Local and and non-local information in decentralized routing strategies

Giovanni Petri,[1,*] Dirk Helbing,[2] Henrik J. Jensen,[3] and John W. Polak[4]

[1]*ISI Foundation, via Alassio 11c, 56010 Torino, Italy*
[2]*Chair of Sociology, in particular Modelling and Simulation, ETH Zürich, Clausiusstrasse 50, 8092, Zürich, Switzerland*
[3]*Complexity and Network Programme, and Dept. of Mathematics, Imperial College London, South Kensington Campus, London SW7 2AZ, UK*
[4] *Centre for Transport Studies, Dept. Civil and Environ. Engineering, Imperial College London, South Kensington Campus, London SW7 2AZ, UK*

We compare three routing models incorporating different information about the network state, showing that local congestion awareness improves the resilience to congestion emergence more than a decentralized dissemination mechanism, which uses dynamically sourced and disseminated information. Before the emergence of congestion, local information is optimal, because it induces the smallest deviation from the shortest path, allowing at the same time to avoid localized congested. When congestion becomes unavoidable, the analysis of the fundamental diagram showed that dynamically sourced information is able to maintain the traffic flow more fluid for a period. With regard to congestion emergence, we found a transition from instability to stability, controlled by the memory parameter. Interestingly, this transition corresponds to a transition for the density field from a very heterogeneous configuration to a temporally and spatially uniform one, that approximates well the Wardrop equilibrium for the system. Finally, our results are in agreement with previous work on the role of traffic spatial and temporal heterogeneity and suggest a two-fold strategy for dealing with congested networks.

PACS numbers: 89.75.-k, 89.40.-Bb, 89.75.-Fb

I. INTRODUCTION

How the spreading of information influences traffic on a network has become the focus of great attention as it touches on a number of organisational, economical and also social issues [1]. Studies on network routing problems, mostly focused on applications to information networks [2, 3], have shown different choices of routing strategy result in different statistical properties of the network traffic, e.g. changing the nature of the transition from a free to a congested network [4–11]. In the specific context of human agents and road traffic, a number of studies are focusing on enabling the exchange of information among agents in order to improve route planning and thus network performances [12–14]. Some of these [20, 21] successfully showed the effectiveness of such approaches in highway traffic, for example in reducing the likelihood of over-critical density perturbations triggering jam formation [20, 22]. At the network level however it is harder to predict what is the actual effect of information dissemination. A few studies proposed optimal routing models for traffic networks [23, 24]. However, these usually require centralized information elaboration and dissemination, which is hard or impossible to fathom in the case of large networks.

Local decentralized approaches are able to solve this problem. Meloni and Gómez-Gardeñes [25] in a particle-hopping model showed that is possible to obtain a global optimisation through "empathic" routing strategy where nodes decide whether to accept particles depending on their own state and that of their first neighbors. Studying a more realistic model with explicit navigation between origins and destinations, Scellato et al. [15] showed that at high traffic densities a simple local congestion-aware rule is able to outperform a navigation model where global (complete) information is used. It is not yet clear however whether any further improvement can be obtained by adopting an intermediate information horizon of the routing strategy, or, conversely, whether too much disseminating information is globally detrimental.

In this article, we take approach the problem of the optimal information scale. In particular, we compare three navigation mechanisms that exemplify different modes of information diffusion and regimes: no information (shortest-path routing), local congestion information (local routing) and, finally, dynamical decentralized information, when the information scale emerges dynamically.

The first can be considered as the null model. The second, routing with local congestion awareness, is akin to those already studied in the literature [15–17], in the context of Internet traffic, particle-hopping models and contact processes [18, 19].

The third model explores the gap between purely local and purely global routing mechanisms. Taking inspiration by navigation methods of ants, we exploit the cooperation between agents and the network infrastructure. In the model, agents are at the same time users, producers and carriers of traffic information spreading. Users, because they can request information to choose their route. Producers and carriers, because they themselves collect and carry information around during their trip through the network.

An example of how a similar cooperative multi-agent mechanism could be implemented in practice was proposed by Adler [26]. Here, however, we focus on the effects, rather than on the details of the realization. In particular, our approach al-

*Electronic address: giovanni.petri@isi.it

Electronic copy available at: http://ssrn.com/abstract=2725066

lows us to study a range of different length scales for information dissemination, through the tuning of a forgetting, or information removal, parameter (similar to the evaporation of ants trail). For fast removal of information, the model tends to local information routing; for slow removal, such information comes effectively into play modifying the environment and therefore the behavior of the agents.

II. ROUTING MODELS

In order to study the effects of different information regimes, it is necessary to define a navigation rule for agents. Consider a agent on route to node d who reaches junction l. To proceed toward its destination, it chooses one of l's outgoing links (l, m) with probability $p^d_{lm} \propto f(\tilde{T}_{ld;m}/T^0_{ld})$, where $\tilde{T}_{ld;m}$ is an estimate of the actual travel time from l to d passing through neighbor m and T^0_{ld} is the reference (free flow) time along the shortest path from l to d. Different information dissemination strategies change the calculation of $\tilde{T}_{ld;m}$, providing different routing schemes.

In general, we can split the $\tilde{T}_{ld;m}$ into $T^{loc}_{lm} + T^{dyn}_{md}$, highlighting the local travel time (the first term) and the travel time information relative to the rest of the journey (second term). In terms of these quantities we can write the probability for a agent at node l to choose the outgoing link (l, k) on her route to d as:

$$p^d_{lm} = \frac{\exp(-(T^{loc}_{lm} + T^{dyn}_{md})/T^0_{ld}}{\sum_{k\in\Gamma_l} \exp(-(T^{loc}_{lk} + T^{inf}_{kd})/T^0_{ld}} \qquad (1)$$

where Γ_l is the neighborhood of l. The $\{T^0\}$ represent the reference travel times along shortest path routes.

a. No information The case of no information coincides with the shortest path routing and thus trivially we have

$$T^{loc}_{lm} \rightarrow T^0_{lm} \qquad (2)$$

$$T^{dyn}_{md} \rightarrow T^0_{md}. \qquad (3)$$

b. Local information In the case of local congestion routing, akin to [15], only the congestion on the next link is considered, thus $T^{dyn}_{md} \rightarrow T^0_{md}$. The travel time along link (l, m) at time t will be proportional to the queue length $Q_{lm}(t)$. For a link with maximal outflow S, we use

$$T^{loc}_{lm}(t) = aQ_{lm}(t)/S, \qquad (4)$$

where a is the average speed through the network relating network distances to the reference travel times (see II A).

c. Dynamical information In the case of dynamical information, the travel time information results from a superposition of the local information and long-range information, that is carried and collected by the agents themselves. Therefore, for T^{loc}_{ld} we adopt Eq. 4. We still need a prescription for T^{dyn}_{md}. Before giving the functional form for T^{dyn}_{md}, it is important to explain how dynamical information is collected and disseminated. The first step in this direction is promoting agents and junctions (network nodes) to interacting agents. Decentralized information dissemination requires travel times to be sourced by the agents along their routes and exchanged with the nodes. In particular, consider two agents, α and β respectively on route from o to d and from d to o. As α proceeds, she collects information about the congestion, hence the cumulative travel time T^{cum}_o, along the links going in the opposite direction (towards o). In this way, when the two agents meet going in opposite directions, β can use the T^{cum}_o gathered by α to decide whether to continue on the same route or divert along another path to o. On a network however there are many possible routes to the same node and therefore we need entities, in our case the nodes, to store, negotiate and share information among agents.

More specifically, each agent α is characterized by:

1. Its origin junction o_α and the time t^α_0 when it entered the system;
2. its destination junction d_α;
3. the cumulative travel time $T^{cum}_{o_\alpha}$ to o_α, gathered as it moves on the network;

The cumulative time $T^{cum}_{o_\alpha}$ is updated as follow: suppose that, on its route to d, a agent follows the link (i, j) reaching j at time t, it updates $T^{cum}_{o_\alpha} \rightarrow T^{cum}_{o_\alpha} + T^{loc}_{ij}(t)$. We stress that each agent carries just the total travel time from its current position back to its origin o along the links it traversed (in opposite direction).

Nodes collect the travel time information carried by agents passing through. At time t, a junction i stores for each node $j \neq i$ the estimated dynamical travel time to j, $T^{inf}_{id}(t)$. This time is a combination of the most recent information received by agents, i.e. the T^{cum}_j of the most recent (largest t_0) agent from j passing through i, and of the reference time T^0_{ij}:

$$T^{inf}_{ij} = (1 - e^{-\lambda(t-t^{upd})})T^0_{ij} + e^{-\lambda(t-t^{upd})}T^{cum}_j(t_o) \quad (5)$$

where t^{upd} is the time at which node i last update its information about j.

The parameter λ controls the memory (or information removal rate) of the system. Small values of λ correspond to retaining for long times the travel time information carried by the agents, while in the limit $\lambda \ggg 1$ Eq. 5 reduces to local information only $T^{inf}_{ij} \sim T^0_{ij}$. Through this setup it becomes possible to study the behavior of the system under a continuous tuning of the memory parameter.

Note that in the navigation models analyzed agents choose at each node which direction to follow, based on the topology of the network and the available information. At variance with the traffic engineering approach of pre-calculating and storing routes, the navigation rules do not lay out routes that agents follow. Each decision is in principle independent from the previous ones. In practice however the navigation rules tend to push agents along minimal travel time routes, similarly to particles following an attractive potential to their own destination

Electronic copy available at: http://ssrn.com/abstract=2725066

point. Routes appear thus as high probability paths, resembling geodesic trajectories in such potential. Information can be then thought as a modulation of this attraction toward the destination.

A. Simulations

Simulations were performed on a square lattice with $n = 100$ nodes and periodic boundary conditions. Each node is a junction and has 4 outgoing and 4 incoming links. This topology because it is closer to the topology of urban networks (e.g. Manhattan), as opposed to scale-free or random network topologies. Each directed link (i, j) has a maximal allowed queue (its buffer) B_{ij}. Nodes service one incoming link per timestep, in order to represent fixed cycle traffic lights. The first link to be served is randomly selected, then the links are served sequentially. When served, a link (i, j) has a maximal outflow S_{ij}, that is up to S_{ij} agents can attempt to continue their trip on another link. However, if the link chosen by the agent is already full, the whole queue on (i, j) is stopped. This mechanism is the main ingredient of congestion spreading. agents are introduced with rate N_c and removed from the system when they reach their destinations. Origin and destination nodes are chosen randomly. All runs last up to $T_{max} = 2.5 \cdot 10^4$ time steps, unless the network population reaches the capacity before. In the simulations we used $B_{ij} = 30$ and $S_{ij} = 15\ \forall i, j$, which that a link at capacity will not be able to completely empty in one timestep. For the capacity parameters above, in a network with vanishing agent density, one finds that the reference times are well approximated by $T^0_{lij} \simeq 2.2 d_{ij}$, where d_{ij} is the topological distance between junctions i and j and therefore $a = 2.2$.

III. RESULTS AND DISCUSSION

A. Dynamical properties

In Figure 1 we compare the performances of the three models under increasing network inflow. The shortest path routing uses no information and produces expectably the shortest times to congestion. Surprisingly, however, the dynamical information strategy does not perform much better. In fact, we find that using local information alone (the local congestion routing) it is possible to push the system to much larger inflows ($N_c \sim 130$) without congestion. The value of λ used for Fig. 1 is 0.1. It is reasonable to expect that tuning λ one might obtain different results. Indeed, this is the case, but increasing N_c also increases the minimal value of $\lambda_c(N_c)$ needed for the system to be stable.

Figure 2 shows a section of the phase space of the normalized time to congestion, T_{cong}/T_{max} for a range of values for N_c and λ. To avoid the emergence of congestion, one needs to increase λ significantly. This is reasonable: in a system were new agents are introduced at a quicker pace, the network state will change more rapidly and therefore old information needs to be removed accordingly. However, it also implies that for larger inflows it is advantageous at the systemic level to shorten the time over which dynamical information actually plays a role, thus converging to the simpler local information strategy.

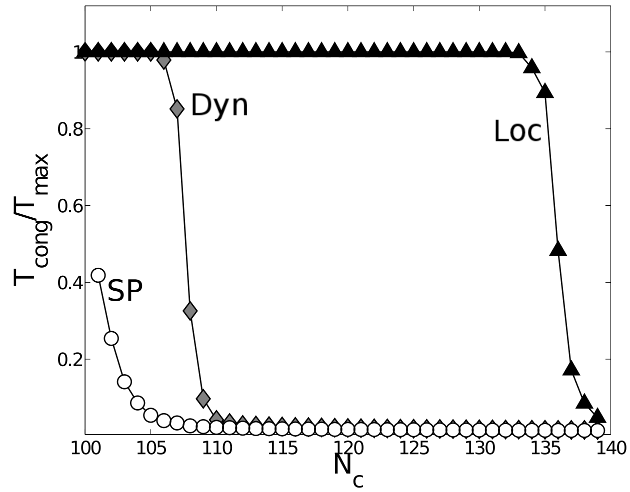

FIG. 1: The time to congestion T_{cong}/T_{max} versus the network inflow N_c for the three cases of shortest path routing (empty circles), local congestion-aware routing (black triangles) and dynamical information routing (gray diamonds) for $\lambda = 0.1$. It can be easily seen that dynamical information is able to outperform shortest path routing for the range shown. However, using only local information provides the best results in terms of the time to the emergence of congestion, significantly increasing the maximum N_c that the network could sustain.

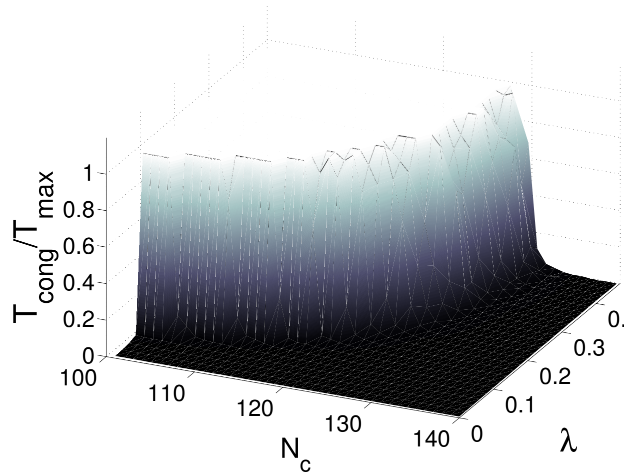

FIG. 2: The time to congestion in the dynamical information routing model over a range of N_c and λ. The system is very prone to develop congestion for low values of λ even when the forcing is small ($N_c \sim 110$). Indeed, increasing λ allows to enter the region of stability where no network-level congestion emerges. The larger the value of λ the faster long-range traffic information is neglected in comparison to local one.

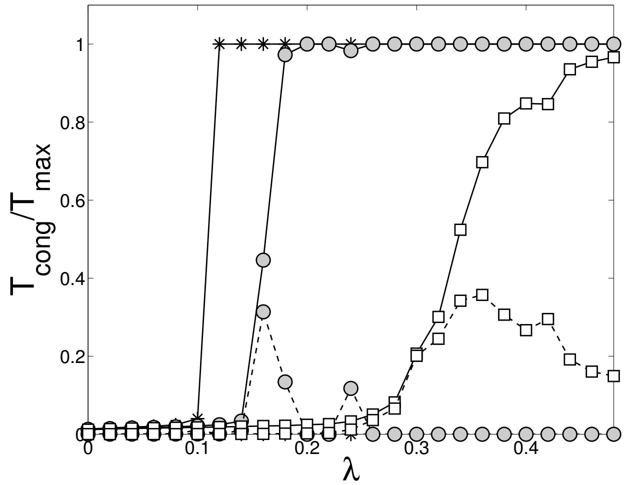

FIG. 3: The normalized time to congestion T_{cong}/T_{max} as a function of λ for different values of the forcing rate N_c= 110 (stars), 120 (circles), 130 (squares). The full lines indicate the expectation values, the dashed lines the fluctuations (over different realizations of the system).

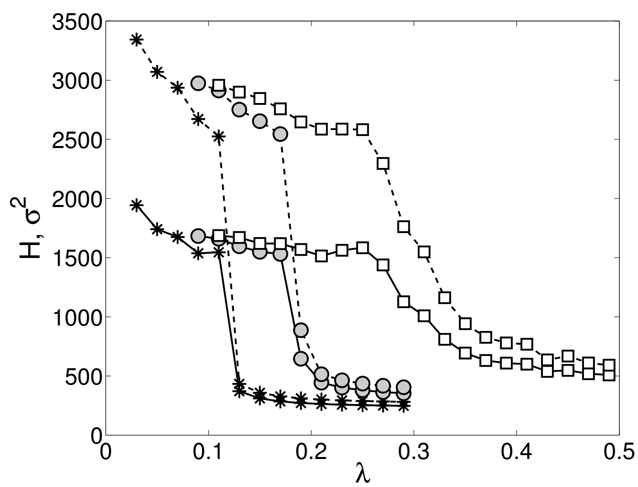

FIG. 4: The quantities H (full lines) and σ (dashed lines) as functions of for different values of the forcing rate N_c= 110 (stars), 120 (circles), 130 (squares). The values shown were averaged over 100 simulations for each data point.

Therefore, in terms of the maximal N_c the system can sustain before jamming, the local information outperforms both the dynamical information strategy and (predictably) the shortest path routing.

This suggests that long-range information can be detrimental to the network state. Different dissemination strategies of travel time information significantly alter the probability of emergence of congestion in a network. However, it is not clear yet which are the dynamical effects, induced by information, that create instabilities in the system. Previous work [27, 28] linked traffic heterogeneity with inefficient use of network capacity, resulting in increased instability and a higher probability for congestion. Intuitively, the most heterogeneous states at a given traffic density are the ones where most of the traffic is concentrated on few links, while the others are scarcely used. The full links therefore initiate small jams which impair the outflow from links upstream and, through spill-over, congestion spreads. On the contrary (and according to Wardrop's principle [29]), a uniform density landscape is less likely to produce the fluctuations needed to break down the flow. Hence, if for some values of λ the system is indeed destabilized, there should be traces in the statistical properties of the queue distributions. To test this hypothesis we consider the following quantities [28]:

$$H = \overline{\langle \rho(t) \rangle^2} - \overline{\langle \rho(t) \rangle}^2 \tag{6}$$

$$\sigma^2 = \overline{\langle \rho(t)^2 \rangle} - \overline{\langle \rho(t) \rangle}^2 \tag{7}$$

where $\langle \cdots \rangle$ stand for time averages, the $\overline{\cdots}$ for spatial averages (over all links) and ρ is defined on each link as the fraction of capacity used ($Q_{ij}(T)/B_{ij}$ for (i, j)). H accounts for the spatial fluctuations across links of the average population: large values characterize heterogeneous configurations. The quantity σ^2 measures both the spatial and temporal fluctuations: for example, a network with high traffic flow volatility will have a significantly larger σ^2 than a system with a very constant flow. There is a direct connection between the couple H and σ^2 and the stability/instability in the time to congestion. In Figures 3 and 4 we report the network properties for three values of $N_c = 110, 120, 130$. First, one see that the transition from small to large T_{cong}/T_{max} is accompanied by small fluctuations for $N_c = 110$, while in the other two cases the fluctuations are significant. This is reminding of a different order of stability/instability phase transition, but is actually an effect of the different angle at which a slice of fixed N_c intersects the surface in Fig. 2.
More interestingly, the transition points $\lambda_c(N_c)$ are the same for T_{cong}/T_{max} and the order measures H and σ^2. For $\lambda \geq \lambda_c$, both quantities decrease rapidly. This means that the average density values on the different links become progressively more and more uniform, producing a flatter landscape. The same effect is apparent in the temporal fluctuations ($\sigma^2 \to 0$), implying that the volatility of the single link densities is reduced: the system tends to a uniform state both in space and time, as non-local information is discarded more and more quickly. Local congestion awareness emerges then as the optimal routing because of it produces the smallest fluctuations.

B. Fundamental Diagram

The results presented up to now have shown a counter-intuitive, if not paradoxical, feature of decentralized information diffusion. Non-local travel time information is likely to reduce the performances of a traffic network, on both the user and system level. The result is surprising as one would naively

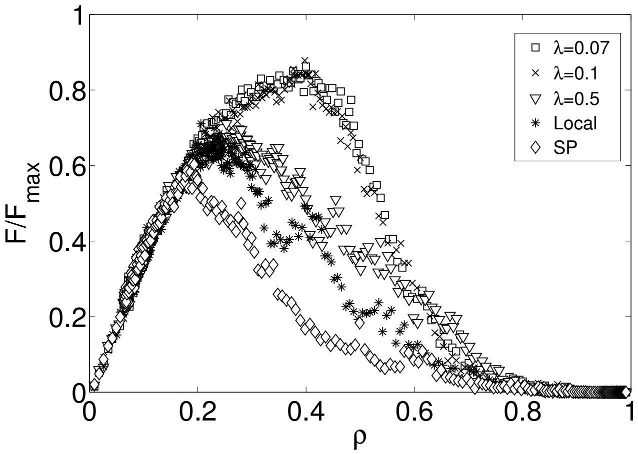

FIG. 5: Fundamental Diagram for the cases of dynamical aging information (blue squares for λ=0.07, red dots for λ=0.1 and gray upside down triangles for λ=0.5), local information only (blue stars), shortest path routing (light grey diamonds).

expect that conveying information about the congestion state upstream would improve the global network flow. It seems odd indeed that the only outcome of more information would be more congestion. The quantities we studied until now give a dynamical picture of the evolution of network traffic. However, they do not describe appropriately how the flow-density relation is modified by the different information mechanisms. In Figure 5 we plot the fundamental diagram (FD) [30, 31] for the dynamical travel time information routing with different values of the parameter λ, for shortest-path navigation and local information navigation. The plots shown refer to a driving rate $N_c = 150$, for which none of the navigation strategies can avoid the emergence of network-wide congestion. This is done in order to probe the effects of the various routing mechanisms during congestion. The density and the flow are normalized by the network capacity and maximal flow (per time step). Interestingly, the FD curve for the local information shows a peak at a traffic density $\rho^{loc}_{peak} \simeq 0.2$, smaller than the one of the dynamical information navigation, $\rho^{dyn}_{peak} \simeq 0.4$ for $\lambda \simeq 0.07$ and 0.1. Also, in the same range of λ, the maximum flow under the dynamical information mechanisms is the largest. As expected, for any further increase in λ, the system approaches the behavior of the local information navigation (see for the curve corresponding to $\lambda = 0.5$ in Fig. 5). Despite being less resilient to larger N_c, the dynamical information mechanism expands the free flow region in the macroscopic FD once congestion has emerged. This happens because, through the long-range information dissemination, incoming agents can steer away, at least for a time, from the growing congestion front. The mechanism therefore does not manage to relieve congestion completely, but can delay the network-wide onset for a time by automatically rerouting agents away from the congestion front.

IV. CONCLUSIONS

We compared three information regimes showing that local congestion awareness improves the resilience to congestion emergence more than a decentralized dissemination mechanism in which information is dynamically sourced and disseminated. The former may be thought as causing the smallest deviation from the shortest path, while at the same time allowing agents to avoid localized congested regions when they are directly encountered. On the other hand, for high network inflows (or equivalently large densities), when congestion becomes unavoidable, the dynamical information mechanisms allows to keep the traffic flow more fluid, as shown by the analysis of the fundamental diagram. With regard to congestion emergence, we found a transition from instability to stability, controlled by the memory parameter. Interestingly, this transition corresponds to a transition for the density field from a very heterogeneous configuration to a temporally and spatially uniform one, that approximates well the Wardrop equilibrium for the system. These results are in agreement with previous work: different information regimes generate different spatial patterns for the flow, which in turn determine the network capacity. Hence, the type of information dissemination can be thought as a property of the network itself, rather than of the agents. For real-world applications, our results suggest a twofold policy regarding information dissemination: for low densities and no congestion, local information should be used (e.g. travel time along the next link after a decision point); when heavy congestion appears, dynamical information should be used in order to avoid (or at least delay as much as possible) the gridlock.

[1] R. Pastor-Satorras and A. Vespignani, *Evolution and Structure of the Internet: A Statistical Physics Approach*, (Cambridge University Press, Cambridge, 2004).
[2] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez and D.-U. Hwang, Phys. Rep. **424**, 175 (2006).
[3] S.N. Dorogovtsev, A.V. Goltsev, and J.F.F. Mendes, Rev Mod. Phys. **80**, 1275 (2008).
[4] M. Argollo de Menezes, and A. L. Barabási, Phys. Rev. Lett. **92**, 028701 (2004).
[5] J. Duch, and A. Arenas, Phys. Rev. Lett. **96**, 218702 (2006).
[6] S. Meloni, J. Gómez-Gardeñes, V. Latora, and Y. Moreno, Phys. Rev. Lett. **100**, 208701 (2008).
[7] T. Ohira, and R. Sawatari, Phys. Rev. E **58**, 193 (1998).
[8] R. Guimerá, A. Arenas, A. Díaz-Guilera, and F. Giralt, Phys. Rev. E **66**, 026704 (2002).
[9] R. Guimerá, A. Díaz-Guilera, F. Vega-Redondo, A. Cabrales, and A. Arenas, Phys. Rev. Lett. **89**, 248701 (2002).
[10] R. V. Solé, and S. Valverde, Physica A **289**, 595 (2001).
[11] S. Valverde and R. V. Solé, Physica A **312**, 636 (2002).
[12] M.W. Ng, S. Travis Waller, Trans. Res. Part C **18**(3), 393–407, (2010).
[13] D. Teodorovic, Trans. Res. Part C, **16**(6), 651–667, (2008).

[14] X. Yang, W. Recker, Trans. Res. C **13**(5-6), 370–390, (2005).
[15] S. Scellato, L. Fortuna, M. Frasca, J. Gmez-Gardees, V. Latora, Eur. Phys. J. B **73**(2),303–308, (2010).
[16] B. Tadi, S. Thurner, Phys. A **346**,183-190, (2004).
[17] Z. Hog-Li, G. Yan-Duong, Z. Chen-Ping, M. Mitrovic, B. Tadić, Proc. 16th Intern. Conf. Digital Signal Proc.,1–8. (2009).
[18] D. De Martino, L. Dall'Asta, G. Bianconi, and M. Marsili, Phys. Rev. E **79**, 015101(R) (2009).
[19] D. De Martino, L. Dall'Asta, G. Bianconi, and M. Marsili, J. Stat. Mech. P080232 (2009).
[20] A. Kesting, M. Treiber, M. Schnhof,D. Helbing, Trans. Res. Part C **16**(6), 668–683, (2008).
[21] M. Schnhof, M. Treiber, A. Kesting, D. Helbing, Trans. Res. Part Board, **1999**, 3–12, (2007).
[22] S. Lämmer, D. Helbing, J. Stat. Mech. P04019, (2008).
[23] L. Engelson, Trans. Res. Part C **11**(2), 161183 (2003).
[24] S. Kim, M.E. Lewis, C.C. White III. IEEE Trans. on Int. Transp. Syst. **6**(2), 178–188, (2005).
[25] Meloni, S., Gómez-Gardeñes, J., *Phys. Rev. E* **82**(5), 056105. (2010)
[26] J.L. Adler, V.J. Blue, Trans. Res. Part C , **10**(5-6), 433–454, (2002).
[27] A. Mazloumian, N. Geroliminis, D.Helbing, Phil. Trans. R. Soc. A**368**:4627–4647, (2010).
[28] A. De Martino, M. Marsili, R. Mulet, Europhys. Lett., **65**(2), 283-289, (2004).
[29] L.C. Davis, Physica A, **388**, 4459–4474, (2009).
[30] N. Gerolimins, J. Sun, Trans. Res. Part B **45**(3), 605–617 (2011).
[31] C.F. Daganzo, Trans. Res. Part B **41**, 49–62, (2007).