Structure of Flow and Noise on Functional Scale-Free Networks

Bosiljka TADIĆ

Theoretical Physics Department, Jožef Stefan Institute, Box 3000, SI-1001 Ljubljana, Slovenia

Topological heterogeneity of structured networks effects course of dynamic processes on them by causing uneven function of nodes and/or links. Within the model of information traffic we demonstrate how dynamically relevant topology emerges on cyclic scale-free Web graph. The dynamic centrality measures are quantitatively characterized through analysis of traffic flow and multichannel noise. The dynamically generated heterogeneity of links is represented by maximum spanning tree. Fluctuations of noise at individual nodes in traffic on the original graph and on its maximum spanning tree (optimal transport) reveal universal scaling features that are related to the underlying link structure and driving conditions. Furthermore, we give some evidences for stochastic ergodicity breaking in traffic on networks.

§1. Introduction

Structured networks, in particular scale-free topologies, influence the dynamic processes which take part on these networks. In recently studied transport of information on structured networks[1] we have shown that in general more efficient processes occur on networks with higher structural complexity. The structural inhomogeneity of the underlying network causes uneven function of different nodes or links. Therefore, functionally relevant topology may differ considerably from pure geometry of links on the network.

Here we discuss how the functional topology emerge within a numerical model of information transport on a cyclic scale-free graph.[1] In particular, we focus on the dynamic analogue of the topological betweenness and betweenness-centrality (betweenness of nodes and links). In real functional networks these properties are known as multichannel noise and network flow.[2] By detailed analysis of traffic flow and noise we show that these dynamic centrality measures exhibit certain universal features that can be related to the network structure. Furthermore, analysis of the occupation probability of nodes reveals evidence of a stochastic ergodicity breaking due to network structure and interaction between traffic streams.

§2. Structure and Traffic on the Web graph

We focus on traffic of information packets on one class of cyclic scale-free networks—the Web graph.[1),4)] On this class of graphs we have shown earlier that efficient traffic emerges when the next-neighbour search rule is applied to navigate packets through the graph. This topology–dynamics matching enables us to study the details of the dynamic process in much wider range of parameters (i.e., posting rate) than on any other scale-free structure. In particular, the traffic remains jamming free until a large density of packets is reached. Some details of structure and traffic rules are described below and in Ref.[1),3),4)]

2.1. *Web graph structure*

The Web graph that we consider belongs to causal type of networks, which are grown in time by adding a node and a link at each time step.[5] The added node is linked to the preceding group of nodes with probabilistic rules introduced in Ref.[4] The linking rules consist of *preferential attachment* and *rewiring* during the graph growth, which leads to the scale-free structure of both *in-coming* and *out-going* links. The parameter α (initial attractiveness of nodes) and $\tilde{\alpha}$ (rewiring probability) control the emergent degree distribution for in- and out-links q_{in} and q_{out} according to[4]

$$P(q_{in}) \sim q_{in}^{-(2+\alpha)} \; ; \; P(q_{out}) \sim q_{out}^{-(1+(1+\alpha)/(1-\tilde{\alpha}))} \; , \qquad (2\cdot1)$$

and graph *flexibility* $\beta \equiv (1-\tilde{\alpha})/\tilde{\alpha}$. Here we use the original one-parameter model[4] with $\tilde{\alpha} = \alpha$. For $\alpha = 1/4$ the structure appears to have two powerful hub nodes, similar to the real Web (see[4] for details). A part of the structure is shown in Fig. 1.

Other structural characteristics of the Web graph are described in the literature.[1),4),7)] Here we mention the connectivity profile $q_{in}(s) \sim s^{-\gamma_{in}}$ for in-linking, and similarly $q_{out}(s) \sim s^{-\gamma_{out}}$for out-linking connectivity, that vary with node's addition time s (proportional to node rank). The exponents relation $\tau_{in} = 1 + \gamma_{in}^{-}1$ and $\tau_{out} = 1 + \gamma_{out}$ holds. In addition, the network's inhomogeneity is characterized by the clustering profile, which measures number of elementary triangles attached to a node. It varies throughout the network as $n_{\Delta}(s) \sim 1/s$. The Web graph is also classified as a *correlated* scale-free structure in which correlation between in-coming and out-going links at neighbouring nodes i, j occurs as $< q_{out}(j) > \sim q_{in}(i)^{-\kappa}$, with $\kappa \approx 0.45$.[7]

For the analysis of traffic on the Web graph, the structure of the graph is determined by its adjacency matrix and kept fixed throughout the simulations.

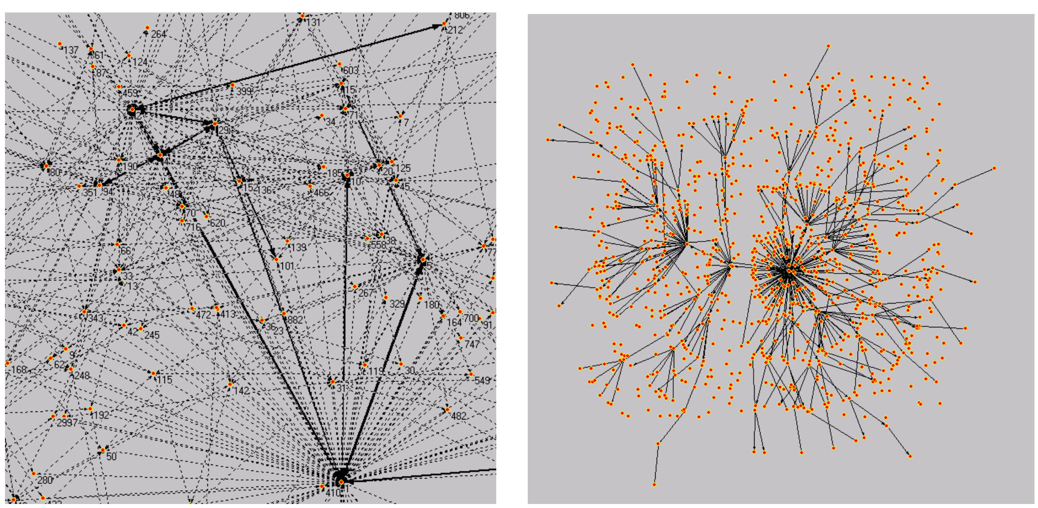

Fig. 1. Core of the Web graph near the main hubs with most often used links shown with strong lines (left) and the maximum spanning tree of the giant component (right) for $N = 1000$ nodes and links and packet creation rate $R = 0.005$. [Plots using *Pajek*.[6]]

2.2. *Information Traffic Model*

Transport of information packets on network of a given architecture[1),3)] consists of the following steps:

- (i) *creation* of a packet at a random node and *assignation* of an address (destination) as another node on the network where it should arrive;
- (ii) *navigation* of a packet through the network, where a visited node directs the packet towards its address using locally available links;
- (iii) *delivery* by arrival at its destination, packet is removed from the traffic.

In step (ii) we apply an advanced local search (CS) in which each node explores its surrounding within next-nearst neighborhood searching for the best direction for a packet to be processed.[1)] The packets are created with a fixed rate R and move simultaneously, making queues at nodes along the path. We employ last-in-first-out queue (LIFO) and impose finite queue capacity $H = 1000$ packets. Details of the numerical code implementation is given in Ref.[7)]

Within the simulations we monitor motion of each packet simultaneously. In particular, we record packet destination, current position on the network and position in the queue, as well as waiting time that a packet spent on each node along its path before arrival to its destination. From these data we make the statistics of the traffic both on global (network) level and on local (individual node and link) level. For instance, the transit time T of each packet is then given by sum of all waiting times along its path from the creation node to delivery at its address, $T_k = \sum_i^k t_w(i)$, where $t_w(i)$ is waiting time at node i on the path of length k. The actual path depends on the network structure and the search algorithm. In addition, the waiting times at each node are created when the packet density increases. The queue lengths are unevenly distributed through the network[3)] with longer queues at more important nodes. We first discuss properties of the traffic at local level.

§3. Traffic Flow and Noise

The information traffic on the Web graph with the rules described above was shown to be stationary for a large range of creation rates $R \leq R_c \approx 0.4$.[1),3)] In the stationary regime, as local measures of the traffic we determine the traffic *flow* $f_{ij}$and *noise* h_i. Flow is defined as number of packet traversing a given link in the network within a given time window T_{WIN}. Similarly, we define (multichannel) *noise* as number of packets that are processed by a given node within T_{WIN}. In this way, f_{ij} and h_i are the dynamic measures of the topological propeties known as betweenness-centrality of links and nodes respectively.

3.1. *Dynamic weights and maximum spanning tree*

The records of flow on links yield different "weights" to links that processed different number of packet within the monitoring time. This is an example of emergent *weighted network* where weights appear dynamically through the network function. In Fig. 1 we show an example with most weighted links appearing near the main hubs of the Web graph.

Uneven distribution of traffic flow on the links of networks can be studied in more quantitative details. One way to characterize the emergent weighted topology is to construct the *maximum spanning tree*, which is the tree structure spanning the connected part of the graph and consisting of *strongest links*. In Fig. 1 we also show the maximum spanning tree corresponding to the traffic on the Web graph when the posting rate is $R = 5 \times 10^{-3}$. It appears to be the scale-free tree, suggesting that the flow on the links is distributed *in accordance*[8] with the original topology of the graph. This supports the idea of matching between the structure of the Web graph and transport with the advanced-local search algorithm.[1]

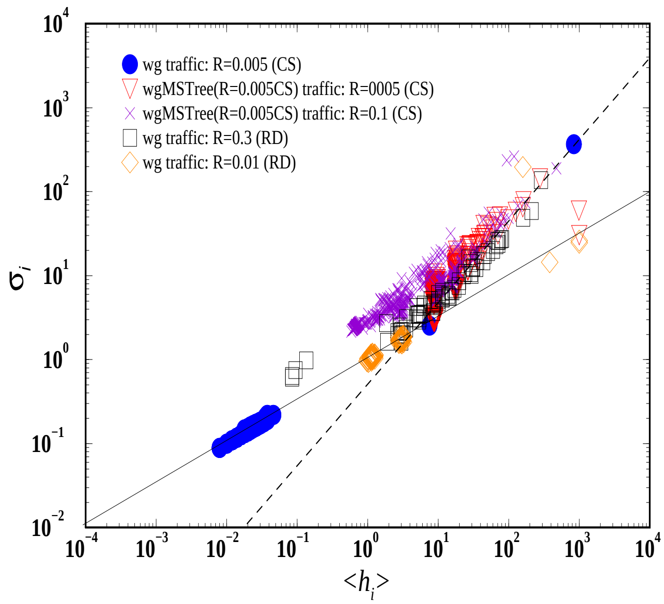

Fig. 2. Universal relations between dispersion σ_i and average occupation $< h_i >$ of nodes $i = 1, \cdots, N$ within time window $T_{WIN} = 1000$ time steps for traffic on the Web graph's giant component and on its maximum spanning tree at driving rate R and for specified search rule.

3.2. *Universality of traffic noise*

We further study the properties of traffic noise by monitoring in parallel the number of packets processed by each node within a fixed time window $T_{WIN} = 1000$ steps. (One time step consists of one parallel update of the whole network.) The number of packets registered within the specified time window fluctuates between different nodes and at the same node at different times. However, the ratio of the fluctuations to the number of processed packets shows some universal features, similar to real transport networks. In particular, the relation

$$\sigma_i \sim < h_i >_T^{\mu} \; ; \tag{3·1}$$

holds, where with μ takes two values $\mu = 1/2$ or $\mu = 1$ in different networks.[9] In Fig. 2 we show the standard deviation σ_i against the average occupation number $< h_i >$ of a node i, where the average is over the fixed time window T_{WIN}. Different curves are for different driving rates R and navigation rules, i.e., advanced nnn-search (AS) or random diffusion (RD). On each curve a point represents one node of the network. As the Fig. 2 shows, different nodes follow the behavior with the slopes $\mu = 1/2$ or $\mu = 1$, which are indicated by lines. For low-density traffic on the Web graph most of the nodes are in the class $\mu = 1/2$, whereas the hub node is on the line with $\mu = 1$. For higher packet density (larger R), however, more nodes experience larger traffic and higher fluctuations of the traffic and consequently align along the $\mu = 1$ curve. Hence the scaling exponent in the ralation (3·1) depends on the driving conditions (packet density), which exploit the network's heterogeneity in different ways. Qualitatively same behavior was found in the case of random diffusion (RD) but with different posting rates, where the stationary flow conditions are fulfilled.

For comparison, we also run the packet traffic along the maximum-spanning tree of the Web graph (results are also shown in Fig. 2). In this respect, the packet motion can be considered as subject of a constraint to only maximum-flow links on the original graph. We find that the constraint does not violate the universal relation (3·1), however, all nodes align along the $\mu = 1$ slope independently on driving rate.

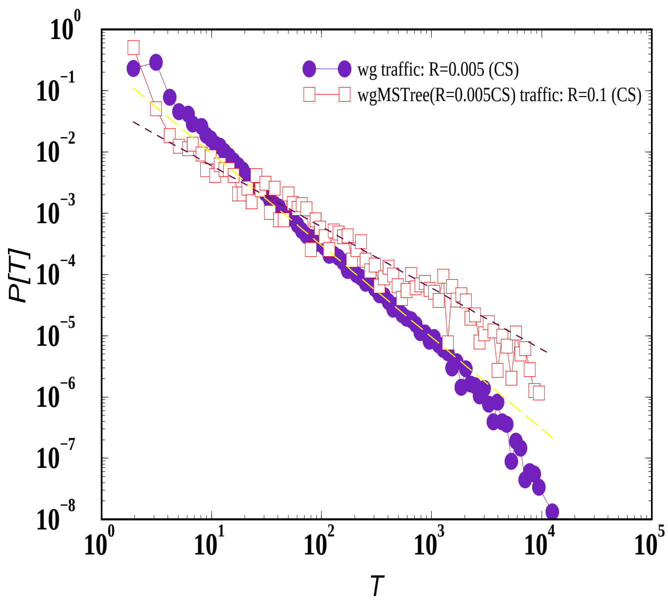

Fig. 3. Distribution of transit times T of packets on the Web graph (wg) and on its maximum spanning tree (wgMSTree) for fixed driving rate R as indicated and advanced local search (CS). Slopes $\tau_T = 3/2$ and $\tau_T = 1$ are indicated by broken lines.

3.3. *Transit times distribution*

As traffic characteristics on global (network) level we consider here the distribution of tranist times and correlations in node activity.[1),3)] Differences in the transport details on the cyclic Web graph and in traffic on its maximum spanning tree (MST) accumulate in the overall time that a packet spends on the way before it arrives to its destination address. In Fig. 3 we show the distributions of the transit times of packets for fixed rate R at the Web graph and its MST. For traffic on the original Web graph the distribution $P(T)$ has a power-law tail $P(T) \sim T^{-\tau_T}$, with the exponent $\tau_T = 1.5$ (within numerical error bars).[1] The observed deviation from the power-law at small times T is due to increased efficiency of the super-structure associated to the hub, when the the destination nodes are at the distance 2—3 nodes, which coincide with the searched depth in the advanced search rule (CS). We also show the results when the transport is constrainted along the maximum flow links corresponding to that rate on the Web graph (traffic on MST). As the Fig. 3 shows, the traffic along the MST intends to have increased probability of larger times compared to the original cyclic graph. This is mainly due to larger distances on the tree and also larger waiting times at the main hub when the traffic density increases. The slope of the curve $P(T)$ for the case of tree is close to $\tau_T = 1$. This ndicates another universality class of the transport processes constrainted on the graph's maximum spanning tree, compared to the underlying original graph with cycles.

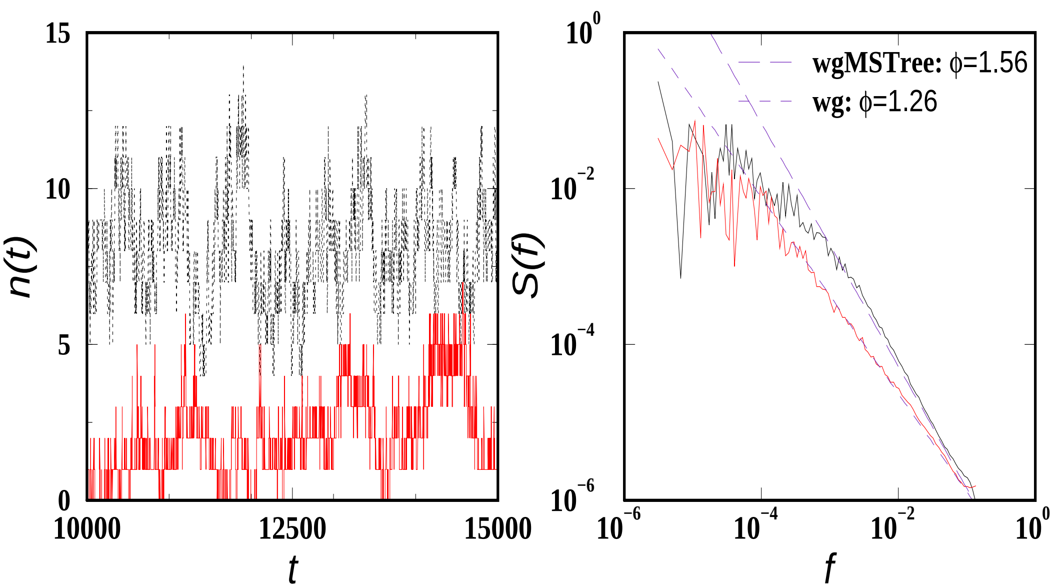

Fig. 4. Part of the time series of the numebr of active nodes $n(t)$ (left) and their power-spectra log-binned curves (right) for traffic on the Web graph (lower curve) and traffic on web-graph's maximum spanning tree (upper curve) for fixed posting rate $R = 5 \times 10^{-3}$.

3.4. *Noise correlations at network level*

More scaling features in the traffic noise can be found when the transport is looked at the global network level. In particular, the network structure supplemented with the navigation rule induces the correlations among the packet streams, a property which was found both in real network traffic[10] and in the traffic models.[1,3,11]

Here we consider only the correlations that arise in the number of simultaneously active nodes $n(t)$ on the network at time t. The time series $\{n(t)\}$ shows the *antiper-*

sistent correlations, as shown in Fig. 4. The exponents of the power-spectrum of the time signal$\{n(t)\}$ is characteristic for the networks structure. We find $\phi = 1.26$ for traffic on the Web graph and $\phi = 1.56$ for traffic on its maximum spanning tree for the driving rate R as indicated on the Fig. 4. More detailed analysis of the scaling properties of noise and the structure of the underlying graph and driving conditions for a class of scale-free graphs can be found in Ref.[?] As a rule, less correlations are found in less structured graph and larger traffic density.[12)]

3.5. *Evidences of stochastic ergodicity breaking in transport on networks*

Detailed analysis of traffic at individual node level reveals additional features of complex systems behavior that we address in the following. Transport of a packet on the network is a (guided) random walk process. Compared to a compact environment, network's topology makes severe constraints to the walker. Hence, by considering the number of hits of a walker to each node in the network one can infer details of the network structural properties, which can be used for exploring complex structures.[13)–16)]

From the dynamic point of view, the number of hits at a node depends not only on the topology but also on the time window in which it is measured. Therefore, it gives information about *how* the process goes in particular parts of the graph. In Fig. 5 we display the distributions of the occupation numbers of each node on the Web graph, collected within a fixed time window $T_{WIN} = 1000$ steps. The distribution shows a power-law tail, which is well pronounced at low traffic density R. As expected,[14)] in this limit the power $\tau_h = 2.28$ is roughly in agreement with the distribution of betweenness of nodes.[17)] Betweenness of a node is pure geometrical characteristic of the graph, which is defined as the number of shortest paths between pairs of nodes on the graph that go through that node.[2),17)]

At large driving rates, however, the distribution develops peaks at specific occupation numbers. In particular, the peak at highest occupation corresponding to $h = T_{WIN}$ appears, which indicates a constantly active hub node. Similarly, one can expect that other peaks that appear at lower occupation numbers correspond to groups of nodes which play a special role in specific parts of the graph or communities at large traffic density.

The broad distribution of the occupation numbers h_i recorded within the time window T_{WIN}, or equivalently of the occupation probabilities of nodes defined as $p_i \equiv h_i/T_{WIN}$, suggests that a dynamic ergodicity breaking occurs in the transport process on networks. Rrecently such ergodicity breaking in the case of a continuous-time random-walk has been discussed and classified in terms of occupation times in Ref.[18)] In our case, the inhomogeneity and sparseness of the underlying network induces spread of the frequency of the node occupation, as shown in Fig. 5. In low traffic density the ergodicity breaking appears where it can be related to the geometry of the network alone. At higher rates R , i.e., higher traffic density, additional effects occur due to nontrivial waiting times[3)] of packets in queues at particular nodes in the network. These waiting times are also related to the structure–dynamics interply.[1),3)] A more detailed analysis of the dynamic ergodicity breaking and role of different network topologies will be given elsewhere.[19)]

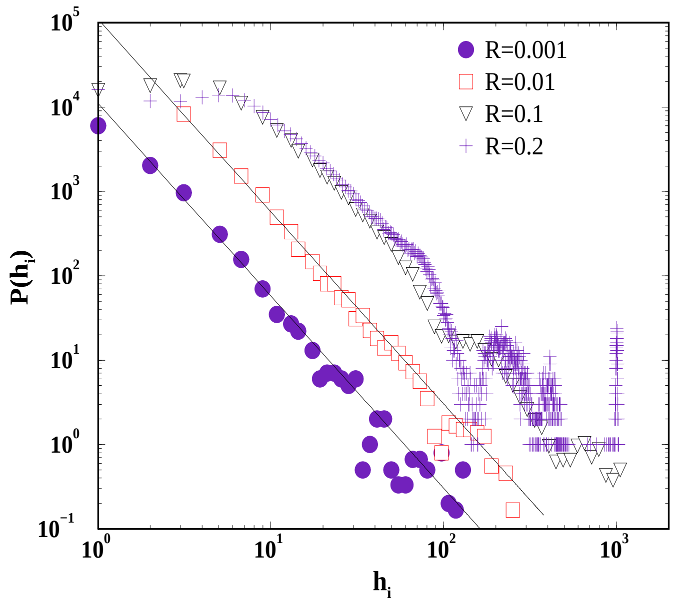

Fig. 5. Distribution of occupation numbers of nodes h_i occurring within a fixed time window $T_{WIN} = 10^3$ steps for different posting rates R. Solid lines slope $\tau_h = 2.28$ within error bars.

§4. Conclusions

We have demonstrated emergence of a functional topology in traffic of information packets on the cyclic scale-free graph (Web graph). Here we have shown results for traffic in the case of constant posting rates R. These topologies are quantitatively described through the statistical properties of the *network flow* and *traffic noise*.

Study of the network flow enables the dynamic differentiation between links and their role in the dynamic process. The systematic record of the traffic flow on the Web graph, as it is done here, represents an example where "weights" on the links are dynamically generated, in contrast to *ad hoc* implementation which is commonly used in study of weighted networks.[8),20)] The functional heterogeneity of the network is then built in the maximum spanning tree. We have shown that the dynamically generated weights result in a scale-free tree in the case of the traffic with the advanced nnn-search rule on the Web graph. Hence the spanning tree preserves the main property of the original graph, suggesting that the network topology is used efficiently by the traffic navigation.

Network's inhomogeneity is further represented by different roles that nodes play in the transport process. The multichannel noise analysis reveals universal noise features, which are expressed by Eq. (3·1). The scaling exponent μ can be related to the underlying graph structure and to traffic density and driving conditions. This is most strikingly demonstrated by comparing the traffic on the graph with the traffic on thereby generated maximum spanning tree. In the case of maximum spanning tree, the traffic can be viewed as traffic on the original graph with applied constraint, in which packets are systematically restricted to the available maximum-flow links.

This constraint results in noise properties which are typical for a tree structure. In particular, the scaling relation in Eq. (3·1) hold with $\mu = 1$ independently on the packet density. In addition, generally larger distances on the tree result in increased probability of longer transit times compared to the original cyclic graph. Therefore the transit time distribution on the maximum spanning tree of the Web graph appears in a new class of stochastic limit processes[21)] with the power $\tau_T \approx 1$, compared to $\tau_T \approx 3/2$ for the Web graph itself.

Finally, a detailed analysis of the occupation probabilities of nodes reveals that the stochastic ergodicity breaking occurs in the traffic on networks. The origin of a broad distribution of the occupation numbers of nodes is in the network topological inhomogeneity, as demonstrated for traffic at low posting rates. Additional effects are due to structurally induced interaction between packet streams, which lead to nontrivial waiting-time statistics at higher posting rates.

Acknowledgements

Supported by the Program P1-0044 of the Ministry of high education, science and technology (Slovenia). Partial support from the project SI-AT/04-05/01 sponsored by the MHEST (Slovenia) and the ÖAD (Austria) is also acknowledged.

References

1) B. Tadić and S. Thurner, Physica A **332** (2004), 566; Physica A **346** (2005), 183.
2) R.K. Ahuja, T. L. Magnanti, and J. B. Orlin, *Network Flows: Theory, Algorithms, and Applications*, Prentice Hall, (1993).
3) B. Tadić, S. Thurner, and G.J. Rodgers, Phys. Rev. E **69** (2004), 036102.
4) B. Tadić, Physica A **293** (2001), 273.
5) S. N. Dorogovtsev and J.F.F. Mendes, *Evolution of Networks: From Biology to the Internet and the WWW*, Oxford University Press, (2003).
6) *Pajek programe for Large Ntework Analysis* by V. Batagelj and A. Mrvar, http://vlado.fmf.uni-lj.si/pub/networks/pajek/doc/pajekman.htm .
7) B. Tadić, *Modeling Traffic of Information Packets on Graphs with Complex Topology*, Lecture Notes in Computer Science, Vol. 2657 Part I, Springer, Berlin (2003).
8) P.J. Macdonald, E. Almas, and A.-L. Barabasi, cond-mat/0405688.
9) M. Argolo de Menezes and A.-L. Barabasi, Phys. Rev. Lett. **92** (2004), 028701.
10) M. Takayashu, H. Takayasu, and T. Sato, Physica A **233** (1996), 824.
11) B. Tadić and G.J. Rodgers, Adv. Complex Systems **5** (2002), 445.
12) B. Tadić, Proceedings of the SFKM 2004, R. Žikić *et al.*, Eds., Institute of Physics, Belgrade (2004).
13) B. Tadić, Eur. Phys. J. B **23** (2002), 221.
14) B. Tadi; *Exploring Complex Graphs by Random Walks*, APS Conference Proceedings *Modeling of Complex Systems*, **661**, 24 (2003).
15) H. Zhou, PRE **67** (2003), 061901
16) J.D. Noh and H. Rieger, PRL **92** (2004), 118701.
17) K.I. Goh, E.S. Oh, H. Jeong, B. Kahng, and D. Kim, Proc. Natl. Acad. Sci. **99** (2002), 12583.
18) G. Bel and E. Barkai, PRL **94** (2005), 240602.
19) B. Tadić, S. Thurner, and G.J. Rodgers, in preparation.
20) A. Barrat, M. Barthelemy, and A. Vespignani, Phys. Rev. E **70** (2004), 066149.
21) W. Whitt, *Stochastic-Process Limits*, Springer, New York (2001)