Congress of the European Regional Science Association

23-27 August 2005, Free University of Amsterdam

Emergent topological and dynamical properties of a real inter-municipal commuting network: perspectives for policy-making and planning

Andrea De Montis[(1)], Marc Barthélémy[(2)], Michele Campagna[(3)], Alessandro Chessa[(4)], Alessandro Vespignani[(5)]

[(1)]Dipartimento di Ingegneria del Territorio, Sezione Costruzioni e Infrastrutture, Università degli Studi di Sassari Via De Nicola – 07100 Sassari, Italy, e-mail: andreadm@uniss.it

[(2)]CEA-Centre d'Etudes de Bruyères-le-Châtel, Département de Physique Théorique et Appliquée BP12, 91680 Bruyères-le-Châtel, France, e-mail: Marc.Barthelemy@th.u-psud.fr

[(3)] Dipartimento di Ingegneria del Territorio, Sezione Urbanistica, Università degli Studi di Cagliari, piazza d'Armi, 16, 09123 Cagliari, e-mail: campagna@unica.it

[(4)]Dipartimento di Fisica, Università degli Studi di Cagliari, Cittadella Universitaria di Monserrato, 09042 Monserrato, Italy, e-mail: alessandro.chessa@dsf.unica.it

[(5)] School of Informatics, Center for Biocomplexity and Department of Physics, Indiana University, Bloomington, USA, e-mail: alexv@indiana.edu

Abstract

In the domains of urban and environmental planning, spatial analysis and regional science, many scholars have showed in the past years an increasing interest for the research developments on complex networks. Their positions range from the theory statements on the need to apply complex network analysis to spatial phenomena (Salingaros, 2001) to the empirical focus on more quantitative research about urban space syntax (Jiang and Claramunt, 2004). As regards transportation systems analysis, interesting results have been obtained in the study of networked subway (Latora and Marchiori, 2002; Gastner and Newman, 2004) and airports (Barrat et al, 2004) systems.

In this paper, the authors aim at the study of the inter-municipal commuting network of Sardinia, Italy, described by the system of study or work-led habitual movements of its citizens. In this complex network, the nodes correspond to urban centres while the links to positive commuting exchanges among municipal towns. Following the analysis developed by Barrat et al. (2004), topological and dynamical properties are investigated as they stem from the underlying complex network. Furthermore, in the perspective of policy-making and planning, the emerging network behaviours are critically compared to the geographical, social and demographical aspects of the transportation system.

JEL classification codes: R15, R40, Z13

Key-words: small worlds, scale free networks, weighted complex networks, commuting networks, socio-economic networks

1 Introduction

The informatics revolution has enabled the systematic gathering and handling of data sets on several large scale networks, leading to the analysis of their detailed structural features. In particular, mapping projects of the WWW and the physical Internet offered the first chance to study the topology and traffic of large-scale networks. Gradually other studies followed describing population networks of practical interest in social science, critical infrastructures and epidemiology (Barabasi and Albert (2000), Dorogovtsev and Mendes 2003, Pastor-Satorras and Vespignani 2004). The possibility of accessing and mining large-scale data sets permits a more detailed statistical analysis and theoretical characterization of correlation patterns, hierarchies and community structure of these networks, along with the setup of new modeling frameworks. In particular, the systematic study of these systems has shown the ubiquitous presence of *complex* features, mathematically encoded in statistical distributions with heavy-tails, diverging fluctuations, self-organization and emerging phenomena and patterns.

The statistical properties induced by these complex features indicate the presence of topological and structural properties that do not find an explanation in the paradigm put forward by Erdos and Renyi with the random graph model (Erdos and Renyi, 1960). Indeed, even if the Erdos-Renyi model rationalizes the small-world property of networks (i.e. the short distance measured in number of links among nodes), it fails in reproducing the high level of local cohesiveness observed in many networks (Watts and Strogatz 1998). Additionally, several of these networks are characterized by a statistical abundance of nodes with a very large degree k; i.e. the number of connections to other nodes. For these "scale-free" networks, the degree probability distribution $P(k)$ spans a wide range of k values which signals the appreciable occurrence of large degree nodes, the "hubs" of the system. Finally, it is important to remark that the topological features of networks turn out to be extremely relevant since they have a strong impact in assessing their physical properties such as their robustness, vulnerability, or their ability to spread a disease (Albert and Barabasi, 2002; Dorogovtsev and Mendes, 2003; Pastor-Satorras and Vespignani, 2004).

More recently, the activity on complex networks has been extended to the characterization of weighted networks. This representation allows the consideration of features pertaining to the dynamics and traffic flows occurring on networks, adding another dimension in the description of these systems. Also in this case, the analysis and the characterization of weighted quantities have pointed out the presence of large scale heterogeneity and non-trivial correlations (Barrat et al, 2004). The weighted graph representation provides valuable elements that open the path to a series of questions of

fundamental importance in the understanding of networks. Among those, it is crucial to address the issue of how dynamics and structure affect each other and their impact on the basic properties of spreading and congestion phenomena. Finally, it spurs the more theoretical questions on how all these properties may be considered in generative networks models.

In the realm of urban and environmental planning, spatial analysis and regional science, the interest for complex networks has noticeably increased during the last years. In his investigation of the structure of the modern urban fabric, Salingaros (2001 and 2003) invokes the concepts of small world network and scale free properties and suggests a parallelism between the city and an ecosystem. Starting from the principle that connective webs are the main source of urban life, Salingaros invites planners to apply complex networks analysis to understand properties such as resilience and self-organization. While these studies are mostly based on conceptual arguments, other scientists (Shiode and Batty, 2000; Batty, 2001 and 2003) promote the development of quantitative analyses with a twofold perspective: (i) explaining the behavior of complex networks such as the world wide web in terms of geography, demography and economics; and (ii) improving the efficiency of advanced spatial analysis GIS-based tools, by integrating the foundation of complex networks analysis into geographic information science. Batty (2001) along with Salingaros acknowledges the existence of a link between efficient urban connectivity and self-preserving behaviors and fosters the development of complex networks to study the functional mechanisms of cities. It is interesting to note that a similar approach has been extremely fruitful in the study of the Internet in relation with the geographical and social environment (Yook et al 2002; Gorman, 2001; Barthélemy et al 2003; Gorman and Kulkarni 2004).

In the domain of geographic information science, Jiang and Claramunt (2004) proposed to use complex network tools to investigate quantitatively the urban space syntax. With the aim at improving the effectiveness of network GIS advanced analysis, they use topological measures of connectivity, average path length and generalised clustering to compare three large street networks of the cities of Gälve, Munich and San Francisco. Fertile research directions are also currently explored in the domain of transportation systems and human traffic analysis. For example, the small-world properties of the Boston subway system have been characterized Latora and Marchiori (2000-2003) and Sen et al (2003) studied the topology of the Indian railway network. At the urban level, a network study has been carried out by Chowell et al (2003) using large scale simulations (with census and demographic data integration) in order to describe urban movements of individuals in

Portland, Oregon (USA). On a larger scale, recent studies have focused on the network properties of major Transportation infrastructure such as the US Interstate highway network and the airport network. (Gastner and Newman, 2004; Guimera et al, 2003). Finally, Barrat et al (2004) have provided the first study of the worldwide airports network including a characterization of the traffic flows and their correlation with the topological structure. In this area, the network approach might provide relevant information on issues such as traffic analysis and risk assessment in case of damages and attacks. It is worth remarking that in this context, the topology is not enough to understand different processes which take place on these networks, urging for a weighted representation of these networks that includes traffic and other network attributes.

In the present work we use a network approach to study the Sardinian inter-municipal commuting network (SMCN), which describes the habitual daily work and study-led movements among 375 municipalities in the Italian region of Sardinia. We obtain a weighted network representation in which the vertices correspond to the Sardinian municipalities and the valued edges to the amount of commuting traffic among them. We provide a detailed quantitative study of the resulting network with the aim of a first characterization of the structure of human traffic at the inter-city level and its relation with the topological structure defined by the connectivity pattern among cities. Very interestingly, while the network topology appears to fit within the standard random paradigm, the weighted description offers a very different perspective with complex statistical properties of the commuting flows and highly non-trivial relations between weighted and topological properties. We discuss these properties in relation with the geographical, social and demographical aspects of the system and show the potentiality of the network approach as a valuable tool at various stages of policy-making and environmental planning processes.

The paper is organized as follows. In the next chapter, we describe the dataset and in the third chapter, we carry out the analysis of the topology of the network. In chapter 4, we analyze the traffic and weight properties of this network and in chapter 5, we discuss the implications of our network analysis results in terms of urban and environmental planning. In chapter 6, we compare topological and weighted network analysis results with socio-economic characteristics of the Sardinian municipalities and we discuss their relations. We finally summarize our results and discuss the different research directions that this work suggests.

2 Setting the case study: Data and geographical features

Sardinia is the second largest Mediterranean island with an area of approximately 24,000 square kilometers and 1,600,000 inhabitants. Its geographical location and morphological characters have determined an important history of commercial and cultural relations with trans-borders external communities and have also favoured the development of a strong social identity and of important political movements toward self-reliance and autonomy. At the date of 1991, the island was partitioned in 375 municipalities, the second simplest body in the Italian public administration, each one of those generally corresponding to a major urban centre (in Figure 1, on the left, we report the geographical distribution of the municipalities). For the whole set of municipalities the Italian National Institute of Statistics (Istat) has issued the origin-destination table (ODT) corresponding to the commuting traffic at the inter-city level. The ODT is constructed on the output of a survey about commuting behaviors of Sardinian citizens. This survey refers to the daily movement from the habitual residence (the origin) to the most frequent place for work or study (the destination): the data comprise both the transportation means used and the time usually spent for displacement. Hence, ODT data give access to the flows of commuters who regularly move among the Sardinian municipalities. In particular we have considered the *external flows* $i \rightarrow j$ which measure the movements from any municipality i to the municipality j and we will focus on total flows of individuals (workers and students) commuting throughout the set of Sardinian municipalities by all means of transportation. This data source allows the construction of the Sardinian inter-municipal commuting network (SMCN) in which each node corresponds to a given municipality and the links represent the presence of a commuting flow among municipalities.

The standard mathematical representation of the resulting network is provided by the adjacency matrix A of elements (a_{ij}). The elements on the principal diagonal (a_{ii}) are set equal to zero, since intra-municipal commuting movements are not considered here. Off-diagonal terms a_{ij} are equal to 1 in the presence of any non-zero flow between i and j ($i \rightarrow j$ or $j \rightarrow i$) and are equal to 0 otherwise. The adjacency matrix is then symmetric $a_{ij} = a_{ji}$ and describes regular bi-directional displacements among the municipalities.

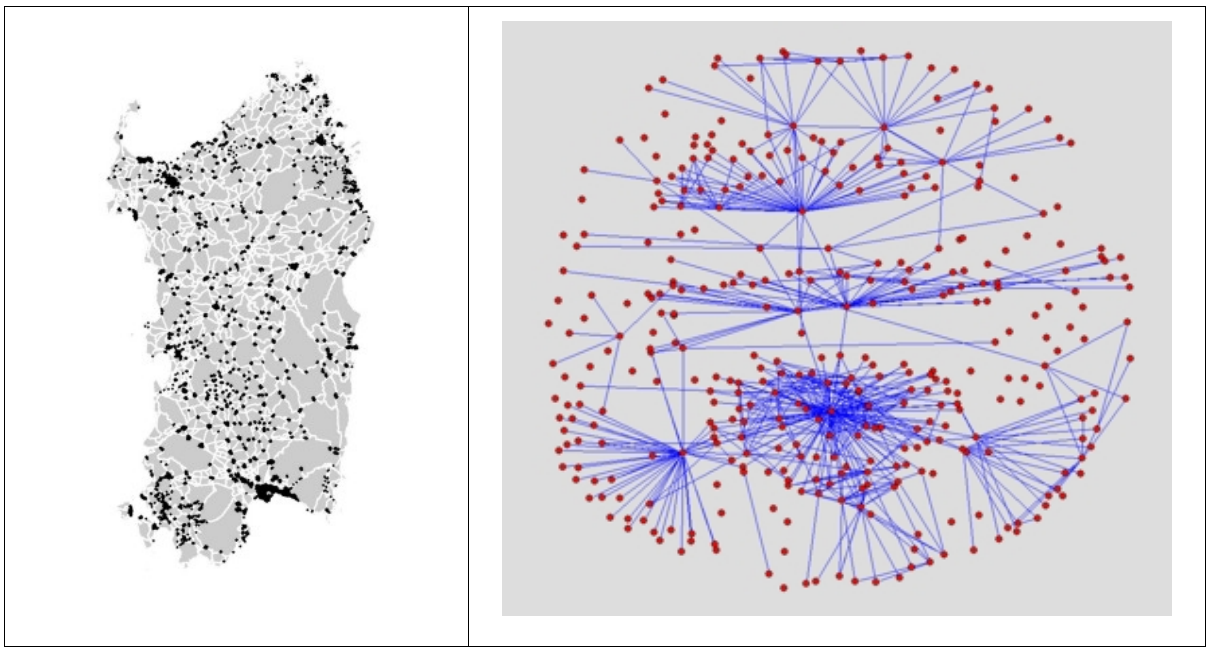

Figure 1: Geographical versus topologic representation of the SMCN: the nodes (red points) correspond to the towns, while the links to a non zero flow of commuters between two towns.

The adjacency matrix contains all the topological information about the network but the dataset also provides the number of commuters attached to each link. It is therefore possible to go beyond the mere topological representation and to construct a weighted graph where the nodes still represent the municipal centres but where the links are valued according to the actual number of commuters. Analogously to the adjacency matrix A, we thus construct the symmetric weighted adjacency matrix $W(w_{ij})$ in which the elements w_{ij} are computed as the sum of the $i \rightarrow j$ and $j \rightarrow i$ flows between the corresponding municipalities (per day). The elements w_{ij} are null in the case of municipalities i and j which do not exchange commuting traffic and by definition the diagonal elements are set to zero $w_{ii} = 0$. According to the assumption of regular bi-directional movements along the links, the weight matrix is symmetric and the network is described as an undirected weighted graph. The weighted graph provides a richer description since it considers the topology along with the quantitative information on the dynamics occurring in the whole network. It is however important to stress that while the nodes correspond to municipalities located in the physical space, the graph representation does not contain any information explicitly related to geographical distances and other spatial characteristic of the network. The definition of a network representation that correlates topological and traffic characteristics with the spatial properties of the SMCN requires more refined investigation and will be presented elsewhere (De Montis et al, forthcoming).

3 The analysis of the topological properties of the SMCN

In this section, we analyze the topology of the SMCN and we propose possible territorial explanations for our findings. This network is relatively small being characterized by *N=375* vertices and E=16,248 edges. The degree *k* of nodes, which measures the number of links of each node, ranges between 8 and 279 and exhibits a high average value $\langle k \rangle = 43.33$. The degree can be considered as a first index of the topological centrality of a vertex: the importance of the corresponding municipal centre as exchange point for commuters in the network. A measure of the topological distance within the network is given by the shortest path length defining the distance between nodes in terms of links to be traversed by using the shortest possible path between the nodes. On the average, the shortest path length is l = 2.0, while the maximum length is just equal to 3. These values are small compared to the number of nodes, in agreement with a small-world behavior for which l typically scales as the logarithm of N.

Further information on the network's topology is provided by the degree distribution *P(k)* defined as the probability that any given node has degree *k*. In this case it is essential to distinguish between an 'exponential' network similar to the usual random graph for which *P(k)* decreases faster than an exponential, and scale-free networks indicative of hub-like hierarchies and for which *P(k)* decreases typically as a power law (Albert and Barabasi, 2002; Amaral et al, 2000). The connectivity probability distribution is skewed and is relatively peaked around a mean value of order 40 (Figure 2).

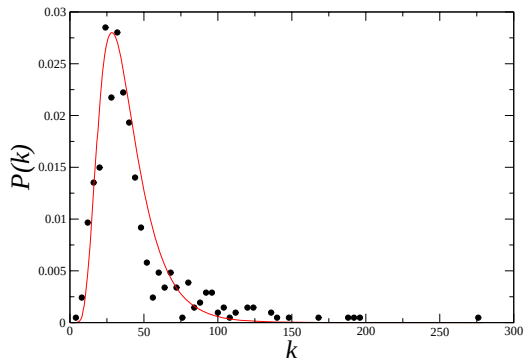

Figure 2: Plot of the probability distribution of connectivity. The red line is a lognormal fit.

The tail of the distribution contains outliers with respect to a regular exponential decay, signaling the presence of municipal hubs, which exchange commuters with many others municipalities. However, a closer inspection of the distribution tail shows a decay not

compatible with a scale–free distribution. The tail has indeed a power-law behavior but with an exponent that ensures the convergence of the relevant moments of the distribution and does not imply a scale-free heterogeneity. In particular, the variance converges to a finite limit for infinitely large systems in contrast with scale-free networks, which have an infinite variance for infinite networks.

The clustering coefficient measures the level of cohesiveness around any given node. It is expressed as the fraction of connected neighbors:

$$C(i) = 2E(i)/k_i(k_i - 1)$$

where $E(i)$ is the number of links between the k_i neighbors of the node i and $k_i(k_i - 1)$ is the maximum number of possible interconnections among the neighbors of the node. This quantity is defined in the interval [0,1] and measures the level of local interconnectedness of the network. If $C(i) = 0$, the neighbors of the node i are not interconnected at all, while $C(i) = 1$ corresponds to the case where all the neighbors are interconnected. A large clustering thus indicates the existence of a locally well-connected neighborhood of nodes. It is often convenient to average $C(i)$ over all nodes with a given degree k leading to

$$C(k) = \frac{1}{NP(k)} \sum_{i/k_i=k} C(i) \quad .$$

In Figure 3, on the left, we plot this coefficient versus the degree: we observe a slow decay (when k varies over 2 decades, $C(k)$ varies from 0.8 to 0.2 approximatively).

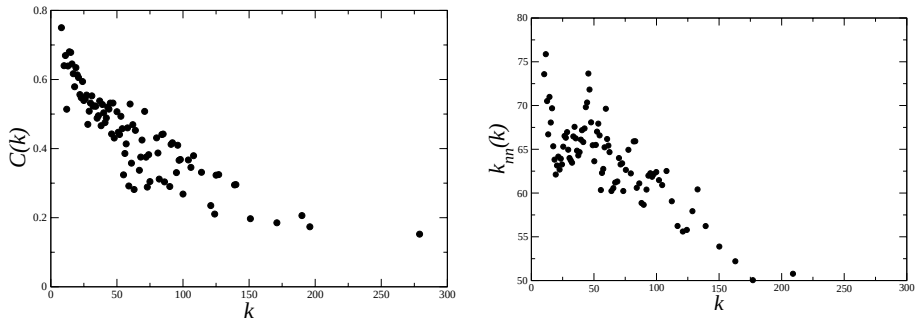

Figure 3 On the left: Scatterplot of the clustering coefficient versus degree. There is an average decay as a slow power law with exponent of order 0.4. On the right: assortativity of the SMCN showing a slight disassortative behaviour.

This indication of a decreasing clustering suggests that the SMCN behaves like other studied real networks, in that low degree municipalities belong to well interconnected working or studying communities while the high degree vertices, the urban hubs, connect otherwise disconnected regions. It is worth noting that this result is similar to the one

obtained by Chowell et al (2003) for urban movement. There is a striking difference, however, since the SMCN is topologically very close to a random graph, while the network studied in Chowell et al (2003) is a scale-free network. This similarity with a random graph is also confirmed by the average value of the clustering coefficient $\langle C \rangle = 0.26$, on the same order of the one computed for the case of a generalized random graph given by

$$\langle C \rangle_{RG} = \left(\langle k^2 \rangle - \langle k \rangle\right)^2 / \left(N \cdot \langle k \rangle^3\right) = 0.24 \quad .$$

Another important property is the similarity of the neighbors of a node which is measured by the assortativity given by the average degree of the nearest neighbors of a given node *i*:

$$k_{nn}(i) = \frac{1}{k_i} \sum_{j \in V(i)} k_j \quad ,$$

where $V(i)$ denotes the set of neighbors of *i*. Similarly to the clustering coefficient, we can average the assortativity over nodes with a given degree leading to

$$k_{nn}(k) = \frac{1}{NP(k)} \sum_{i / k_i = k} k_{nn}(i) \quad ,$$

where $NP(k)$ is the total number of nodes of degree *k*. If $k_{nn}(k)$ increases with *k*, nodes have a tendency to connect to nodes with a similar degree and the network is said to be assortative. In contrast, if $k_{nn}(k)$ is decreasing with *k*, small degree nodes connect preferentially to hubs and the network is disassortative. The result for the SMCN is shown in Figure 2, on the right, and displays a relatively constant k_{nn} with a small decreasing trend. This result seems to indicate that the SMCN belongs to the class of disassortative mixed networks. This disassortative behavior is typical for technological or transportation networks for which hubs tend to connect to smaller nodes.

4 The SMCN weighted network analysis

The ODT dataset provides the number of commuters among municipalities, allowing the construction of the weighted graph representation of the SMCN. The analysis of the weights provides information on the traffic and the study of the topology-traffic correlations shed some light on the global characteristics of the human inter-cities traffic. The analysis of these commuting flows informs us on the system dynamical behavior and on the needs of the

commuters in terms of regional transportation services and infrastructure. This information might provide relevant indication to researchers and decision-makers for addressing environmental policies and planning.

4.1 Weight analysis

The values of the weights between pairs of vertices ranges between 1 and $w_{\max} = 13{,}953$ while its average value $\langle w \rangle = 23.32$ is much lower than $w_{\max}$. This fact is a signature of a high level of heterogeneity in the SMCN, since a peaked distribution of weights would very unlikely give a maximum value of 3 orders of magnitude larger than the average. Indeed, the probability distribution *P(w)* that any edge has a weight *w* displays a power-law decay- $P(w) \sim w^{-\gamma_w}$ with an exponent γ_W of order 2 (Figure 4, on the left).

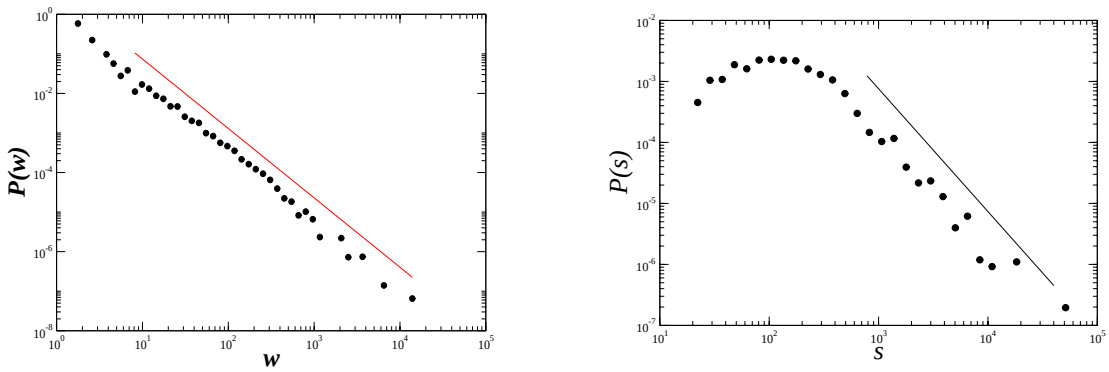

Figure 4 On the left: loglog plot of the probability distribution of the weights (the straight line is a power law fit of exponent 1.8). On the right: loglog plot of the probability distribution of the strength (straight line: as a guide to the eye, we plotted a power law of exponent 2).

The heterogeneity in the weight values implies that commuters' flows are distributed in a broad spectrum with a large heterogeneity. Another relevant quantity that characterizes the traffic is given by the strength of a node, defined as

$$s(i) = \sum_{j \in V(i)} w_{ij} \quad .$$

This quantity corresponds to the total traffic commuting in the town *i* and can be considered as a measure of the traffic centrality of a municipality. For the SMCN, the strength varies between 1 and 64,834 while its average value is equal to 1,010. The probability distribution of the strength *P(s)* that any given node has a strength *s* is well fitted by a power-law regime- $P(s) \sim s^{-\gamma_s}$ with exponent $\gamma_s \approx 2$ (Figure 4, on the right). The strength provides an intuitive parameter for the description of the centrality of Sardinian municipalities according to the actual traffic handled generated and received by the municipality.

It is worth commenting the implications of a power-law behavior of the statistical distributions *P(s)* and *P(w)*. This behavior defines the so-called heavy tailed distributions which have a virtually unbounded variance. A peculiar fact about a distribution with a heavy tail is that there is a finite probability of finding vertices with weight or strength much larger than the average value. In other words, the consequence of heavy tails is that the average behavior of the system is not typical. The characteristic weight (or strength) is the one that, picking up a vertex at random, should be encountered most of the times. In the distributions shown in Figure 4 most of the times vertices will have a small degree, but there is an appreciable probability of finding vertices with large degree values. Yet all intermediate values are probable and the average degree does not represent any special value for the distribution. This is clearly opposite to bell-shaped distribution observed for the degree of the SMCN, in which the average value is very close to the maximum of the distribution and represents the most probable value in the system. The power-law behavior and the relative exponent thus represent a quantitative measure of the level of heterogeneity of the network's traffic.

Another important question concerns the distribution of the traffic among the different roads. All the connections could carry a similar flow or on the contrary one connection could dominate. A convenient measure of this is given by the disparity (Barthélemy et al, 2005), defined as

$$Y_2(i) = \sum_j \left(\frac{w_{ij}}{s_i} \right)^2 \quad .$$

For hubs (k>>1), this quantity enables to distinguish situations (Figure 5, on the right) for which all weights are of the same order ($Y_2 \approx 1/k$ <<1) from situations where only a small number of connections dominate (Y_2 is of order 1/n where n is of order unity).

We computed the value of $Y_2(i)$ for each node and we average these quantities for each degree in order to obtain $Y_2(k)$. The result is shown on Figure 5, on the left. We obtain here an average behavior of the form $kY_2 = k^{1-\theta}$ with $\theta \approx 0.4$. This result shows that the weights on the links attached to a given node are very heterogeneous. In other words, there are only a few dominant connections, while the traffic on all the other roads is very small.

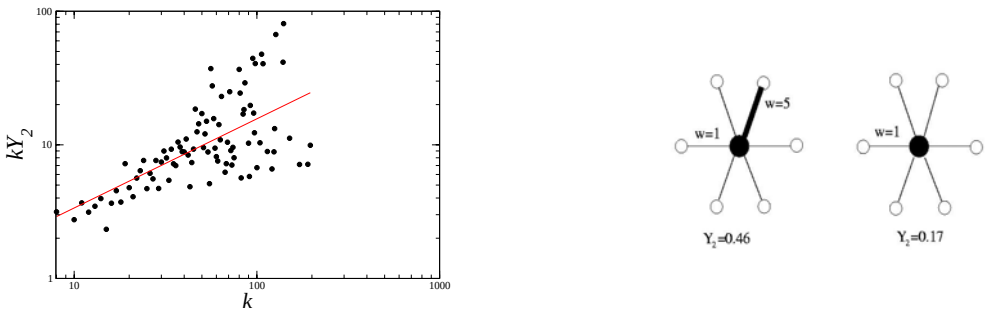

Figure 5 On the left: loglog plot of the disparity versus connectivity for the SMCN; on the right: illustration of the disparity for two very different cases. When a few connections dominate, Y_2 is of order 1; in contrast if all connections have the same weight, Y_2 is of order 1/k.

4.2 *The analysis of traffic-topology correlations*

As noted above, it is possible to rank cities according to their degree k or their strength s leading to the same result with some re-ordering differences. In order to inspect in more detail the relation between the degree and the traffic, we show in Figure 6 the strength of nodes as a function of their degree. Despite the existence of inevitable fluctuations, we observe over a wide range of degrees a power-law behavior of the form $s(k) \sim k^{\beta}$ with an exponent $\beta \approx 1.90$. Independent weights and connectivities would give a value $\beta = 1.0$ (Barrat et Al., 2004) and the result obtained here reveals that there is strong correlation between the traffic and the topology. The strength of the nodes grows faster than their degree: the more a municipality is connected with other centres the much more it is able to exchange commuters' flows or, in other terms, the traffic per connection is not constant and increases with the number of connections.

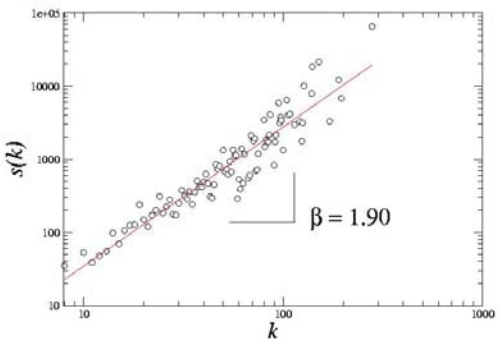

Figure 6 Average strength of the municipalities as a function of the degree.

The usual topological clustering coefficient used in the previous section could lead to incorrect conclusions since it doesn't take the actual traffic into account. In order to understand the weight structure and their relation with topology, the weighted clustering coefficient can be defined as follows (Barrat et al, 2004)

$$C^W(i) = \frac{1}{s_i(k_i - 1)} \sum_{j,h} \frac{(w_{ij} + w_{ih})}{2} a_{ij} a_{ih} a_{jh}$$

where a_{lm} is an element of the adjacency matrix and where $s_i(k_i - 1)$ is a normalization factor which ensures that $C^W(i)$ belongs to [0,1] (averaging over all nodes of same degree k gives then $C^W(k)$). This weighted quantity counts for each triple formed in the neighborhood of the vertex i the weight of the two participating edges starting from i. In this way, we are not just considering the number of closed triangles but also their total relative weight with respect to the vertex' strength. In the case of random networks $C^W = C$ but in real weighted networks, we can however face two different situations. If $C^W < C$ the topological clustering is generated by edges with low weight and therefore the cohesiveness is less important in terms of traffic properties. On the contrary, if $C^W > C$ we are in presence of a network in which the interconnected triples are more likely formed by edges with larger weight. We show in Figure 7 (on the right) an example of such a situation. It is clear that the interconnected triple has a major role in the network dynamics and organization, and that the clustering properties are clearly underestimated by a simple topological analysis. The weighted clustering coefficient takes into account this property by an appropriate mathematical form.

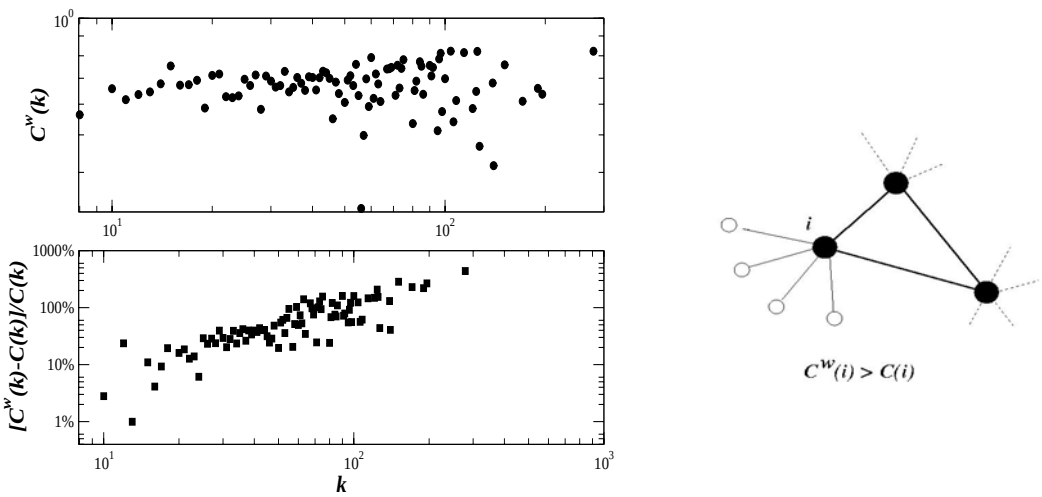

Figure 7 On the left, top: weighted clustering coefficient versus connectivity; on the left, bottom: relative difference between the weighted and the topological clustering coefficients versus connectivity; on the right: large cities are well interconnected by large flow links (bold edges must be considered with a larger weight).

In the SMCN the values of the weighted clustering coefficient are larger than the corresponding topological values $C^w \rangle C$ (Figure 7, on the left) over the entire range of

degree values. In contrast with the topological clustering, the weighted clustering coefficient is approximately constant over the whole range of connectivity.

This result reveals the existence of a rich-club phenomenon in that important cities form a group of well interconnected nodes and that weight heterogeneity is enough in order to balance the lack of topological clustering. The rich-club phenomenon implies that in the SMCN the interconnected triplets are more frequently built by edges with a higher weight (Figure 7, on the right): in terms of local cohesiveness of workers' and students' commuting system, two different destinations available from a given city are more likely to be connected if the traffic flow leading to them is large.

Similarly to the weighted clustering, we need to integrate the information on weights in a sensible definition of assortativity. This can be easily done by introducing the weighted average degree of the nearest neighbors of a given node *i*:

$$k^{w}{}_{nn}(i) = \frac{1}{s_i} \sum_{j \in V(i)} w_{ij} k_j .$$

In this case, we perform a local weighted average of the nearest neighbor degree according to the normalized weight of the connecting edges. This definition implies that $k^{w}{}_{nn} > k_{nn}$ if the edges with the larger weights are pointing to the neighbors with larger degree and $k^{w}{}_{nn} < k_{nn}$ in the opposite case. The $k^{w}{}_{nn}$ thus measures effective affinity to connect with high or low degree neighbors according to the magnitude of the actual interactions. In the SMCN the weighted average degree of the nearest neighbors displays an increasing pattern, which is a signature of assortativity (Figure 8).

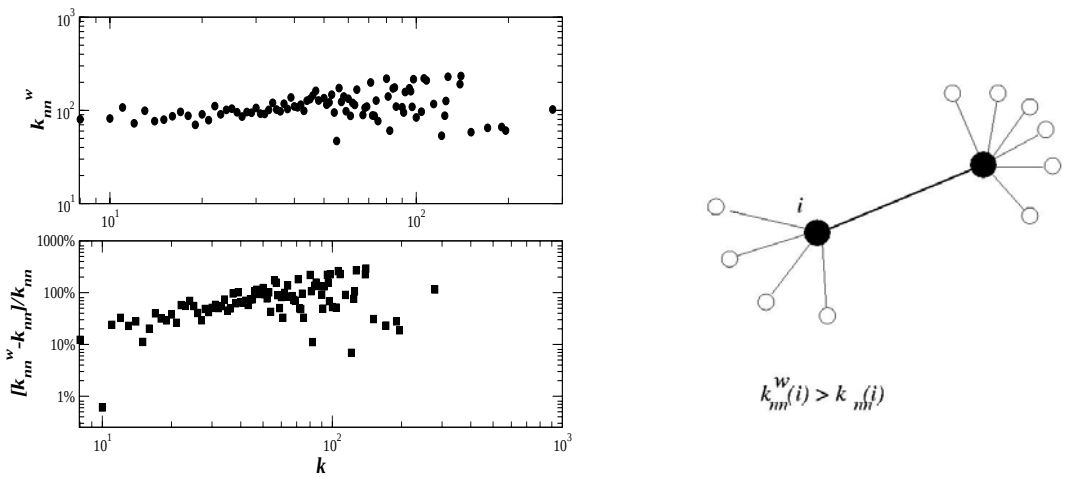

Figure 8 On the left, top: weighted assortativity; on the left, bottom: relative difference between the weighted and the topological assortativity; on the right: a difference between the weighted and the topological assortativity indicates that large cities are connected by large flow links.

This result is in sharp contrast with the behavior of the topological assortativity and indicates an affinity of high-degree municipalities with other large centers, which exchange a proportionally high number of commuters: weighted hubs are linking more frequently to same or higher-order municipalities. This phenomenon is illustrated on Figure 8, on the right: large weights connect hubs while the connection with nodes of smaller degree has smaller weights.

5 Discussion: the SMCN and its relation with real environmental systems

In this section, we propose a discussion of the underlying environmental characteristics of Sardinia in the light of the results obtained about the topological and weighted properties of the SMCN.

5.1 Topology and environmental properties

The ranking connected to the degree (Table 1), an index of topological centrality, confirms the existence of administrative and socio-economic hierarchies: among the hubs, Cagliari is the most populated and regional capital town, Nuoro, Oristano and Sassari are province capital towns, Macomer and Quartu Sant'Elena are emerging productive and residential centres. So, as a rule of thumb, the degree k of a town is positively correlated to its size.

Table 1: Ranking of Sardinian municipalities by their degree k.

Rank	Sardinian Municipal Centres	*K*
1	Cagliari	279
2	Nuoro	196
3	Oristano	190
4	Macomer	171
5	Sassari	151
6	Quartu Sant'Elena	140
7	Assemini	139
8	Selargius	127
9	Villacidro	125
10	Ottana	124

While this network is not scale free, the analysis of the cumulative probability distribution of connectivity still reveals the presence of a relatively large number of well-connected towns that attract commuters from a fairly high number of satellite centers.

The behavior of the clustering coefficient versus the degree, we can see that for $k< 50$ the clustering is relatively high, which would be a signature of the fact that small towns (small k) form very well connected clusters of students and workers, while for $k>50$, hub-towns (large k) are poorly inter-connected. We will see however in the next chapter that this

interpretation needs to be revised since the heterogeneity of the traffic is very large and an analysis solely based on topology is necessarily misleading.

The indication of the weighted average degree of the nearest neighbor suggests that the SMCN seems to belong to the class of dissassortative mixed networks. In this case, well-connected (high degree) hub-towns display links preferentially with poorly connected (low degree) small towns. This topological analysis would suggest the existence of inter-municipal commuting districts of small towns, which pivot around a few urban poles and could be interpreted as the signature of a phenomenon often observed in urban and transportation systems, when top functional rank towns attract commuters from small centers, behaving as urban poles: in this environment, citizens of the satellite centers, being interested to regional-level services such as public administration, health and finance, prefer commuting to the pole rather than to a peer level town. The heterogeneity detected in the pattern of municipal centers could be seen as functional to an efficient behavior of the whole network, since each urban hub exchange commuters, and thus goods, services and revenues, with several small size towns. Turning to the geographical displacement of the hub centers, we can note that as expected they are located far away from each other. The distance between Cagliari and Sassari is equal to 210 Km, while from Cagliari and Nuoro to 160 Km, with respect to a total length of the island equal to 230 Km.

5.2 Traffic and environment in the weighted characterization of the SMCN

If all weights were of the same order the information they carry would be irrelevant. The heterogeneity of the weights is as we saw in the previous chapters very high, which implies that the links are not all equivalent and thus forbids any interpretation based on topology only. The analysis of the weighted network introduces thus new insights in the interpretation elaborated on the result of the topological properties under a variety of perspectives.

Table 2: Ranking of pairs of Sardinian municipal centres by the weight of their connections.

Rank	Pairs of Sardinian Connected Municipal Centres	*W*
1	Cagliari-Sassari	13,953
2	Sassari-Olbia	7,246
3	Cagliari-Assemini	4,226
4	Porto Torres-Sassari	3,993
5	Cagliari-Capoterra	3,731

Since the weights are distributed in a very heterogeneous pattern, the corresponding commuters' flows display a wide range of values, which reflects the existence of a territorial hierarchy embedded and supported by the underlined transportation network. Precise

indications emerge about dominant relations between some of the major poles of the island (Table 2): Cagliari, Sassari, Olbia, Assemini, Porto Torres and Capoterra.

In particular, the analysis of the strength probability distribution shows that towns that are detected as topological poles behave also as attractors of traffic flows. The strength is broadly distributed along the nodes, revealing a strong hierarchical dominance of the dynamic hubs over other smaller level centres. The analysis of the relation between strength and degree shows that traffic grows superlinearly with the degree confirming the fact that the more the topological hub-towns are connected and the larger their dynamic centralities (ie. their capacity to attract Sardinian commuting students and workers). The ranking of Sardinian municipalities according to their strength is reported in Table 3.

Table 3: Ranking of Sardinian municipalities by their strength.

Rank	Sardinian Municipalities	*s*
1	Cagliari	64,834
2	Sassari	21,437
3	Quartu Sant'Elena	18,431
4	Oristano	12,130
5	Selargius	10,084
6	Assemini	7,915
7	Porto Torres	6,886
8	Nuoro	6,834
9	Carbonia	6,616
10	Iglesias	6,479

From the behavior obtained for the disparity (Fig. 5), we see that only few connections emerging from the hub-towns carry a large amount of traffic, indicating a strong hierarchy also in the structure of the flows connecting large towns of the SMCN. This mirrors an underlined road network structure, which comprises a very small set of main highways (state roads), a large set of medium-sized roads (provincial roads) and a very large set of local roads (municipal roads). In particular, the link Cagliari-Sassari corresponds to state road n° 131, named "Carlo Felice": this highway crosses longitudinally the island and can be considered a sort of dorsal spine within the SMCN structure to which the lower hiearchical levels transportation structures connect.

The divide detected between weighted and topological clustering coefficients displays evidence for a rich-club-like phenomenon: the hubs tend to aggregate more frequently when their connections convey a large number of commuters. This implies that the SMCN comprehends a few clusters of urban poles, which exchange a lot of commuters. On the opposite side, small towns are mostly connected each other through small traffic

links, which correspond to the underlined second or third order roads. This, again, is a signature of the presence of a strong hierarchy of the commuting relations in this network.

The analysis of the weighted versus topological assortativity shows strong evidences of the hierarchical structure of the SMCN. Thus, it is possible to observe a diffuse trend for towns to exchange commuters with same or higher order centres which is a completely new scenario compared to the one predicted by the topological analysis. In the SMCN, a large number of small-size cities tend to become satellites of higher rank municipalities, shaping an overall network structure widely punctuated with star-like subsystems pivoting around important urban poles connected through dominant roads.

6 Relating topology and traffic to socio-economic phenomena

In this section, we provide a discussion of the emergent topological and traffic properties in relations to the demographic and economic properties of the towns of the SMCN. In the previous sections, we observe that the values of degree k and of the strength s, as indexes of topological and weighted centrality, confirm quantitatively historical structural hierarchies among Sardinian towns. The heterogeneity of the strength and the degree can be visualized as shown in the thematic map representations of Figure 9.

Figure 9 On the left: map of the degree k*; on the right: map of the strength* s *for each Sardinian municipality (Scale: the darker the shade, the higher the value).*

We explore now in quantitative terms the relation between the complex network properties and environmental and economical indicators. One goal is to inspect whether the commuting network hubs correspond to economically relevant towns. As a measure of social and economic centrality of each town, we consider two variables: the total resident population (*pop*) in 1991 (Istat, 1991) and the average monthly income (*mmi*). The average monthly income, an index of urban aggregated wealth, is defined as the product of the average monthly income per worker times the number of workers (Carcangiu et al, 1993). A map representation using these indicators is shown in Figures 9 and 10.

Figure 10 On the left: map of the resident population pop*; on the right: map of the municipal monthly income* mmi *for each Sardinian municipality in 1991 (Scale: the darker the shade, the higher the value).*

In Figure 11, we plot the variable *pop* as a function of the degree *k* and the strength *s* and we observe a clear positive correlation. The association between the population and these variables has a power-law behavior with two different exponents related by a factor 1.9. This is expected since we have already observed that traffic and degree are related by an empirical law $s \sim k^{1.9}$. This implies that if the population has a scaling $pop \sim s^{a}$, the simple substitution of variable yields $pop \sim k^{1.9a}$. This is confirmed in the analysis where the slopes of the two curves are related by a factor close to two. It is interesting to note that the population scales almost linearly with the traffic which suggest that a constant fraction of the population of every city commutes to another town, irrespectively of the size of the city.

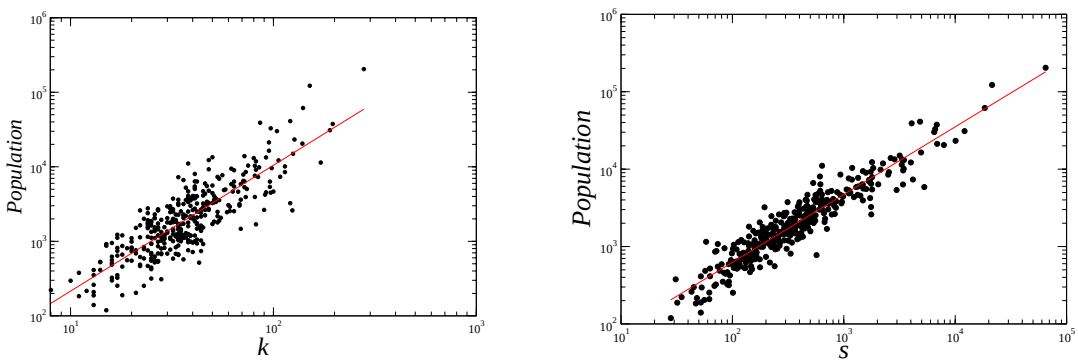

Figure 11 On the left: log log plot of population over the degree k*, slope coefficient equal to 1.70; on the right: log log plot of population over the strength* s*, slope coefficient equal to 0.90.*

In Figure 12, the monthly municipal income (*mmi*) is plotted versus the degree *k* and the strength *s*. A positive correlation is signaled by a linear behavior on a log-log scale which indicates the presence of a power law relation. Also in this case, the exponents of the two relations confirm the general result of a quadratic relation between strength and degree.

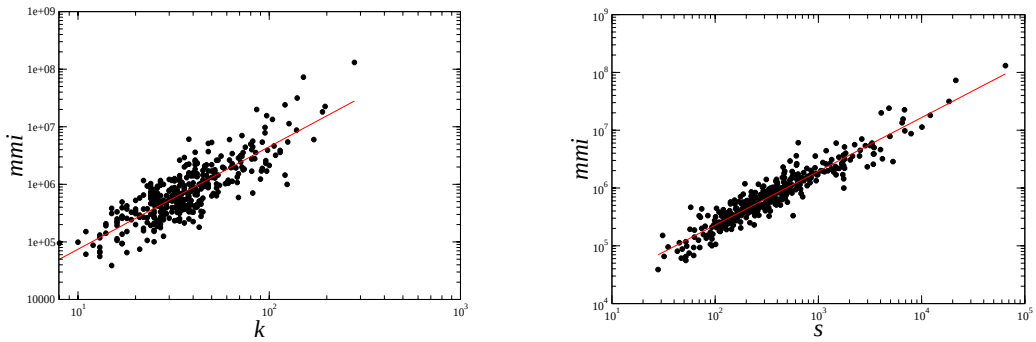

Figure 12 On the left: log log plot of mmi *over the degree* k*, slope coefficient equal to 1.80; on the right: log log plot of* mmi *over the strength* s*, slope coefficient equal to 0.90.*

The results of this section confirm that the centrality measures (strength and degree) display trends consistent with the behavior of social and economic indicators such as population and income. In this sense, the higher the network centrality of a node, the higher is its demographic and economic size. The aforementioned territorial size measures, *pop* and *mmi* grow much faster with the degree *k* than with the strength *s*. We can therefore identify a group of variables related to population (*s*, *pop, mmi*) related by association laws with a behavior very close to the linear one. Their relation with the degree is however highly non-linear indicating that these variables grows much faster than the connectivity of municipalities. Indeed, inverting the relation between degree and traffic (or population) we have $k \sim \sqrt{s}$. The degree of a municipality is both an indicator of the level of economic and social exchange with other municipality and the transport infrastructure. The above relation implies that the economic and transportation infrastructure seem to grow at a pace slower than population indicators. It is not clear if this is a peculiar characteristic of the Sardinia region, due to political and urban planning decision, or a general property of population flows indicating a kind of adaptive equilibrium of a societal system. It is interesting to note that the worldwide airport network exhibits association trends among the network infrastructure and the population size very similar to those observed in this study, suggesting a possible level of universality. However, the detailed study of other commuting and transportation networks is needed in order to gain further insight on this issue.

7 Conclusions

In this paper, the inter-urban commuting system of the island of Sardinia has been considered as a weighted network. This representation allows to characterize quantitatively the attraction on workers and students of each urban center. We compared the different levels of hierarchies emerging in this network with the pattern of demographic and economic poles and, in the light of this framework, we obtained the following results.

At the topological level, the SMCN is relatively small and dense and is similar to a small-world random graph-like network (but is not a scale-free network). The clustering organization of the SMCN similarly to other technological networks, such as the airline network, suggests a hierarchy of nodes where small towns are well connected, while hub towns connect parts of the islands otherwise disconnected. This network seems to belong to the class of disassortative mixed networks: the hub towns are found to connect with a slight tendency to smaller order urban centers.

However, the topology is not enough to describe accurately all aspects of the inter-municipal traffic and we analyzed the weight structure of the network. We found that both the weights and the strengths are very broadly distributed; this property confirms the necessity of including those measures in a realistic description of the SMCN. In particular, the cumulative probability distribution of the strength displays a power-law regime over a wide range of degree values: the SMCN behaves as a scale-free weighted network. In addition, the traffic is strongly correlated to the connectivity and moreover, as shown by the disparity, it is largely concentrated on a very few links. This indicates that the more a town is connected, the much more it attracts commuters and this along a small number of dominant roads. Our analysis of weighted clustering suggests that important towns form a group of well-interconnected centers. It also shows that the SMCN is overall weighted assortative: weighted hub centers are more frequently connected to same or higher order municipalities. In terms of environmental system analysis, a large number of small-size centers are commuting satellites of higher rank municipalities. The SMCN appears as widely punctuated with star-like subsystems pivoting around important urban poles connected preferentially through high traffic roads. In addition, the fact that the weighted clustering coefficient is large for the whole range of connectivity suggests that at a smaller scale there are micro-commuting basins, interpretable as inter-municipal districts of highly clustered small towns, which pivot around a few urban poles belonging to macro commuting basins.

Interestingly enough, there are striking similarities for the weight distribution and the clustering, with the results obtained for urban movements (Chowell et al, 2003): this comparison suggests the existence of basic mechanisms leading to these robust features. In Table 4, the main characteristics of different transportations networks at different scales are shown. In particular, it seems that while the topological properties can vary, the broadness of the traffic distribution is a common feature to all these networks. Another common feature is the existence of correlation between the traffic and the topology.

Table 4 Comparison of properties of transportation networks.

Network	$P(k)$	$P(s)$	s	Y_2
Global: WAN	Heavy Tail	Broad	β=1.5	Θ=1.0
Inter-Cities: SNCM	Fast tail	Broad	β=1.9	Θ=0.4
Intra Urban (Eubank et al, 2003)	Heavy tail	Broad	β=1.0	__

Finally, the existence of hierarchies among towns that we have uncovered with the help of a weighted network analysis is confirmed by comparing demographic and economic urban polarization in Sardinia. Population and wealth, as proxies of the endowment of local resources, display a positive correlation with topological and dynamic centrality.

8 Outlook

The results of this investigation open further questions that are currently under study. First, the indication of network centrality provides a powerful measure of similarity of the towns in attracting workers and students daily. Further studies will address the problem of the detection of communities of workers/students commuting in Sardinia. These results may provide insights for bottom-up processes of decision-making and planning based on the definition of existing emerging behaviors of the citizens. Third, the SMCN is obviously not independent from geography and space. It refers to a physical network, mostly based on road systems, which display precise geographical characteristics and further study, which includes spatial aspects of the network, are currently under progress. These results will hopefully provide decision-makers and planners with relevant indications on the global effect of new infrastructure.

References

Albert R and Barabasi AL, 2002, "Statistical mechanics of complex networks", Rev. Mod. Phys. 74, 47-97.

Amaral LAN, Scala A, Barthélemy M, Stanley HE 2000, "Classes of small-world networks", Proc. Natl. Acad. Sci. (USA) 97, 11149-11152.

Barabási AL, Albert R, Jeong H, 2000, "Scale free characteristics of random networks: The topology of the World Wide Web", Physica A 281, 69-77.

Barrat A, Barthélemy M, Pastor-Satorras R and Vespignani A, 2004, "The architecture of complex weighted networks" Proceedings of The National Academy of Sciences 11, 3747-3752.

Barthélemy M, Barrat A, Pastor-Satorras R and Vespignani V, 2005, "Characterization and modelling of weighted networks" Physica A 346, 34-43.

Batty M, 2001, “Cities as small worlds”, Editorial Environment and Planning B: Planning and Design 28, 637-638.

Batty M, 2003, “Network geography” Paper 63, Centre of Advanced Spatial Analysis (CASA), UCL.

Carcangiu, R, Sistu, G., Usai, S, 1999, “Struttura socio-economica dei comuni della Sardegna. Suggerimenti da un analisi cluster”, Working Papers, 3, CRENOS, Cagliari.

Chowell, G, Hyman, JM, Eubank, S and Castillo-Chavez, C, 2003, “Scaling laws for the movement of people between locations in a large city” Physical Review E 68, 066102.

De Montis A, Barthélémy M, Campagna M, Chessa A and Vespignani V, forthcoming, “Topological and dynamic properties emerging in a real inter-municipal commuting network: perspectives for policy-making and planning”, Congress of the European Regional Science Association, Amsterdam, 23-27 August 2005.

Dorogovtsev SN and Mendes JFF, 2003, “Evolution of networks: From Biological nets to the Internet and WWW” Oxford Univ. Press, Oxford.

Erdös P and Rényi P, 1960, “On the evolution of random graphs” Publ. Math. Inst. Hung. Acad. 5, 17-60.

Gastner MT and Newman MEJ, 2004, “The spatial structure of networks” Cond-mat 0407680.

Gorman SP, 2001, “The network advantage of regions: the case of the USA, Europe and China” paper presented at the International Regional Science Association Meeting, Charleston, SC.

Gorman SP, Kulkarni R, 2004, “Spatial small worlds: new geographic patterns for an information economy” Environment and Planning B: Planning and Design 31, 273-296.

Guimera, R, Mossa, S, Turtschi, A and Amaral, LAN, 2003, “Structure and efficiency of the World-wide Airport network” Cond-mat 0312535.

Jiang B and Claramount C, 2004, “Topological analysis of urban street networks” Environment and Planning B: Planning and Design 31 151-162.

Latora V and Marchiori M, 2000, “Harmony in the small-world” Physica A 285, 539-546.

Latora V and Marchiori M, 2001, “Efficient behaviour of Small-World Networks” Physical Review Letters 87,198701.

Latora V and Marchiori M, 2002, “Is the Boston subway a small-world network?” Physica A 314, 109-113.

Latora V, and Marchiori M, 2003 “Economic small-world behavior in weighted networks” The European Physical Journal B 32, 249-263.

Pastor-Satorras R and Vespignani A, 2004, “Evolution and Structure of the Internet” Cambridge University Press, Cambridge, USA.

Salingaros NA, 2001, “Remarks on a city’s composition” Resource for Urban Design Information.

Salingaros NA, 2003, “Connecting the Fractal City” Keynote speech, 5th Biennial of town planners in Europe, Barcelona.

Schintler, AL, Gorman SP, Reggiani A, Patuelli R, Gillespie A, Nijkamp P and Rutherford J, forthcoming, “Scale-free phenomena in telecommunications networks” Networks and Spatial Economics.

Sen P, Dasgupta S, Chatterjee A, Sreeram PA, Mukherjee G and Manna SS, 2003, “Small World properties of the Indian Railway network” Phys. Rev. E 67, 036106.

Shiode N and Batty M, 2000, “Power law distribution in real and virtual worlds” Paper 19, Centre of Advanced Spatial Analysis (CASA), UCL.

Watts DJ and Strogatz SH, 1998, “Collective dynamics of ‘small-world’ networks” Nature, 393, 440-442.

Yook SH, Jeong H, and Barabási, AL, 2002, “Modeling the Internet's large-scale topology” Proc. Natl. Acad. Sci. (USA), 99, 13382-13386.