

# Theoretical approach and impact of correlations on the critical packet generation rate in traffic dynamics on complex networks

P. Fronczak<sup>a</sup>

Faculty of Physics, Warsaw University of Technology, Koszykowa 75, 00-662 Warsaw, Poland

Received 25 January 2012 / Received in final form 25 May 2012

Published online 22 October 2012 – © EDP Sciences, Società Italiana di Fisica, Springer-Verlag 2012

**Abstract.** Using the formalism of the biased random walk in random uncorrelated networks with arbitrary degree distributions, we develop theoretical approach to the critical packet generation rate in traffic based on routing strategy with local information. We explain microscopic origins of the transition from the flow to the jammed phase and discuss how the node neighbourhood topology affects the transport capacity in uncorrelated and correlated networks.

Transport phenomena in real networked communication systems, such as the Internet [1] and WWW [2], has turned recently more and more attention in physical and computational science. Since the rapid development of society entails demands for high transport efficiency, scientists strive to develop methods for understanding and controlling traffic congestion on communication systems.

In the basic models frequently used to mimic transport phenomena in communication networked systems [3–7], all nodes in a network are equally considered as hosts and routers for generating and delivering packets. Then, the traffic dynamics is defined as follows:

- at each time step, there are  $R$  packets generated in the system, with randomly chosen sources and destinations;
- during the next time steps packets travel around the network, and look for their destination-nodes. Once a packet arrives at its target, it is removed from the system;
- to navigate packets, each node performs a local search among its neighbors. If the packet's destination is found within the search area, it is delivered directly to the target. Otherwise, the particle is forwarded to the next node according to the prescribed strategy;
- at each time step every node can distribute/deliver at most  $C$  packets towards their destination (the fixed value of  $C$  reflects the limited router bandwidth);
- the queue length of each node is assumed to be unlimited and the FIFO (first in first out) discipline is applied.

In these models, one can distinguish between two kinds of strategies:

- in the first kind, each node has the global topological information about the network, which allows

packets to be forwarded either following the shortest path [3,4,8,9] or using the concept of betweenness [6,10,11], which measures the number of total shortest paths that pass through the given node. This kind of strategy may be practical for small or medium size networks, but not for very large networks in real communication systems such as the Internet, WWW, peer-to-peer networks [12] or urban transportation systems [13,14];

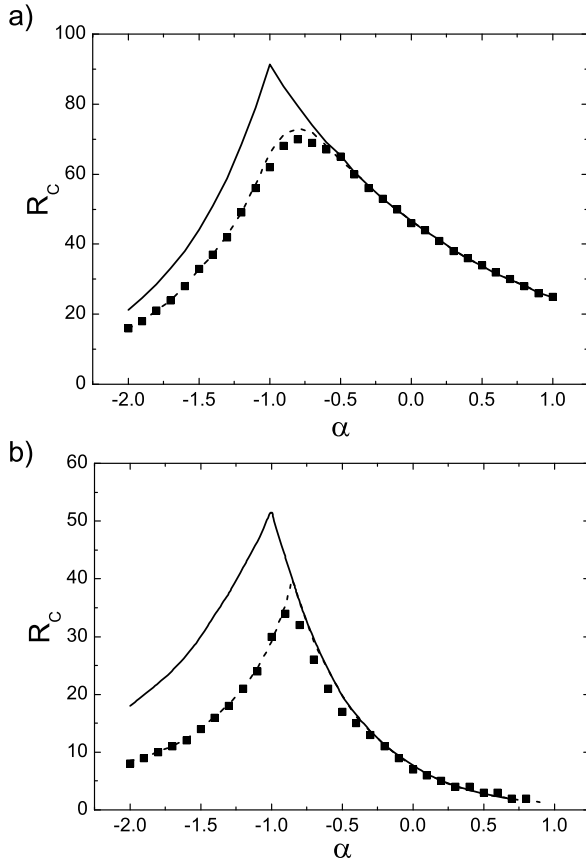
- the strategies of the second kind base on local information (each node only knows its neighbourhood) and are favored in very large networks due to heavy communication cost of searching.

One of the most important measurements for transport performance of a network is the traffic capacity,  $R_c$ , i.e. the critical packet generation rate. At  $R_c$ , the network undergoes a phase transition from the free flow state to the congested state. When the packet generation rate  $R$  is below  $R_c$ , the number of generated and delivered packets are balanced and therefore the network is in free flow state. On the other hand, when  $R$  goes beyond  $R_c$ , the number of packets keeps on increasing with time and leads to congestion, simply because nodes cannot deliver too many packets at each time step due to limited delivering capacity.

Although a number of empirical strategies for improving transport efficiency has been proposed (see the review [15] and references therein, and also [16]), the theoretical background of traffic congestion phenomena is not well developed. With reference to this theoretical line of research we would like to highlight reference [4], where the estimation of traffic capacity for the strategies of the first kind has been provided.

In this paper, we present a theoretical approach to the critical packet generation rate applied to the strategy of the second kind proposed by Wang et al. [17]. The strategy

<sup>a</sup> e-mail: fronczak@if.pw.edu.pl



**Fig. 1.** The critical packet generation rate  $R_c$  calculated for different values of the parameter  $\alpha$  with network size  $N = 1000$  and capacity  $C = 10$  in (a) classical random graphs and (b) scale-free networks  $P(k) \sim k^{-\gamma}$  with the characteristic exponent  $\gamma = 3$ . Relatively large value of average node degree  $\langle k \rangle = 12$  in case (a) and  $\langle k \rangle = 9.3$  in case (b) ensures connectivity of the networks, i.e. there exists a path between each pair of nodes. Black squares represent results of numerical simulations, solid lines correspond to theoretical prediction of equation (7), while dashed lines have been calculated from equation (7) with  $P_i^\infty$  replaced by  $P_{i\Delta}^\infty$ .

is based on the biased random walk. In this strategy the next position of the packet (node  $j$ ) is chosen according to the prescribed preferential transition probability  $w_{ij}$

$$w_{ij} = \frac{k_j^\alpha}{\sum_{m=1}^{k_i} k_m^\alpha}, \quad (1)$$

where the sum in the denominator runs over neighbors of the node  $i$ , which represents the current position of the packet, and the exponent  $\alpha$  is the model free parameter. Note that according to the formula (1) the transition probability from  $i$  to  $j$  depends only on the connectivity of the next-step node  $j$ . Note also that for  $\alpha = 0$  we recover the ordinary unbiased random walk studied by Noh and Rieger [18]. In the model, the so-called path iteration avoidance is assumed, which means that no link can be visited twice by the same packet.

The main result of this model is presented in Figure 1. The black squares (numerical simulations) show that the

optimal performance of the system (the largest traffic capacity) corresponds to  $\alpha = -1$ , which represents the anti-preferential transition probability  $w_{ij} \sim 1/k_j$ . An interesting finding arises from the comparison of Figures 1a and 1b showing results for classical random graphs and networks with power law degree distribution, respectively. In the former case, the function  $R_c(\alpha)$  has a smooth shape. In the later case the character of the relationship changes sharply. In what follows, we explain the observations with the help of a simple theoretical approach.

Let us start with the simple observation: the congested phase occurs when the average number of packets  $N_p(i, t, R)$  arriving at a certain node  $i$  at time  $t$  exceed its processing capacity  $C_i$ . Thus the critical value of the posting rate  $R_c$  should be somehow found from the balance condition

$$C_i = N_p(i, t, R_c), \quad (2)$$

which, for  $C_i = \text{const.}$ , can be also rewritten as

$$C = N_p(i, t, R_c). \quad (3)$$

Note that in the free flow state one has  $C > N_p(i, t, R)$  whereas in the congested phase there is  $C < N_p(i, t, R)$ . In the considered model,  $N_p(i, t, R)$  can be written as

$$N_p(i, R) = P_i^\infty N_p(R), \quad (4)$$

where the stationary occupation probability,  $P_i^\infty$ , describes the probability that the particle is located at the node  $i$  in the infinite time limit (for that reason we have omitted time dependence of  $N_p$ ) and  $N_p$  is the total number of packets in the network. It has been shown [19], with the help of biased random walk formalism, that

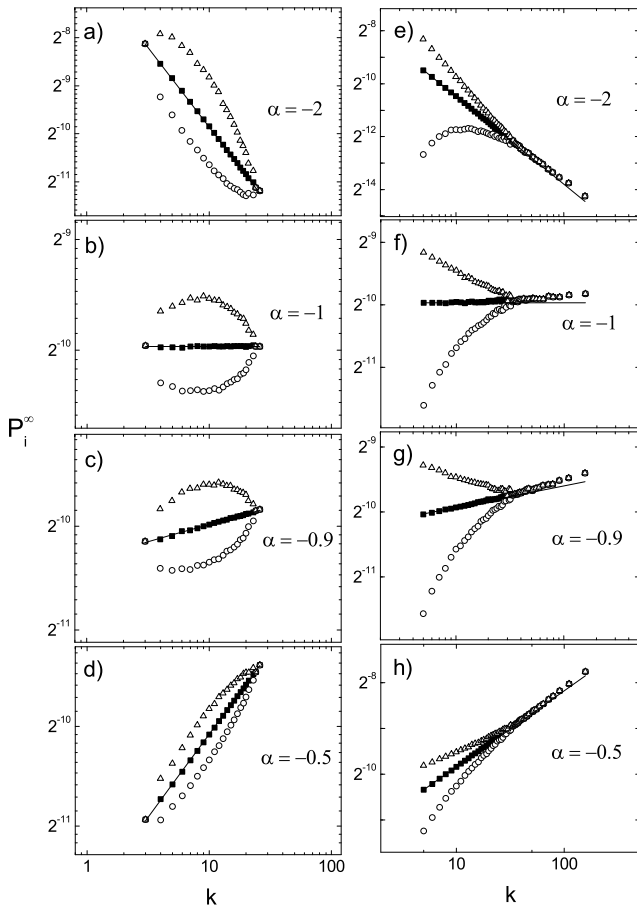
$$P_i^\infty = \frac{k_i^{\alpha+1}}{N \langle k^{\alpha+1} \rangle}, \quad (5)$$

where  $N$  is the network size. Note that for  $\alpha = 0$ , which stands for the unbiased random walk, the stationary distribution is, up to normalization, equal to the degree of the node  $i$ , i.e.  $P_i^\infty \sim k_i$ . It means that the more links a node has, the more often it will be visited by a random walker. Note also that for  $\alpha = -1$ , the stationary occupation probability is uniform  $P_i^\infty = 1/N$ . The same scaling behavior as given by equation (5) was found in reference [17] for the number of packets moving simultaneously on BA networks [20] in the free flow state. It means that in the free flow, and also in the critical point which is the limiting case of free flow state, the packets may be considered as non-interacting particles (i.e. independent biased random walkers).

The last observation allows us to calculate the total number of packets distributed over the whole network  $N_p$  in the free flow state as:

$$N_p = R \langle T_{ij} \rangle, \quad (6)$$

where  $\langle T_{ij} \rangle$  stands for the mean first-passage time averaged over all pairs of nodes.  $\langle T_{ij} \rangle$  can be understood as the mean lifetime of a packet and can be calculated theoretically from equation (19) in [19].



**Fig. 2.** Stationary occupation probabilities calculated for different node degrees  $k$  and different values of parameter  $\alpha$  in classical random graphs (a)–(d) and scale free networks (e)–(h). Solid lines correspond to theoretical prediction of equation (5). Black squares, triangles and circles are the results of numerical simulations of  $P_{\blacksquare}^{\infty}(k)$ ,  $P_{\triangle}^{\infty}(k)$  and  $P_{\circ}^{\infty}(k)$ , respectively.

Now, combining equations (3), (4) and (6) one can find the critical value of packets generation rate  $R_c$ :

$$R_c = \frac{C}{P_i^{\infty} \langle T_{ij} \rangle}. \quad (7)$$

The last equation shows that the critical value of packets generation rate is, due to  $P_i^{\infty}$ , a function of a node degree  $k_i$ . It means that in heterogeneous networks nodes of different degrees become congested for different values of  $R$ . It also means that although the system as a whole enters the jammed state even if one node is congested, it still possesses partial capacity for forwarding packets in this phase.

To find the critical value of  $R_c$ , which reflects simulation results presented in Figure 1, from the whole set of different degree dependent values of  $R_c$  one has to choose the smallest one. Since  $R_c$  is inversely proportional to  $P_i^{\infty}$ , choosing minimal value of  $R_c$  corresponds to taking nodes with the highest  $P_i^{\infty}$ . In Figure 2,  $P_i^{\infty}$  is presented as a function of a node degree  $k$  for different values of the

parameter  $\alpha$  and for two different network topologies. Solid lines correspond to theoretical prediction of equation (5) whereas black squares are results of numerical simulations. In both kinds of networks for  $\alpha < -1$ ,  $P_i^{\infty}$  is maximal for nodes with the smallest degrees and in this range of the parameter  $\alpha$  these nodes become congested first. For  $\alpha > -1$  the situation changes and congestion starts in nodes with the highest degrees.

Solid lines shown in Figure 1 represent theoretical estimation of  $R_c(\alpha)$  as given by equation (7). The lines have been calculated taking into account  $P_i^{\infty}(k_i = k_{min})$  for  $\alpha < -1$  and  $P_i^{\infty}(k_i = k_{max})$  for  $\alpha > -1$ . The mean lifetime of a packet  $\langle T_{ij} \rangle$  in equation (7) has been calculated numerically because its theoretical estimation given by averaging equation (19) in reference [19] due to applied approximation (see Eq. (16) in Ref. [19]) gives correctly only the order of magnitude of  $\langle T_{ij} \rangle$ .

Although the character of theoretical lines in Figure 1 reflects the shape of numerical results and confirms the numerically found optimal performance of the system for  $\alpha = -1$ , it is still far from perfection. The discrepancies are caused by the fact that equation (5) describes only the average occupation of a node with a given degree  $k$ . In reality,  $P_i^{\infty}(k)$  may differ among the nodes of the same degree. In general, the differences depend on the topology of a node neighbourhood. Therefore one can write formally

$$P_i^{\infty}(k) \equiv P_{\blacksquare}^{\infty}(k) = \langle p_{i,\Gamma}^{\infty}(k) \rangle, \quad (8)$$

where  $p_{i,\Gamma}^{\infty}(k)$  is a stationary occupation probability for a node with particular neighbourhood topology  $\Gamma$  and the average is calculated over all possible such topologies. The black square symbol in equation (8) has been used to reflect the fact that we are talking about results marked by black squares in Figure 2.

The triangles (circles) in Figure 2 represent stationary occupation probability calculated for the most (the least) frequently occupied node among all nodes with the same degree (the results are averaged over 100 network realizations to get rid of fluctuations)

$$\begin{aligned} P_{\triangle}^{\infty}(k) &= \max_{\Gamma} \{ p_{i,\Gamma}^{\infty}(k) \}, \\ P_{\circ}^{\infty}(k) &= \min_{\Gamma} \{ p_{i,\Gamma}^{\infty}(k) \}. \end{aligned} \quad (9)$$

The explained above scenario of congestion means that the nodes which become congested first are those represented by triangles rather than by black squares. This observation allows to understand the difference in shape of  $R_c(\alpha)$  between classical random graphs and networks with power law node degree distribution (cf. Fig. 1). In the former case, along with the increase of the parameter  $\alpha$  the initial congestion affects nodes with gradually larger and larger degrees. In the later case of scale free networks large negative values of  $\alpha$  correspond to congestion of the nodes with the lowest degrees. Nearby  $\alpha \approx -0.9$  (not  $\alpha = -1$  as it was estimated in previous studies [17]) situation changes suddenly and the highly connected nodes become responsible for the congestion (without the intermediate participation of nodes with middle degrees). The dashed line in Figure 1

calculated from equation (7) with  $P_i^\infty$  replaced by the discussed above maximal stationary occupation probability shows excellent agreement with numerical simulations.

As we have stated before, stationary occupation probability  $P_i^\infty$  calculated for nodes of the same degree is in fact the average over all possible topologies of a node neighbourhood (cf. Eq. (8)). To show the impact of neighbourhood topology on the probability we have calculated the average degree of the nearest neighbor,  $k_{nn}$ , for three groups of nodes: those marked in Figure 2 by squares (all the nodes), triangles (the most frequently occupied nodes) and circles (the least frequently occupied nodes). The results for SF networks and for  $\alpha = -1$  (corresponding to the scenario of unbiased random walk shown in Fig. 2f) are presented in Figure 3a. Although, in the case of uncorrelated networks,  $k_{nn}$  calculated for all nodes does not depend on  $k$  (what is confirmed by horizontal character of black squares), Figure 3a shows that the neighbourhood of the nodes which are responsible for congestion (i.e. in case of  $\alpha = -1$  the most frequently occupied nodes with  $k = k_{min}$ ) is composed of weakly connected nodes. Then one can suggest that increasing their  $k_{nn}$  (i.e. making their neighbourhood disassortatively correlated) may improve transport capacity of the network.

To check the hypothesis one has to introduce degree correlations to the network. From many methods for generating correlated networks (e.g. [21,22]) we select one of the simplest model proposed by Noh [23] and Newman [24]. The model belongs to a class of the exponential random graph family [25,26]. In this class, a network model is defined as a Gibbsian ensemble of networks with an associated network Hamiltonian

$$H(G) = -\frac{J}{2} \sum_{i,j=1}^N a_{ij} k_i k_j, \quad (10)$$

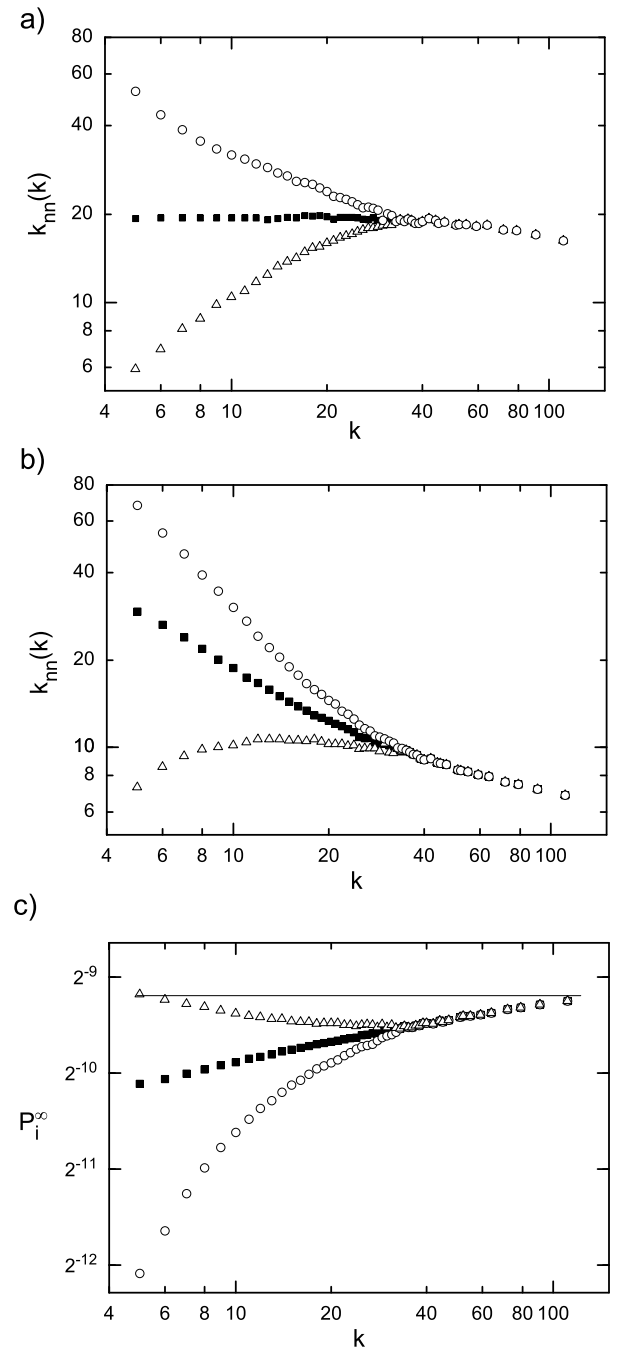
where  $J$  is a control parameter and  $a_{ij}$ , element of the adjacency matrix, takes the value  $a_{ij} = 1$  or  $0$  if nodes  $i$  and  $j$  are connected or not. A positive (negative) correlation is favored by a positive (negative) value of  $J$ . The Monte Carlo dynamics are based on updating network configurations via the link rewiring process, which preserves the degree of each node. The dynamics lead to the Gibbsian ensemble in the stationary state.

The assortativity of the network [24] is measured by the Pearson correlation coefficient of the degrees at either ends of an edge:

$$r = \frac{\langle k k' \rangle_l - \langle (k + k')/2 \rangle_l^2}{\langle (k^2 + k'^2)/2 \rangle_l - \langle (k + k')/2 \rangle_l^2}, \quad (11)$$

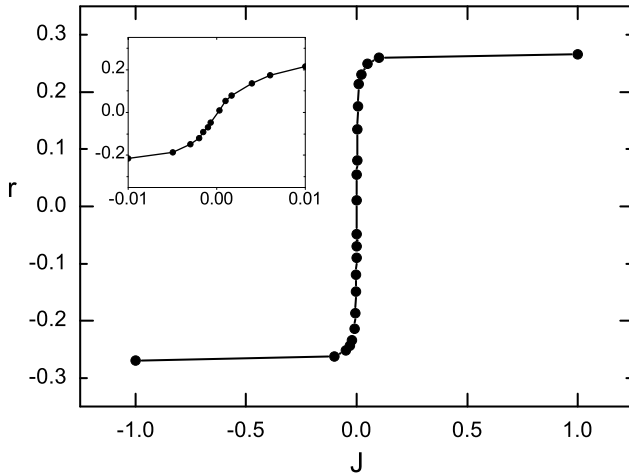
where  $\langle \cdot \rangle_l$  denotes the average over all links, whereas  $k$  and  $k'$  represent the degrees of two nodes at either end of links. The sign of  $r$  indicates a positive (assortative) or negative (disassortative) degree correlation. It vanishes for uncorrelated networks.

The stationary state value of the assortativity for SF network is presented in Figure 4. The possible values of  $r$  in correlated SF networks, which can be obtained with the



**Fig. 3.** The average degree of the nearest neighbor for all the nodes (squares), the most frequently occupied nodes (triangles) and the least frequently occupied nodes (circles) for  $r = 0.00$  (a) and  $r = -0.20$  (b). Stationary occupation probabilities calculated for different node degrees  $k$  for  $\alpha = -1$  in SF networks (c).

help of thi model belong to the range  $-0.27 \leq r \leq 0.27$ . This differentiates SF networks from networks with the Poisson degree distribution for which  $-1 \leq r \leq 1$  (cf. Fig. 2b in [23]). It seems that the possible patterns of correlation in scale-free networks are restricted by the power law degree distribution.

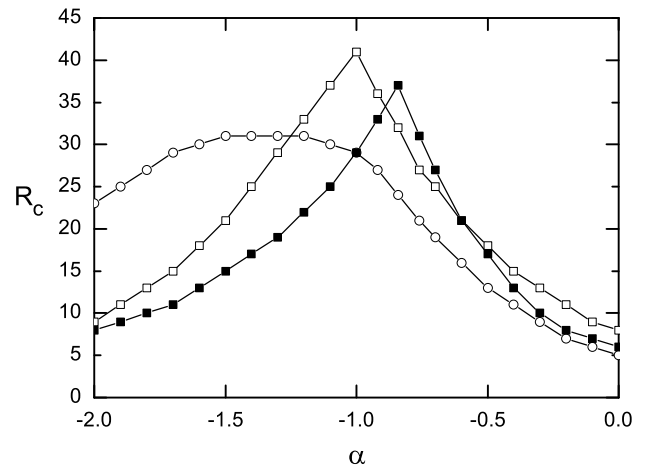


**Fig. 4.** Stationary state values of the assortativity as a function of  $J$  in SF network.

Having the tool for generating correlated networks in hands, one can repeat the calculations of  $k_{nn}$  for such networks. In Figure 3b we have shown results for the case of SF networks with  $r = -0.20$ . In such networks  $k_{nn}$  for the nodes responsible for congestion has increased from about 5.9 (in uncorrelated networks) to about 7.3. Unfortunately, further increasing of negative correlations will have an undesirable effect on the network capacity: detaching weakly connected nodes from other weakly connected nodes one have to attach them to the highly connected ones. As a consequence, the transport abilities of hubs decrease. In Figure 3c we have shown corresponding occupation probabilities for that case. As one can see,  $P_i^\infty$  for the most frequently occupied nodes with  $k_{min}$  and  $k_{max}$  equalize (what is emphasized by the horizontal line). Therefore, the case of  $r = -0.20$  is optimal, because any change of network correlations can only increase the maximal occupation probability (through increasing  $P_{i\Delta}^\infty(k_{min})$  or  $P_{i\Delta}^\infty(k_{max})$ ).

Finally, in Figure 5 we have shown the profiles of the critical packet generation rate  $R_c$  calculated for three different values of the parameter  $r$ : in the case of uncorrelated network ( $r = 0.00$ ), the case of optimally correlated network ( $r = -0.20$ ), and the case of highly correlated network ( $r = -0.27$ ). The interesting observation is that although in the most optimal case,  $k_{nn}$  calculated for the nodes responsible for congestion has increased just about one degree (cf. Fig. 3b), in the same time the critical packet generation rate  $R_c(\alpha = -1)$  has increased from 29 to 41, i.e. about 40%.

The observation, that transport capacity of a network can be enhanced for both assortative mixing and disassortative mixing has been recently reported in reference [27]. In the mentioned paper, authors have studied routing strategies with the global topological information. In our case, assortative mixing cannot improve the network capacity, since it only decrease  $k_{nn}$  of nodes responsible for congestion.



**Fig. 5.** Profiles of the critical packet generation rate  $R_c$  calculated for the case of uncorrelated network ( $r = 0.00$ ) (black squares), the case of optimally correlated network ( $r = -0.20$ ) (open squares), and the case of highly correlated network ( $r = -0.27$ ) (open circles).

In summary, using the formalism of the biased random walk in random uncorrelated networks with arbitrary degree distributions, we have developed the theoretical approach to the critical packet generation rate in traffic dynamics with the local routing strategy as proposed by Wang et al. [17]. We have shown that the random walk approach can be used to give microscopic explanation of the phase transition from free flow to the jammed phase. We have also discussed the effect of degree correlations and node neighbourhood topology on the properties of transport in complex networks.

The author wishes to thank Dr. Agata Fronczak for her valuable comments and suggestions. This work was financially supported by internal funds of the Faculty of Physics at Warsaw University of Technology.

## References

1. R. Pastor-Satorras, A. Vazquez, A. Vespignani, Phys. Rev. Lett. **87**, 258701 (2001)
2. R. Albert, H. Jeong, A.-L. Barabasi, Nature **401**, 103 (1999)
3. R. Guimera, A. Arenas, A. Diaz-Guilera, F. Giralt, Phys. Rev. E **66**, 026704 (2002)
4. L. Zhao, Y.C. Lai, K. Park, N. Ye, Phys. Rev. E **71**, 026125 (2005)
5. B. Tadic, S. Thurner, G.J. Rodgers, Phys. Rev. E **69**, 036102 (2004)
6. R. Guimera, A. Diaz-Guilera, F. Vega-Redondo, A. Cabrales, A. Arenas, Phys. Rev. Lett. **89**, 248701 (2002)
7. G. Mukherjee, S.S. Manna, Phys. Rev. E **71**, 066108 (2005)
8. L. Zhao, K. Park, Y.C. Lai, Phys. Rev. E **70**, R035101 (2004)
9. P. Holme, B.J. Kim, Phys. Rev. E **65**, 066109 (2002)

10. M. Barthlemy, Eur. Phys. J. B **38**, 163 (2003)
11. W. Huang, T.W.S. Chow, Chaos **19**, 043124 (2009)
12. W.X. Wang, C.Y. Yin, G. Yan, B.H. Wang, Phys. Rev. E **74**, 016101 (2006)
13. M.B. Hu, R. Jiang, Y.H. Wu, W.X. Wang, Q.S. Wu, Eur. Phys. J. B **63**, 127 (2008)
14. S. Scellato, L. Fortuna, M. Frasca, J. Gmez-Gardenes, V. Latora, Eur. Phys. J. B **73**, 303 (2010)
15. S.Y. Chen, W. Huang, C. Cattani, Math. Probl. Eng. **2011**, 732698 (2011)
16. D. De Martino, L. Dall'Asta, G. Bianconi, M. Marsili, Phys. Rev. E **79**, R015101 (2009)
17. W.X. Wang, B.H. Wang, C.Y. Yin, Y.B. Xie, T. Zhou, Phys. Rev. E **73**, 026111 (2006)
18. J.D. Noh, H. Rieger, Phys. Rev. Lett. **92**, 118701 (2004)
19. A. Fronczak, P. Fronczak, Phys. Rev. E **80**, 016107 (2009)
20. A.L. Barabasi, R. Albert, H. Jeong, Physica A **272**, 173 (1999)
21. M. Boguna, R. Pastor-Satorras, Phys. Rev. E **68**, 036112 (2003)
22. A. Fronczak, P. Fronczak, Phys. Rev. E **74**, 026121 (2006)
23. J.D. Noh, Phys. Rev. E **76**, 026116 (2007)
24. M.E.J. Newman, Phys. Rev. Lett. **89**, 208701 (2002)
25. J. Park, M.E.J. Newman, Phys. Rev. E **70**, 066117 (2004)
26. P. Fronczak, A. Fronczak, J.A. Holyst, Eur. Phys. J. B **59**, 133 (2007)
27. J.T. Sun, S.J. Wang, Z.G. Huang, Y.H. Wang, Physica A **388**, 3244 (2009)