NETWORKS AND HETEROGENEOUS MEDIA doi:10.3934/nhm.2012.7.463

Volume **7**, Number **3**, September **2012** pp. **463–471**

IDENTIFYING CRITICAL TRAFFIC JAM AREAS WITH NODE CENTRALITIES INTERFERENCE AND ROBUSTNESS

Giovanni Scardoni

University of Verona, Center for BioMedical computing
Verona, Italy

Carlo Laudanna

University of Verona, Center for BioMedical computing, Department of Pathology
Verona, Italy

Abstract. We introduce the notions of *centrality interference* and *centrality robustness*, as measures of variation of centrality values when the structure of a network is modified by removing or adding individual nodes from/to a network. Centrality analysis allows categorizing nodes according to their topological relevance in a network. Thus, *centrality interference* analysis allows understanding which parts of a network are mostly influenced by a node and, conversely, *centrality robustness* allows quantifying the functional dependency of a node from other nodes in the network. We examine the theoretical significance of these measures and apply them to classify nodes in a road network to predict the effects on the traffic jam due to variations in the structure of the network. In these case the interference analysis allows to predict which are the distinct regions of the network affected by the function of different nodes. Such notions, when applied to a variety of different contexts, opens new perspectives in network analysis since they allow predicting the effects of local network modifications on single node as well as global network functionality.

1. **Introduction.** Study of complex networks currently spans several disciplines, including biology, pharmacology, economy, social science, computer science and physics [5]. One of the major goals of modern network science is the quantitative characterization of network structure and functionality with the purpose of inferring emergent properties of complex systems, abstracted as networks and represented as graphs [4]. Notably, since network structure always affects function [16], the topological analysis approach allows understanding networks functionality through the analysis of their specific architecture. For instance, the topological structure of the road network affects critical traffic jam areas, the topology of social networks affects the spread of information and disease, and the topology of power grids affects the robustness and stability of energy distribution. Remarkable results have been reached in this field and, even if far from being complete, several key notions have been introduced. These unifying principles underlie the topology of networks belonging to different fields of science [2],[9],[12],[11],[17],[10]. Notably, network analysis mainly focuses on global network properties and on their global modifications[3][8][1][6][15]. In this context, network centralities, such as

2000 *Mathematics Subject Classification.* 05C85, 90C35, 90B10, 90B06.

Key words and phrases. Graph theory, graph algorithms, complex networks, transportation networks, network centralities.

degree, eccentricity, closeness, betweenness, stress, centroid and radiality [14] are topological parameters allowing understanding the importance of single nodes in a network. Here, we introduce the notion of *centrality interference* and *robustness*, as measurements of changes in the local topological structure of the network as a consequence of single nodes removal or addition, in order to quantify the influence of single nodes in different parts of the network. Our approach allows addressing the following question: "if we remove or add one node in the network, how do other nodes modify their functionality because of this removal?". In some cases, such as in social and financial networks, the structure of the network is naturally modified over time; in other cases this can be due to specific network changes: power grid failures, traffic jam or work in progress in a road network, temporary closure of an airport in an airline network and so on. In a biological network one or more nodes (genes, proteins, metabolites) are possibly removed from the network because of gene deletion, pharmacological treatment or protein degradation. Understanding the topological consequences of such changes in the network means to understand how the network functionally rearranges. For instance in the case of a drug treatment, we can potentially predict side effects of the drug by looking to the topological properties of nodes in a drug-treated network, meaning with that a network in which a drug-targeted node (protein) was removed. Similarly, we can understand new critical traffic points in a road or airline network after a modification of its structure. Notably, our perspective concerns node-by-node modifications: a single node modification can be irrelevant to the overall organization of the network (for instance its scale-free structure), but can profoundly modify the properties of one or more nodes in different regions of the network, thus changing, for instance, the network modular structure. Since centralities are single-node properties, the effects of single node alterations can be calculated by analyzing modifications of centralities values due to single node alterations. As the centrality value of a node is strictly dependent on the network structure and on the properties of other nodes in the network, if we add or remove a node in the network the consequences of this modification on the network structures are reflected on the centrality values of all the other nodes. Such a situation, similarly to the case of interference for computer programs [7] can be analyzed introducing the notion of *centralities interference* and *robustness*. We first introduce the *betweenness interference* and then we extend the notion to other centrality measures. All definitions consider connected networks (i.e. networks where each node is reachable from all the others) and that remain connected also after nodes deletion. This hypothesis is in agreement with results in attack tolerance for scale-free networks [1].

2. **Betweenness Interference.** We consider a network as a graph $G = (N, E)$ where N is the set of nodes and E is the set of edges. Betweenness of node n is defined as $Btw(G, n) = \sum_{s \neq n \in N} \sum_{t \neq n \in N} \frac{\sigma_{st}(n)}{\sigma_{st}}$ where σ_{st} is the number of shortest paths between s and t and $\sigma_{st}(n)$ is the number of shortest paths between s and t passing through the vertex n. We consider the relative value of betweenness normalizing it as $relBtw(G, n) = \frac{Btw(G,n)}{\sum_{j \in N} Btw(G,j)}$ in order to have the fraction of betweenness of each node with respect to the rest of the network. Consider the example in figure 1a. If we remove node k from the network, node b becomes the only node connecting a to all the other nodes in the network (fig. 1b), so its betweenness value will increase. This is a case of *betweenness interference* since removing node k from the network "interferes" with the betweenness value of node b and can be

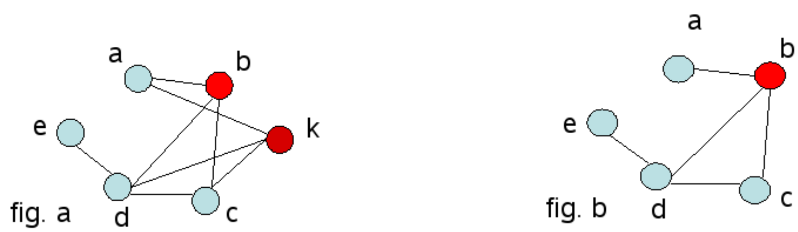

FIGURE 1. **a**. Node k and b are in the shortest paths from node a to the other nodes. **b**. Node k have been removed. Node b is now essential for connecting node a to the rest of the network: it is the only node in the shortest paths connecting a to the other nodes: node b betweenness increases.

measured as follow. $G_{|i}$ is the network obtained from G removing node i and all its edges from the network. The *betweenness interference* of node i with respect to node n in the network G is $Int_{Btw}(i,n,G) = relBtw(G,n) - relBtw(G_{|i},n)$. The measure shows which fraction of betweenness value a node loses or gains with respect to the rest of the network when the node i is removed from the network. The *interference* value can be positive or negative. If it is negative, it means that the role of node n in the network is higher when the node i is not present in the network. So we can say that node i has *negative interference* on node n, in the sense that the presence of node i in the network is "negative" for the node n to play a "central role" in the network. If the interference value is positive, it means that betweenness value of node n is higher if node i has been added to the network. In this case we say that i has *positive interference* on node n, in the sense that the presence of node i is "positive" for node n to play a "central role" in the network. The meaning of negative and positive interference strictly depend on the kind of network they are applied.

Note. Even if interference is calculated removing a node from the graph, it is a measure of the influence of this node with respect to the rest of the network. Besides it can also be used to model some frequent situations where nodes are added or removed to/from a network. In these cases adding a node means to add a node whose interactions are known. As example, adding a protein to a protein-protein interactions network we exactly know its interactions with other proteins (the new edges to add to the graph).

3. **Centralities Interference definitions.** The notion of interference can be easily extended to other centrality values and other interference based measures as *modulus interference* ($ModInt_{Btw}(i,n,G) = |Btw(G,n) - Btw(G_{|i},n)|$) *absolute interference* ($AbsInt_{Btw}(i,n,G) = Btw(G,n) - Btw(G_{|i},n)$) can be used to enrich the analysis. Finally, a further step for a complete analysis of interference is to quantify the interference of a single node with respect to the entire network. The question is: How node i is important for the functionality of the entire network? A node can interfere with high value with respect to few nodes and can have low interference value with respect to many others. Alternatively one node can interfere with significant values with respect to the most of the nodes in the network.

In the second case the node can have importance for the entire network functionality and not only for one or few nodes. In order to quantify the interference with respect to the entire network we introduce the $max\ interference$ value and the $global\ interference$ value. The $betweenness\ max\ interference$ value of node i is defined as $maxInt_{Btw}(i,G) = \max_{n \in N_{|i}} \{Int_{Btw}(i,n,G)\}$. If high at least one node is consitently affected by node i. The $betweenness\ global\ interference$ value of node i is $Int_{Btw}(i,G) = \sum_{n \in N_{|i}} (Int_{Btw}(i,n,G))$. If high the nodes interferes with high values with respect to the most of nodes in the network. In order to compare different networks this two values can be normalized dividing them by $|N|-1$ where $|N|$ is the number of nodes of the network.

4. **Centralities Robustness, Dependence and Competition value.** We approach now the reverse problem of interference: we know that a node has a central role in the network and we'd like to know if its functionality can be affected by other nodes and how much. The question is, conversely to interference: "which are the nodes affecting node n?". To answer to this question we introduce the notion of robustness, competition and dependence value of a node. The betweenness robustness of node n is defined as $Rob_Btw(n,G) = \frac{1}{\max_{i \in N_{|n}} \{|Int_{Btw}(i,n,G)|\}}$. Robustness depends on the maximum interference value that can affect the betweenness value of the node. If it is low, the node can be easily "attacked" by removing or adding particular nodes. If it is high, the node is "robust", i.e. there is no node removal or adding that can affect its betweenness value and consequently its functionality. Notice that we consider the modulus value of interference. Similarly to $interference$, positive and negative *robustness* can be defined but it is more intuitive to consider their reciprocal values respectively *dependence* and *competition* values. *dependence value* is $Dep_{Btw}(n,G) = \max_{i \in N_{|n}} \{Int_{Btw}(i,n,G)\}$ where $Int_{Btw}(i,n,G) \geq 0$. If high, this value means that the node is "central" because of the presence of at least another node in the network: if that node is removed then node n looses a consistent part of its central role (its centrality measures decreases). If low the central role of node n is not dependent on other nodes and there is no node removal that can consistenly affects its relevance in the network. Similarly we define the *competition value* as $Comp_{Btw}(n,G) = \max_{i \in N_{|n}} \{|Int_{Btw}(i,n,G)|\}$ where $Int_{Btw}(i,n,G) \leq 0$. High competition value means that the central role of node n can be "improved" removing a particular node from the network (node n betweenness increases). In this sense the two nodes, node n and the removed one are "competitors" in the network. If low, the central position of the node cannot be improved removing a particular node from the network. Because our specific focus on single node analysis, the betweenness variation due to robustness, competition and dependence can be related to the betweenness value of the node in the starting network (the network with no node deletion) ($relRob_{Btw}(G,n) = \frac{Rob_{Btw}(G,n)}{relBtw(G,n)}$, $relDep_{Btw}(G,n) = \frac{Dep_{Btw}(G,n)}{relBtw(G,n)}$, $relComp_{Btw}(n,G) = \frac{Comp_{Btw}(n,G)}{relBtw(n,G)}$ Similarly to the interference definitions, total robustness dependence and competition value can be also used as global parameters in order to characterize the entire network. All *robustness*, *competition* and *dependence* definitions can be extended to other centrality values. Next example shows the role of node centrality robustness, dependence and competition value. Consider the network in figure 2a. Node3 and node6 have the highest values of betweenness (25.64), node4 and node5 presents the third highest value (12). A Robustness analysis of node3 and node4 allows to understand if and how much their

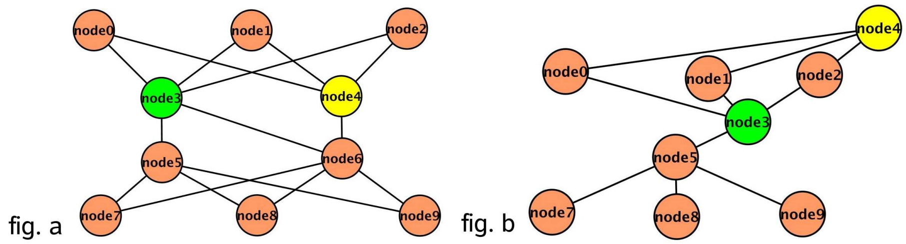

FIGURE 2. Node3 and node6 have highest betweenness (25.64). Betweenness value of node4 and node5 is 12.

high betweenness values depends on other nodes of the network. Node3 has higher robustness value (1.4824) than node4 (0.5385). In fact node4 is in the shortest paths connecting node0, node1 and node2 with node7, node8 and node9 (fig. 2a), but if we remove node6, node4 looses this role and become a "peripheral node" connecting only node0, node1, node2 between them (fig. 2b). This can not happen to node3 since it is connected to both node6 and node5. Node3 has highest dependence on node5 equals to 0.09999. The relative dependence value is 0.3118 indicating that node3 looses about the 31% of its starting betweenness value if node5 is removed from the network. Indeed, if we delete node5 the betweenness value of node3 becomes the same of node4, since they connect the same nodes through the same paths: those passing through node6. But dependence of node4 on node6 is higher (0.1143, with relative dependence 0.7619 i.e it looses about 76% of its starting betweenness value if node6 is removed from the network): as previously seen, if we remove node6 then node4 becomes a "peripheral" node and node3 becomes the only way to connect the "top" of the network with the "bottom". Also the competition value of both nodes is very informative. The highest values of node3 depends on deletion of node4 and the highest value of node4 depends on node3. In this sense they are really "competitors" in the network. But this also means that, missing one of the two nodes, its role can be replaced by the other one. If we remove node3 then node4 becomes the only connection between the "top" and "bottom" of the network. The same for node3 if we remove node4. But node4 competition values is higher (1.875). This is due to the fact that starting betweenness value of node4 is lower (12) than node3 value. So the increase of betweenness of node4 is higher, the 185% of the starting value.

5. **Interference in a transportation network: the case of Italy north-east highway.** We applied interference to the highway network of the north-east of Italy, the region included between Milan, Bologna and Trieste (see fig. 3). The network, containing 136 nodes and 144 edges have been compiled with the distance in minutes between each highway exit as reported by the official Italy highway website [13]. We choosed three highway exit as example to evaluate betweenness interference: Melegnano, Como, Mestre. First ten positive and negative values or each of these exits are reported in figures 4 and 5. Firstly we analyzed Melegnano betweenness interference: Melegnano is a critical node to connect the Milano area with the Bologna one. Closing the highway in this point means to stop the main traffic from Milano to Bologna. As expected the first ten positive interference values are all the towns between Milano Sud and Parma (see fig. 6b, the blue road). This

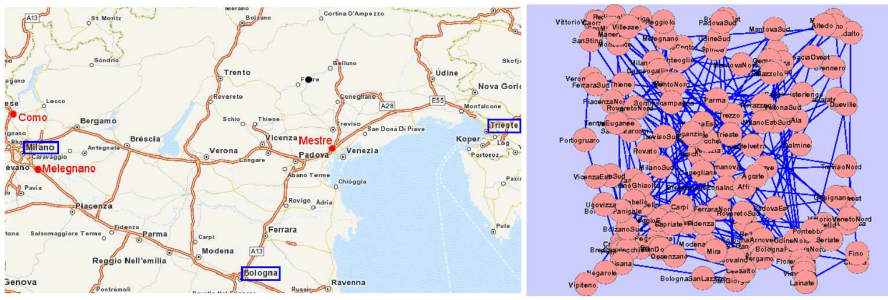

FIGURE 3. The north-east of Italy highway network and ita representation as a graph.

Melegnano

Node name	Betweenness Interference
BresciaCentro	-0.23
BresciaOvest	-0.2
Ospitaletto	-0.2
Grumello	-0.19
Manerbio	-0.19
Palazzolo	-0.19
Ponteoglio	-0.19
Pontevico	-0.19
Rovato	-0.19
Bergamo	-0.18

Como

Node name	Betweenness Interference
PadovaEst	-0.03
Grisignano	-0.02
Mestre	-0.02
Mira	-0.02
Mirano	-0.02
Montebello	-0.02
Montecchio	-0.02
PadovaOvest	-0.02
VeronaSud	-0.02
VicenzaEst	-0.02

Mestre

Node name	Betweenness Interference
Spinea	-1.72
Preganziol	-1.66
Venezia	-1.06
Ala	0.0
Belluno	0.0
BolognaArcoveggio	0.0
BolognaFiera	0.0
BolognaPanigale	0.0
BolognaSanLazzaro	0.0
BolzanoNord	0.0

FIGURE 4. First ten negative interference value of Melegnano, Como and Mestre

Melegnano

Node name	Betweenness Interference
MilanoSud	0.33
Lodi	0.25
Sesto	0.24
Casalpusterlengo	0.23
PiacenzaNord	0.2
Fidenza	0.18
Fiorenzuola	0.18
Parma	0.18
PiacenzaSud	0.18
ReggioEmilia	0.18

Como

Node name	Betweenness Interference
Fino	0.08
Lainate	0.07
MilanoNord	0.05
Agrate	0.04
Cavenago	0.04
MilanoEst	0.04
Monza	0.04
Trezzo	0.04
Bergamo	0.03
Capriate	0.03

Mestre

Node name	Betweenness Interference
Mira	1.77
PadovaEst	0.06
Mirano	0.05
Grisignano	0.04
Montebello	0.04
Montecchio	0.04
PadovaOv...	0.04
Soave	0.04
VeronaEst	0.04
VicenzaEst	0.04

FIGURE 5. First ten positive interference value of Melegnano, Como and Mestre

region is the one that is more affected by Melegnano. If Melegnano is part of the network, these towns are the shortest way to connect Milano and its area with Bologna and its area. This is perfectly captured by the positive interference of Melegnano with the highway exits of this region. If Melegnano is removed from the network, for example if it is blocked by a road accident, the road between Milano sud and Parma can not connect Milano and Bologna. To understand the alternative paths, we consider the negative interference of Melegnano. As reported in the table, the first then negative values belong to the region around Brescia Centro. As expected, if Melegnano is blocked, the interference analysis predicts that Brescia Centro is the new critical point to connect Milano and Bologna, through the highway from Brescia Centro to Fiorenzuola (see fig. 6 a, the green road). Even if these nodes are far from Melegnano in the network, the interference analysis can easily predict that they are undirectly influenced by Melegnano. As second example we

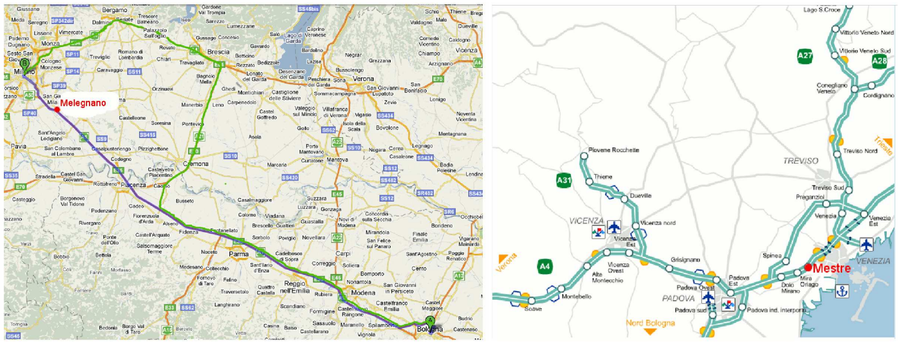

fig. a fig. b

Figure 6. **fig. a** The shortest road from Milano to Bologna pass through Melegnano (blue road). If Melegnano is blocked, the shortest road is the one passing through Brescia (green road). This behaviour is exactly predicted by the interference analysis. **fig. b**. Mestre is a critical node to connect Trieste with the rest of Italy. Notice the recently built alternative path passing through Spinea, called “passante of Mestre”.

computed Como interference. In the north-east highway Como is only a peripheral node (see fig. 3). As expected its interference value is high only for its neighbour, and susbstantially smaller than interference of Melegnano (Como interference max value = 0,08, Melegnano interference max value 0,32). This shows that interference analysis really reflects the importance of undirect interaction between nodes. As third example we computed interference for Mestre (see fig. 6 b). Mestre is well known as an important connection between Trieste (the extreme east of Italy) and important nodes as Milano and Bologna. Its interference analysis results in high negative interference with respect to Spinea, Preganziol, Venezia. This is totally in agreement with the real situation: The road passing through Spinea, Preganziol and Venezia called “passante of Mestre”, was recently built in order to solve traffic jam problem of the Mestre Area, always congested because of traffic from Milano and Bologna to the Venice port and to Trieste. To confirm this analysis we can modify the distance in minutes between Mira and Mestre. In high traffic condition, as for example during summer week-end when a lot of people moves to the Venice area for holidays, the real distance between Mira and Mestre is more than 20 minute. In this case the shortest path connecting Trieste with Milano and Bologna is the one passing through Spinea. We modified the distance of the network according with these value. An interference analysis of Spinea in the updated network, shows high negative interference (= 1.7) with respect to Mira and Mestre. As expected, according to the interference analysis, the role of Spinea is exactly to reroute the traffic of Mestre: if Spinea is not part of the network, its negative interference with Mestre and Mira predict that Mestre and Mira are more congested as it was before the “passante of Mestre” building.

6. **Further considerations and conclusions.** As a further implication of our approach, we can consider centrality interference and robustness as natural generators of network modularity. Indeed, a new clusterization algorithm can be derived if we group nodes depending on their interference value. Given a node, we may compute its interference activity of the network and, then, we may group in the same cluster all nodes having high interference values. This interference-based modular decomposition of a network allows grouping of nodes according to their response to the deletion (or addition) of specific nodes in the network. Importantly, this approach may lead to a less purely mathematical, but more contextual-oriented method of network modularization. Notably, it is well known that scale free networks are not easily affected by randomly removing single nodes [1][3][8]. So a possible scenario of application of interference analysis implies removal of groups of nodes. Definition of interference can be easily adapted to such a situation, where removing a subset of nodes is considered. The method can be limited by the fact that sometimes removing one ore more nodes can result in disconnected components. To solve this problem, future works should redefine the betwenness and other centrality values in order to allow the calculation also in the case of disconnected components. Anyway disconnected components after node removal are notified to the user by the Interference software used for the computation (http://www.cbmc.it/%7Escardonig/interference/Interference.php). The betweenness is only one example of the possible usage of the interference notion that is a more general notion. The Interference software allows to calculate interference also to other centralities as radiality, closeness, stress, eccentricity and centroid value, but it could be applied also to other centralities or node oriented measures as eigenvector centrality, clustering coefficient, assortativity and so on (see [5] for an excellent review of all these measures). In conclusion, the introduction of centrality interference and robustness allows understanding how a network locally rearranges itself when nodes are removed or added from/to a network, a common situation in several applications of networks analysis. The main issue is the node oriented point of view: a potential user should not be interested in global behaviour or properties of the network, but in modeling single situations of interest. The possible scenario is that a potential user has one or more targets nodes, and want to evaluate their area of influence and the effects of their removal in the network. An interference analysis allows also to identify which parts of the networks are influenced by single nodes or by modification on the functionality of such nodes; with robustness and related notions (dependence and competition values) we may infer how much the central role of a node can be affected by other nodes in the network.

REFERENCES

[1] R. Albert, H. Jeong and A.-L. Barabási, *Error and attack tolerance of complex networks*, Nature, **406** (2000), 378–382.

[2] A.-L. Barabási and R. Albert, *Emergence of scaling in random networks*, Science, **286** (1999), 509–512.

[3] A.-L. Barabási and Z. N. Oltvai, *Network biology: Understanding the cell's functional organization*, Nature Reviews Genetics, **5** (2004), 101–113.

[4] U. S. Bhalla and R. Iyengar, *Emergent properties of networks of biological signaling pathways*, Science, **283** (1999).

[5] G. Caldarelli, "Scale-Free Networks: Complex Webs in Nature and Technology (Oxford Finance)," Oxford University Press, USA, June 2007.

[6] P. Crucitti, V. Latora, M. Marchiori and A. Rapisarda, *Error and attack tolerance of complex networks*, News and expectations in thermostatistics, Phys. A, **340** (2004), 388–394.

[7] J. A. Goguen and J. Meseguer, *Security policies and security models*, Symposium on Security and Privacy, IEEE Computer Society Press, (1982), 11–20.
[8] H. Jeong, S. P. Mason, A. L. Barabási and Z. N. Oltvai, *Lethality and centrality in protein networks*, Nature, **411** (2001), 41–42.
[9] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai and A. L. Barabási, *The large-scale organization of metabolic networks*, Nature, **407** (2000), 651–654.
[10] D. Koschützki, K. A. Lehmann, L. Peeters, S. Richter, D. T. Podehl and O. Zlotowski, *Centrality indices*, in "Network Analysis: Methodological Foundations" (eds. U. Brandes and T. Erlebach), Springer, (2005), 16–61.
[11] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii and U. Alon, *Network motifs: Simple building blocks of complex networks*, Science, **298** (2002), 824–827.
[12] M. E. J. Newman, *Modularity and community structure in networks*, Proceedings of the National Academy of Sciences, **103** (2006), 8577–8582.
[13] The official, *Autostrade per l'Italia*, http://www.autostrade.it/, 2011.
[14] G. Scardoni, M. Petterlini and C. Laudanna, *Analyzing biological network parameters with CentiScaPe*, Bioinformatics, **25** (2009), 2857–2859.
[15] C. M. Schneider, T. Mihaljev, S. Havlin and H. J. Herrmann, *Suppressing epidemics with a limited amount of immunization units*, Physical Review E, **84** (2011), 061911+.
[16] S. H. Strogatz, *Exploring complex networks*, Nature, **410** (2001), 268–276.
[17] Duncan J. Watts and Steven H. Strogatz, *Collective dynamics of 'small-world' networks*, Nature, **393** (1998), 440–442.

Received December 2011; revised July 2012.

E-mail address: giovanni.scardoni@gmail.com
E-mail address: Carlo.Laudanna@univr.it