
[2]*The Abdus Salam ICTP, Strada Costiera 11, 34014, Trieste, Italy*
(Received 5 August 2008; published 15 January 2009)

We define a minimal model of traffic flows in complex networks in order to study the trade-off between topological-based and traffic-based routing strategies. The resulting collective behavior is obtained analytically for an ensemble of uncorrelated networks and summarized in a rich phase diagram presenting second-order as well as first-order phase transitions between a free-flow phase and a congested phase. We find that traffic control improves global performance, enlarging the free-flow region in parameter space only in heterogeneous networks. Traffic control introduces nonlinear effects and, beyond a critical strength, may trigger the appearance of a congested phase in a discontinuous manner. The model also reproduces the crossover in the scaling of traffic fluctuations empirically observed on the Internet.

DOI: 10.1103/PhysRevE.79.015101 PACS number(s): 89.75.Hc, 02.50.Ey, 89.20.Ff, 89.75.Fb

The welfare and security of modern societies increasingly depend on the correct functioning of large networked infrastructures, such as the Internet, power grids, and socioeconomic and transportation networks. Such networks often lack a centralized administration and design, and their functioning relies on decentralized local heuristics devised in order to optimize the performances or prevent misfunctioning. However, local heuristics can hardly provide the best global performances without a deep understanding of the collective behavior of the system.

Statistical mechanics can help to understand the cooperative phenomena taking place on networks of large infrastructures [1], providing means to predict and prevent failures. Examples range from avalanches of failures in power grids [2] to credit contagion in networks of firms [3], the diffusion of email viruses [4], traffic jams in urban road networks [5,6], or congestion events in the Internet and communication networks [11–16].

Our focus here is on congestion phenomena in general. For the sake of simplicity, we shall, however, first specialize to the specific case of information networks and then discuss how our results extend to other cases. In the case of packet-based communication networks, misfunctioning is usually due to congestion phenomena that slow down the traffic, clogging large regions of the network. Congestion has been observed in wireless networks [7], in multimedia networks [8], and, more importantly, on the Internet [9]. The most common experimental signature of Internet congestion is the observation, over some interval of time, of heterogeneous distributions of the round-trip time of Ping-like signals between the same source and destination [10]. Apart from these indirect measures, congestion events are difficult to monitor and study, so that a clear phenomenological picture is still missing.

Congestion as a phase transition phenomenon has been recently studied by several authors [11,12]. Echenique *et al.* [13] have found in numerical simulations that the nature of the congestion transition depends on the type of routing rules. The effect of routing rules on network performance has also been addressed in [14] as well as the scaling of traffic fluctuations [15].

In this Rapid Communication, we put forward a minimal model of traffic, based on the idea that packet dynamics could be described by random walks on a queueing network. The model preserves all interesting features previously observed in real data and simulations, but it is simple enough to be studied analytically, shedding light on the mechanisms responsible for congestion phenomena in networks. We observe continuous or discontinuous phase transitions depending on the routing protocol, providing a theoretical insight into the findings in Ref. [13]. The discontinuous transition is observed together with a hysteresis of the order parameter indicating a region of metastable free-flow phase. We find that while traffic control can make the system more stable against congestion (especially in highly heterogeneous networks), it can also make a congested state coexist with an uncongested one, turning the transition discontinuous. Traffic control is instead deleterious in homogeneous networks. Our model also offers a simple theoretical explanation for the scaling of traffic fluctuations observed in Ref. [15].

Let us consider a network of N nodes, and let $v(i)$ denote the set of neighbors of node i. We describe traffic dynamics as a continuous-time stochastic process, in which packets are generated at each node i with a rate p_i. Each node is endowed with a first-in first-out queue. Let n_i be the number of packets in the queue of node i. If $n_i>0$, node i attempts to transmit packets at a rate r_i, which represents the bandwidth, to one of the neighbors $j \in v(i)$. We assume the following probabilistic routing protocol. First, the node j is chosen at random among the neighbors $v(i)$ of i. Second, the fate of the packet being transmitted depends (a) on whether j is the destination node for that packet and (b) on the state of congestion of node j. We model both as probabilistic events: (a) we call μ_j the probability that node j is the destination of the packet being sent, meaning that with probability μ_j the packet is "absorbed" in the transfer, and (b) we assume that the transfer is refused by node j with a probability $\eta(n_j)$, which is nondecreasing with n_j; in this case, the packet does not leave node i.

Random-walk routing looks quite different from the process assumed in other models of traffic in networks [13,16], where the packets follow a path from the source to the destination trying to minimize times, taking into account information about distance and the local traffic. In shortest pathlike routing a node i of degree k_i is visited with a probability $\propto k_i^{\beta}$, with $\beta \approx 2$, whereas in the random-walk protocol the probability of visiting node i is proportional to the degree k_i. In both, high-degree nodes are more exposed to events of congestion, which is why starlike topological structures are particularly vulnerable to congestion and perform optimally only at low traffic [16]. Our model can be easily generalized in order to accommodate for this statistical feature—for example, by considering degree-biased random walks, such as in [17]. Nevertheless, we believe that the general scenario remains unchanged. The second important ingredient of the model is the presence of traffic-aware routing protocols. This mimics *congestion avoidance* schemes elaborated by computer scientists for the Internet [18]. This class of algorithms is based on a feedback mechanism that relies on the exchange between routers of acknowledgement signals (ACKs) carrying information on the local level of traffic. When the round-trip time of ACKs sent in a given direction becomes too large, the node decreases the rate with which packets are forwarded in such a direction. As a result, congested nodes have a lower probability to receive packets, as postulated by the function η in our model.

In order to study the phase transition from the free-flow regime to a jammed phase, we consider the order parameter [11]

$$\rho = \lim_{t\to\infty} \frac{\mathcal{N}(t+\tau) - \mathcal{N}(t)}{\tau P}, \tag{1}$$

where $\mathcal{N}(t)=\Sigma_i n_i(t)$ is the total number of packets in the system at time t, $P=\Sigma_i p_i$ is the rate of creation of packets, and τ is the observation time. This order parameter represents the fraction of not adsorbed packets per unit of time in the asymptotic state. Note that a local order parameter, replacing $\mathcal{N}(t)$ by $n_i(t)$ and P by p_i, can be defined in the same way. Based on this, we define congested a node in which the average number of packets increases with time ($\langle \dot{n}_i \rangle > 0$), in the stationary state.

The asymptotic regime of the dynamics can be efficiently analyzed within a mean-field approximation, in which we assume that the packet distribution factorizes on the nodes, $\mathcal{P}(n_1,\ldots,n_N)=\Pi_i \mathcal{P}_i(n_i)$. The transition rates for n_i only depend on the number of packets in the neighboring nodes; hence, the corresponding set of master equations for $\mathcal{P}_i(n_i)$ can be solved by message passing-type algorithms on every specific network and set of parameters μ_i, p_i, r_i, and $\eta(n)$ [19].

Our focus here is on the generic nature of the congestion phase transition, and its dependence on network topology and on routing schemes, rather than on specific examples. A major insight, in this respect, is obtained rephrasing the problem in terms of ensembles of graphs. We consider uncorrelated random graphs with degree distribution $P(k)$, so that n_k represents now the average queue length of nodes in classes of degree k. We focus on the simple case $\mu_i=\mu$, $p_i=p$, and $r_i=1$ for all i, and routing protocol $\eta(n)=\overline{\eta}\theta(n-n^\star)$, where $\theta(x)$ indicates the step function. We define $q_k=P\{n_i=0|k_i=k\}$ as the probability that a node of degree k has empty queue and $\chi_k=\overline{\eta}P\{n_i \geq n^\star|k_i=k\}$ as the probability that a node of degree k refuses packets. The mean-field transition rates for nodes with degree k are

$$w_k(n\to n+1) = p + (1-\mu)(1-\overline{q})\frac{k}{z}[1-\overline{\eta}\theta(n-n^\star)],$$

$$w_k(n\to n-1) = \theta(n)(1-\overline{\chi}), \tag{2}$$

where z is the average degree, $\overline{q}=\Sigma_k q_k P(k)$, and $\overline{\chi}=\Sigma_k \frac{k}{z}\chi_k P(k)$. The average queue length n_k follows the rate equation

$$\dot{n}_k = p + (1-\mu)(1-\overline{q})\frac{k}{z}(1-\chi_k) - (1-q_k)(1-\overline{\chi}). \tag{3}$$

Note that summing over k and dividing by p we obtain a measure of the order parameter $\rho(p)$. Since $\dot{n}_k$ depends linearly on k, high-degree nodes are more likely to be congested. Therefore, in the stationary state for a given p, there exists a real-valued threshold k^* such that all nodes with $k > k^*$ are congested, whereas nodes with degree less than k^* are not congested. Congested nodes ($k>k^*$) have $q_k=0$ and $\chi_k=\overline{\eta}$. The probability distribution for the number of packets in the queue of noncongested nodes with connectivity $k < k^\star$ can be extracted by calculating the generating function $G_k(s)=\Sigma_n P_k(n_k=n)s^n$ from the detailed balance condition. This takes the form

$$G_k(s) = q_k\left\{\frac{1-(a_k s)^{n^\star}}{1-a_k s} + \frac{(a_k s)^{n^\star}}{1-(a_k-b_k)s}\right\}, \tag{4}$$

corresponding to a double exponential, where $a_k=[p+(1-\mu)\frac{k}{z}(1-q)]/[1-\overline{\chi}]$ and $b_k=\overline{\eta}[(1-\mu)\frac{k}{z}(1-q)]/[1-\overline{\chi}]$. From the normalization $G_k(1)=1$ and the condition $\dot{n}_k=0$, we get expressions for q_k, χ_k and, finally, for $\overline{q}$, $\overline{\chi}$. The value $k^\star$ is self-consistently determined, imposing the condition that nodes with $k=k^\star$ have $q_{k^\star}=0$, $\chi_{k^\star}=\overline{\eta}$, and $\dot{n}_{k^\star}=0$, which translates into the equation

$$k^\star = \frac{1-p-\overline{\chi}}{(1-\mu)(1-\overline{\eta})(1-\overline{q})}. \tag{5}$$

The above set of closed equations can be solved numerically for any degree distribution $P(k)$, and $\rho(p)$ can be accordingly computed. The solution is particularly simple for regular graphs: $k_i=z$, $\forall i$ in the limit $n^\star \gg 1$. Then the congestion-free solution with $\rho=0$ has $q_k=\overline{q}=1-p/\mu$ and $\chi_k=\overline{\chi}=0$ and it exists for $p\leq\mu$. In the congested phase, instead, all nodes have $n_i\to\infty$—i.e., $\overline{\chi}=\overline{\eta}$ and $\overline{q}=0$. This solution has $\rho=\dot{n}/p=1-(1-\overline{\eta})\mu/p$ and exists for $p\geq(1-\overline{\eta})\mu$. Therefore, in the interval $p\in[(1-\overline{\eta})\mu,\mu]$ both a congested and a free phase coexist, as shown in the inset of Fig. 1. The behavior of ρ as a function of p exhibits hysteresis: the system turns from a free phase to a congested one discontinuously as p increases at $p=\mu$, and it reverts back to the free phase from a congested phase at $p=(1-\overline{\eta})\mu$ as p decreases. This simple case also shows that traffic control is useless in homogeneous graphs, as it does not enlarge the stability region of the free phase, while making a congested phase stable for $p\in[(1-\overline{\eta})\mu,\mu]$ (see inset of Fig. 1). The case of heterogeneous graphs instead is much richer. Figure 2 com-

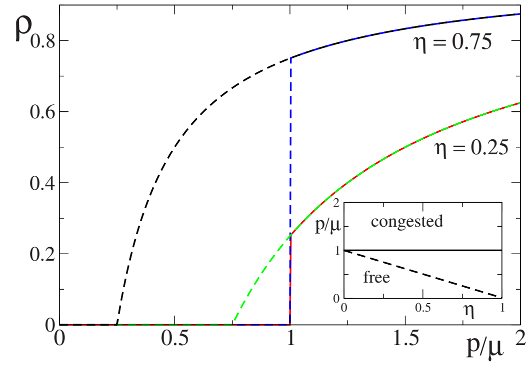

FIG. 1. (Color online) $\rho(p/\mu)$ for a homogeneous graph from theoretical predictions for η=0.25, 0.75. Inset: phase diagram for the same graph.

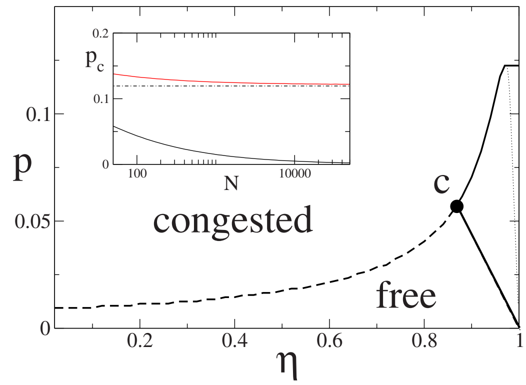

FIG. 3. (Color online) (η, p) phase diagram for an uncorrelated scale-free graph $[P(k) \propto k^{-3}$, k_{min}=3, k_{max}=110, N=3000], μ=0.2, and $n^{\star}=\infty$ from theoretical predictions. The inset shows $p_c(N)$ for $\bar{\eta}$=1 (upper line) and $\bar{\eta}$=0 (lower line).

pares $\rho(p)$ obtained from simulations (points), for a scale-free network, with our theoretical prediction (solid line). The agreement is very good, and the behavior of the curves reproduces the scenario observed in [13]. The figure is obtained for μ=0.2 and $n^{\star}$=10, but the behavior does not qualitatively change for different values of these parameters. The dependence on $\bar{\eta}$ brings instead qualitative changes. Increasing $\bar{\eta}$ from 0.05 to 0.95, the transition becomes discontinuous and p_c increases. Hysteresis is still present in case of discontinuous transition (see inset of Fig. 2).

We computed the phase diagram (see Fig. 3) in the plane $(p, \bar{\eta})$ for the same uncorrelated scale-free random networks considered in Fig. 2 for $n^{\star} \to \infty$. The dashed line represents the continuous phase transition, separating the free-flow regime from congestion. At the point C, the critical line splits into two branches that define a coexistence region. The upper solid line represents the discontinuous transition from the free-phase to the jammed state, whereas the lower indicates the opposite transition from congestion back to free flow. The dotted line decreasing from the maximum of the critical line is an unphysical branch of the analytic solution. Indeed, in the limit $n^{\star} \to \infty$, if a free-flow phase is stable for a given value of $\bar{\eta}$, this will persist for larger values of $\bar{\eta}$ because traffic control only affects congested nodes: a free stationary state cannot become congested if we increase $\bar{\eta}$. This phenomenology crucially depends on the tail of the degree distribution. Since k_{max} depends on the system's size, we expect that p_c depends on N as well; the inset of Fig. 3 shows that the critical rate of packet creation goes to zero as $p_c(N) \propto 1/\sqrt{N}$ for $\bar{\eta}$=0, but it goes to a constant for $\bar{\eta}$=1 in the limit of large N. Hence point C separates two regions in $\bar{\eta}$ with distinct behavior of finite size effects.

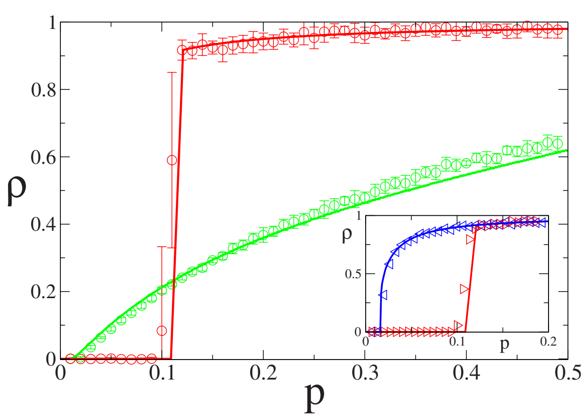

FIG. 2. (Color online) $\rho(p)$ for an uncorrelated scale-free graph $[P(k) \propto k^{-3}$, k_{min}=3, k_{max}=110, N=3000], μ=0.2, n^*=10, and $\bar{\eta}$=0.05 (below) and $\bar{\eta}$=0.95 (above), from both simulations and theoretical predictions. Inset: hysteresis circle for the same graph for $\bar{\eta}$=0.95.

We point out an important general result concerning the statistics of the number of packets at each node. In the free-flow stationary state, we generally expect the distribution of the number of packets to be exponential. Indeed, Eq. (4), with $a_k < 1$ and neglecting terms of order $a_k^{n^*} \ll 1$, yields $G_k(s) \simeq q_k/(1-a_k s)$. As a consequence, the average n_k and the variance σ_k^2 of the number of packets at a node of degree k in the free-flow stationary state stand in the relation $\sigma_k^2 = n_k(1+n_k)$. At low traffic levels ($n_k \ll 1$) we have $\sigma_k^2 \approx n_k$, whereas $\sigma_k^2 \approx n_k^2$ when the traffic increases ($n_k \gg 1$). Furthermore, since n_k increases linearly with k, we expect the first scaling to hold in low-degree nodes, whereas $\sigma_k^2 \approx n_k^2$ in nodes with large k, which is precisely the behavior reported in Ref. [15].

In conclusion, we have proposed a minimal model to study the emergence of congestion in information networks. For uncorrelated random graphs, the analysis can be performed analytically at the ensemble level, revealing that the interplay between the feedback process induced by traffic-aware routing and the topological structure of the network (in the tail of the degree distribution) generates a rich phenomenology of phase transitions [19]. Traffic-aware routing is useful only in heterogeneous networks, where it expands the region of stability of the congestion-free state. However, when its effects are strong enough, a congested phase may arise abruptly, and once it arises, it may persist even under lower traffic loads. The mechanism triggering the emergence of congestion is somewhat reminiscent of jamming or bootstrap percolation, where a node is occupied if the number of occupied neighbors exceeds a given threshold. Also in these models, as the threshold increases, the transition turns from continuous to discontinuous [20].

The work presented here can be extended in several interesting directions: First, the analysis of fluctuation phenomena in the congested phase requires an approach going beyond the mean-field approximation. The analogy with models of urban traffic is also a promising avenue. In the dual representation of the road network [5], nodes represent segments of roads and links denote junctions. A model in which adaptive drivers on a road network try to avoid congested roads exhibits a similar phenomenology [6]: as the level of traffic increases, drivers find ways to avoid overloaded streets, thus distributing as uniformly as possible the traffic load. However, at a critical threshold a congestion phase transition takes place beyond which the system is plagued by strong traffic fluctuations. Similarly, in the present model, traffic-aware routing makes the traffic load uniform in a large part of the network. The occurrence of a discontinuous phase transition can indeed be traced back to this homogenizing effect: when traffic control is strong, a finite fraction of the network can suddenly become congested upon increasing the traffic load (p). Modeling the complex adaptive behavior of human users in communication networks, such as the Internet, is a further challenge. There, users control the rates of packet production in response to network performances and they face the social dilemma of maximizing their own communication rates, maintaining the system far from the congested state [21]. In such a situation, the presence of a continuous transition may allow the system to self-organize at the edge of criticality, whereas a discontinuous transition may have catastrophic consequences. Finally, the possibility to solve the model on a given network with realistic parameters [19] allows one to draw the phase diagram of congestion phenomena for a single instance, be it the Internet map or a city's road network. This would provide both specific predictions and hints for the design of systems less vulnerable to congestion phenomena.

ACKNOWLEDGMENTS

We want to acknowledge Grant No. PRIN 2007JHLPEZ (from MIUR), which partially supported this work.

[1] R. Pastor-Satorras and A. Vespignani, *Evolution and Structure on the Internet: A Statistical Physics Approach* (Cambridge University Press, Cambridge, England, 2004).

[2] A. E. Motter, Phys. Rev. Lett. **93**, 098701 (2004).

[3] J. P. L. Hatchett and R. Kühn, J. Phys. A **39**, 2231 (2006).

[4] R. Pastor-Satorras and A. Vespignani, Phys. Rev. Lett. **86**, 3200 (2001).

[5] S. Porta, P. Crucitti, and V. Latora, Physica A **369**, 853 (2006) and references therein.

[6] A. De Martino, M. Marsili, and R. Mulet, Europhys. Lett. **65**, 283 (2004).

[7] B. Hull, K. Jamieson, and H. Balakrishnan, *ACM SenSys, Baltimore, 2004* (ACM, New York, 2004).

[8] S. Chatterjee and M. A. Bassiouni, *Proceedings of 18th IEEE Conference on Local Computer Networks, Minneapolis, 1993* (IEEE, New York, 1993).

[9] V. Jacobson, ACM SIGCOMM Comput. Commun. Rev. **25**, 157 (1988).

[10] I. Csabai, J. Phys. A **27**, L417 (1994); R. Percacci and A. Vespignani, Eur. Phys. J. B **32**, 411 (2003).

[11] A. Arenas, A. Díaz-Guilera, and R. Guimerá, Phys. Rev. Lett. **86**, 3196 (2001).

[12] T. Ohira and R. Sawatari, Phys. Rev. E **58**, 193 (1998); R. V. Solé and S. Valverde, Physica A **289**, 595 (2001); M. Takayasu, H. Takayasu, and T. Sato, *ibid.* **233**, 824 (1996); A. Y. Tretyakov, H. Takayasu, and M. Takayasu, *ibid.* **253**, 315 (1998).

[13] P. Echenique, J. Gómez-Gardeñes, and Y. Moreno, Europhys. Lett. **71**, 325 (2005).

[14] B. Tadić, S. Thurner, and G. J. Rodgers, Phys. Rev. E **69**, 036102 (2004); C. Yin, B. Wang, W. Wang, T. Zhou, and H. Yang, Phys. Lett. A **351**, 220 (2006); G. Yan, T. Zhou, B. Hu, Z.-Q. Fu, and B.-H. Wang, Phys. Rev. E **73**, 046108 (2006); A. T. Lawniczak and X. Tang, Eur. Phys. J. B **50**, 231 (2006).

[15] M. A. de Menezes and A.-L. Barabasi, Phys. Rev. Lett. **92**, 028701 (2004); S. Meloni, J. Gómez-Gardeñes, V. Latora, and Y. Moreno, *ibid.* **100**, 208701 (2008); J. Duch and A. Arenas, *ibid.* **96**, 218702 (2006); Z. Eisler, J. Kertesz, S.-H. Yook, and A.-L. Barabasi, Europhys. Lett. **69**, 664 (2005).

[16] R. Guimerà, A. Díaz-Guilera, F. Vega-Redondo, A. Cabrales, and A. Arenas, Phys. Rev. Lett. **89**, 248701 (2002).

[17] A. Fronczak and P. Fronczak, e-print arXiv:0709.2231.

[18] S. Floyd and V. Jacobson (unpublished); R. Jain and K. K. Ramakrishnan (unpublished).

[19] D. De Martino, L. Dall'Asta, G. Bianconi, and M. Marsili (unpublished).

[20] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, Phys. Rev. Lett. **96**, 040601 (2006).

[21] B. A. Huberman and R. M. Lukose, Science **277**, 535 (1997).