Grouping Complex Systems
a Weighted Network Comparative Analysis

Andrea De Montis(1) *, Alessandro Chessa(2), Michele Campagna(3),
Simone Caschili(3), Giancarlo Deplano(3)

(1) *Dipartimento di Ingegneria del Territorio, Sezione Costruzioni e Infrastrutture, Università degli Studi di Sassari, via De Nicola, Sassari 07100 - Italy*
(2) *Dipartimento di Fisica, INFM, Università degli Studi di Cagliari, Complesso Universitario di Monserrato, Monserrato 09042 - Italy*
(3) *Dipartimento di Ingegneria del Territorio, Università degli Studi di Cagliari,Piazza d'Armi 16, Cagliari 09123 - Italy*

Abstract

In this study, the authors compare two inter-municipal commuting networks (MCN) pertaining to the Italian islands of Sardinia and Sicily, by approaching their characterization through a weighted network analysis. They develop on the results obtained for the MCN of Sardinia (De Montis et al. 2007) and attempt to use network analysis as a mean of detection of similarities or dissimilarities between the systems at hand.

Key-words: complex networks, commuters dynamics, weighted networks, spatial networks, GIS

---

* Corresponding author.
E-mail address: andreadm@uniss.it

1 Introduction

Complex network theory (CNT) dates back its origin to the elaboration of graph theory by Euler[1] and in the 1960's has been recalled and renewed by Erdős and Rényi (1959, 1960), who in seminal works studied the structure of random graphs, networks where a couple of nodes whatsoever has a probability p ranging from 0 to 1 to be associated a connection (edge). In this case, the probability distribution of the degree k of a node (the number of its connections to the first neighbours) display a bell shape and is featured by a finite and characteristic mean value.

At the end of last millennium, the availability of even larger data sets and the parallel explosion of computer processing power has enabled a systematic and intensive application of CNT to the study of very large networks (Pastor Satorras and Vespignani, 2004; Albert and Barabàsi, 2002). A major property of these networks has been found by Watts and Strogatz (1998) and termed small world effect: the average shortest path length *l* between a pair of nodes whatsoever has been found to scale very slowly with the number of nodes N ($l \sim \log N$). This is a signature of a relatively high cohesion of small world networks, that are in fact systems able to organize themselves and resist to attacks and brake-downs (Amaral et al, 2000). This cohesion is provided by on average an appreciable probability to find high degree nodes, the hubs of the system, providing a relevant number of shortcuts in the network. Another important discovery has been that sometimes in these systems the probability distribution of the degree k does not provide any characteristic value in presence of a diverging measure of statistical fluctuations (variance). In this case, a scale free behaviour has been found through the detection of a power law trend of the probability distribution of the degree k. These networks are said to be scale free, since they show invariant statistical properties over the entire range of degree values. Moreover, the presence of heavy tails in the probability distribution curve often is a signature of a non null probability to encounter very large degree nodes, the "hubs" of these networks. Another important property of these network is described by their growth mechanism that obey to a preferential attachment (Barabàsi and Albert, 1999) attitude of new nodes. As new comers, they tend to grasp the highest advantage from the system they join to by linking to nodes that have a very large degree value, again the "hubs" of the network.

The last research development of CNT is constituted by studies directed to inspect the dynamics of a network, by analysing the patterns of distribution of a quantity (the weight) attributed to each edge (Barrat et al, 2004; Barthélemy et al, 2005). In this approaches, authors adopt a generalization

[1] One of the greatest mathematicians ever: applying graph theory for the first time in its Solutio problematis ad geometriam situs pertinentis (1736) he demonstrated that it is impossible to complete a leisure walk of the city of Königsberg by crossing its seven bridges only once.

of the series of network measures developed to study pure topological characteristics. Moreover, the inspection of the probability distribution of the strength, the generalization of the degree k summing the weights attached to the edges converging to a given node, has often revealed a fairly goodness of fit to a power-law line, a clear signature of a sort of dynamic scale free behaviour of the network. Other relevant results come from the analysis of the interplay between dynamic and topological properties, which often leads to uncover super proportional correlations, a sign of the fact that the informational/traffic properties of a complex network can not be fully explained by the sole contribution of its topology.

Complex Network Theory has been applied to a number of both simulated and real systems. In the case of the latter, CNT provides with insights into a wide range of questions regarding food webs, human interactions, the Internet, the world wide web, the spread of diseases, population genetics, genomics and proteomics. In each of these cases, it is possible to start by inspecting recurrent structures embedded in complex systems characterized by not identical elements (the nodes) connected through different kinds of interactions (the edges). For a review of these applications, see Albert and Barabàsi (2002) and Newman (2003).
Recently also in many fields grouped under the realm of regional science, a number of scholars have begun applying the paradigm of complex network analysis for modeling urban (Batty, 2001; Jiang and Claramunt, 2004), regional and socio-economic systems (Latora and Marchiori, 2003; Schintler et al, 2005). Such network models can be conceived of as interlaced compositions of individual entities (the nodes) and their multiple interactions (the links). Many authors attempt to extend the analysis beyond the consideration of topology and dynamics by inspecting the influence of geographical space onto the network properties (Gorman and Kulkarni, 2004; Gastner and Newman, 2004; Crucitti et al, 2006; Campagna et al, 2007). With a specific interest for this paper, a number of applications refers to the study of infrastructures and of commuters' complex behaviour (Guimera et al, 2003; Latora and Marchiori, 2002; Chowell et al, 2003; Sen et al, 2003). These works are often developed on the assumption that the emergence of scale free properties is a signature of efficiency in the system general behaviour. Examples are the hub-and-spoke structure invoked for transportation systems by O'Kelly and, in particular, for airline networks by Reggiani et al (in press). In the field of the analysis of commuters' behaviour, De Montis et al (2007) have developed a weighted network approach to the system of inter-municipal habitual movements of the inhabitants of the Italian region of Sardinia, the second largest island of the Mediterranean.

In the light of these background remarks, in this paper the authors aim at extending the work developed by De Montis et al (2007) to the inspection of the network properties of the other major Italian region of Sicily, again an island, yet very close to the mainland. Starting from the assumption that geographical similarities should lead to the emergence of ubiquitous statistical properties and vice versa, the authors develop a comparative analysis of the two insular inter-municipal commuters' systems. The arguments are reported as follows. In the next section, the main methodological aspects and the most relevant results obtained in the study of commuting behaviour of Sardinia are reported. The latter constitutes the starting hint of the comparison developed with respect to the island of Sicily. In section three, analyses are developed for the Sicilian inter-Municipal Commuting Network (SiMCN) for both the topological and the weighted network characteristics. In this section, an analysis of the interplay between dynamic, demographic and topologic properties of the SiMCN is also represented. In section four, the latter resulting statistics are confronted to the correspondent results obtained for the case of Sardinia; in this comparative framework, the authors attempt at commenting the emergence of geographical similarities or dissimilarities between the two island. In section five, the overall conclusions and outlook remarks of this paper are proposed.

2 Recalling the properties of SMCN

This study is inspired by the analyses worked out by De Montis et al. (2007) about the interplay between traffic dynamics and socio-economics in the inter-municipal commuting system of the Island of Sardinia, Italy. In that study, the authors inspected workers and students daily movements, by adopting a network representation, the Sardinian inter-municipal commuting network (SMCN) characterized by N=375 vertices, each one corresponding to a town, and E=8124 edges, each one representing the existence of commuters exchange between two extreme towns.

The authors have developed a number of measures on this undirected graph and on the corresponding representation obtained by featuring each link with a weight explaining how many commuters flow through that connection. This weighted undirected network is actually built by processing information conveyed in the regional origin destination table, a dataset that reports the daily work and study-led movements among Sardinian municipalities. In the remaining, we report the most important results obtained in that study.

The topological analyses show that the SMCN belongs to the classes of small-world random networks, since the cumulative distribution of the degree k displays a bell-shape behaviour. The study of the clustering coefficient reveals a divergence from usual random network behaviour and uncovers properties common in other technological networks. In particular, small (with small k) municipalities are locally densely interconnected, while large municipalities (even hubs of the network) provide a large set of connections for remote regions otherwise disconnected. This evidence is confirmed also by the analysis of the average degree of the nearest neighbours, which signals a disassortative mixed behaviour: the hub towns are preferentially connected to small degree (less central) municipalities acting as star-like vertices of the SMCN.

On the side of the analysis of the weighted network, the study finds that the probability distributions of both weights and strengths (total commuter traffic handled by the municipalities) display a power-law regime over a wide

spectrum of degree values. In this case, no characteristic value of the distribution is found and the SMCN can be included in the class of scale-free weighted networks.
The spectrum of the strength averaged over the values of the degrees reveals a super linear behaviour implying that the higher the number of connections to a town the much larger the traffic per connection handled. This means that most likely there are hidden properties that control and describe the behaviour of the network. The inspection of the disparity of a node, which measures eventual inequalities in the distribution of the traffic flow among the connections of each node, confirms the actual structure of the real network: a fairly large amount of commuters are exchanged between hub towns through a very small number of backbone connections constituting the dorsal highways of that systems.

3 Inspecting the properties of SiMCN

While in the previous section the properties found for the SMCN have been briefed, in this section a network analysis is developed to inspect topologic and dynamic characteristics of the Sicilian inter-Municipal Commuting Network (SiMCN).

3.1 Setting the case study and dataset

Sicily is an administrative region of Italy, which belongs to the macro-region also known as “Mezzogiorno”. It extends for 25707 squared kilometres and is the largest Mediterranean island; during the ages of the Greek colonization for its similarity to a triangle this island was named “Trinacria”. Sicily hosts a resident population that amounts to 5 millions inhabitants; comparing to Sardinia, this island is on average about three times as much densely populated. While Sardinia displays a neat geographical separation from the mainland, Sicily by contrast is divided from the coast of the Italian region of Calabria by the strait of Messina, which has a minimum width equal to about 3 Kilometres.

Figure 1 A global view of Italy. The two islands are marked in dark gray

As a matter of fact, the main transportation system still currently adopted consists of a system of frequent ferries able to carry persons, cars and trains from a coast to another. The exceptional geographical settings and the need to benefit of a stabile and definitive connection between the island and the mainland have motivated so far a great interest of researchers, professionals, scholars and recently also of central government officials. Many projects have been studied in the last fifty years; a project for a single span bridge has been proposed and almost definitely approved during the last legislation. Yet, the new government elected in 2006 has promoted the development of further studies to assess the impact of this relevant construction over the environment.

A broad description of commuters' behaviour is provided by the Origin Destination Table (ODT), a census dataset that describes work and study led movements of inhabitants from the residential origin urban area to the most habitual destination town (Istat, 1991). Commuters' movement are also described with reference to travel time period and means of transportation. Following the outline of De Montis et al (2007), the ODT provides the main indication for constructing the adjacency matrix [A], the standard mathematical representation of a network, whose general term a_{ij} is equal to 1, if at least one person commutes from the origin town i to the destination town j, and is equal to zero otherwise. In this case, [A] has null diagonal elements (a_{ii}=0), since movements of commuters within the same municipality are not considered. In this analysis, commuters' movements are inspected regardless the means adopted for transportation.
Even though Sicilian geographical settings would suggest an extension of the commuter network at least to the continental municipal centre of Reggio Calabria, this analysis refers just to commuters movements among Sicilian municipalities. This is motivated by the evidence that the contribution of Reggio Calabria to the total traffic of the SiMCN is equal to 805 commuters exchanged with Sicilian towns, a negligible figure with respect to the whole range of traffic values.

3.2 Topological properties of the SiMCN: analysis and interpretation

As introduced in the previous section, in the following the system of commuters' movements among "Sicilian" municipalities will be analysed as a network, with a focus first for its topological properties.
The SiMCN is constituted by N=391 nodes, which correspond to the set of Sicilian municipalities, and by E=9993 edges, which correspond to the pattern of commuters' exchanges among those towns.
The first topological measure adopted in this study is the degree k of a given node i, which has this expression

$$k_i = \sum_{j \in V(i)} a_{ij} \tag{1}$$

where $V(i)$ denotes the set of neighbors of i.
The analysis of the probability distribution of the degree k provides the reader with a proxy indication of the centrality of the nodes, in terms of number of first neighbours connected to each node. In table 1, a ranking of the tenth most central Sicilian towns is reported. In the highest positions the main towns of the island are located.

Table 1 Ranking of Sicilian municipalities by their degree

Rank	Municipal centres	Degree k
1	Palermo	280
2	Catania	228
3	Messina	220
4	Caltanisetta	154
5	Enna	132
6	Termini Immerese	132
7	Bagheria	123
8	Giarre	122
9	Gela	120
10	Milazzo	119

The analysis of the cumulative probability distribution of the degree P(k) of the nodes (figure 2) reveals the emergence of an exponential line fit with a fast decaying tail and a characteristic mean value at $<k>=51$, while the maximum value of k is equal to 280. The latter is a signature of a random graph behaviour of the SiMCN. The low values of the average ($<l>=1.98$) and maximum ($l_{max}=4$) path length between a pair whatsoever of nodes points out that in the SiMCN can be considered a small world random network, under a strict topological point of view.

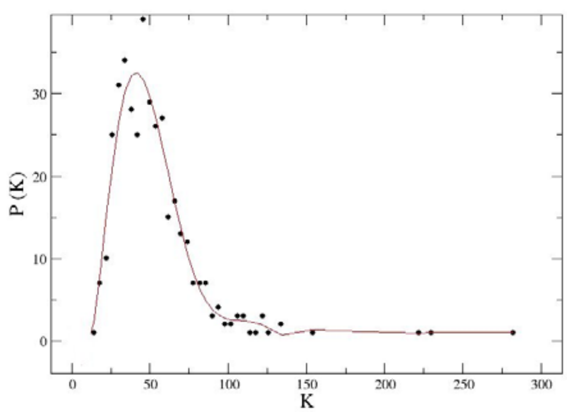

Figure 2 Plot of the probability distribution of the degree k for the SiMCN. The line fit is an exponential curve. This behaviour signals the emergence of a random graph structure.

Another relevant quantity is the clustering coefficient, a measure of the level of local cohesiveness of a node that obeys to the following relation:

$$C(i) = \frac{2E(i)}{k_i(k_i - 1)} \quad (2)$$

where $E(i)$ is the number of links between the k_i neighbors of the node *i* and $k_i(k_i - 1)/2$ is the maximum number of possible interconnections among the neighbors of the node. The clustering coefficient ranges in the interval [0,1]: values close to 1 are a signature of a very high local connectedness around a node, while the opposite is valid for values approaching to zero.
It is often preferable to consider an averaged measure of the clustering coefficient $C(i)$ for all nodes with a given k value, by managing the following spectrum of the clustering coefficient versus the degree:

$$C(k) = \frac{1}{NP(k)} \sum_{i/k_i = k} C(i) \quad (3)$$

where $NP(k)$ is the total number of nodes of degree k. Figure 3 refers to the latter spectrum revealing a downward sloping trend of C(k) over the whole range of degree values.
This is a signature of a common property in transportation systems: the hubs of this network have first neighbour nodes that on average are not connected each other, while small degree nodes' first neighbours are much more interconnected. This implies that large hub towns are linked to many satellite towns that are disconnected each other, while the latter small towns display fewer neighbours by contrast very often connected each other. The origin of this pattern can be explained by the search of efficiency in the dynamics of the entire network: commuters often prefer to commute from satellite centers to hub towns seeking higher lever services provided in the main towns.

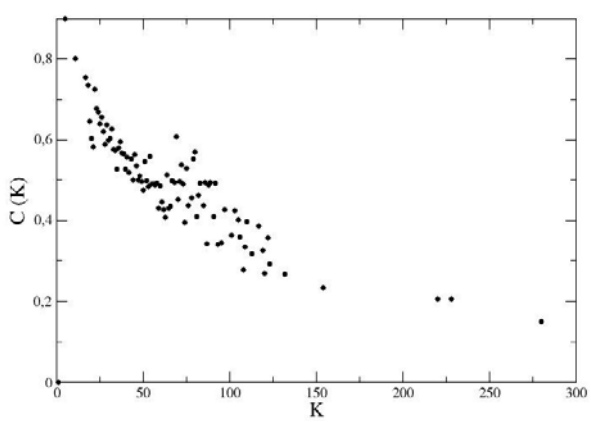

Figure 3 Plot of the clustering coefficient versus the degree k for the SiMCN. It is possible to detect a downward sloping behaviour common in many transportation systems.

This pictured can be confronted with the insights provided by means of the analysis of the degree similarity of the neighbors of a node, which is measured by the average degree of the nearest neighbors of a given node *i*:

$$k_{nn}(i) = \frac{1}{k_i} \sum_{j \in V(i)} k_j \qquad (4).$$

where *V*(*i*) denotes the set of neighbors of *i*. Analogously to the clustering coefficient, it is possible to average the assortativity over nodes with a given degree leading to

$$k_{nn}(k) = \frac{1}{NP(k)} \sum_{i / k_i = k} k_{nn}(i) \qquad (5).$$

where *NP*(*k*) is the total number of nodes of degree *k*. This spectrum measures the tendency of vertices to be connected with vertices with the same degree properties. An increasing trend in this spectrum is the signature of a phenomenon called assortativity: nodes on average tend to connect to other nodes with a equal or higher degree value. The opposite behavior is typical of disassortative mixed networks.
In the case of the SiMCN, the spectrum reported in figure 4 displays a downward sloping trend with a very low coefficient and the network can be very hardly classified as a disassortative mixed network. The latter corresponds to a very diffused property in transportation systems, whose hubs (high degree nodes) are designed to connect preferentially to low degree nodes in order to achieve an efficiency as high as possible.

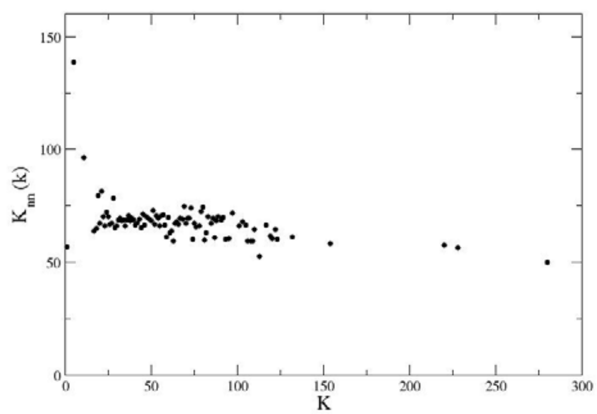

Figure 4 Spectrum of the average degree of the neighbours versus the degree for the SiMCN. It is possible to detect a very weak tendency of the SiMCN to behave as a disassortative mixed network.

In order to inspect with a finer approach the level of centrality of a node (town), the next topological quantity considered is the betweenness centrality (BC) (Freeman, 1977). This variable characterizes the inter-centrality of nodes (or edges) in large networks and is defined as the total fraction of shortest path going through a given node *i*. More precisely the BC is given by

$$g(i) = \sum_{s \neq t} \frac{\sigma_{st}(i)}{\sigma_{st}} \quad (6)$$

where $\sigma_{st}(i)$ is the number of shortest path going from node *s* to node *t* and passing through node *i* and σ_{st} is the number of shortest paths from *s* to *t*. The higher the value of BC, the higher the inter-centrality of the nodes. In other terms, high BC nodes are part of more shortest paths within than less important nodes and act as "bridges" between different parts of the network.

As figure 5 shows, there is a positive correlation between BC and the degree *k* signalling in a sense that, in this case, BC does not add any extra information about the centrality of a node with respect to the indication already conveyed in the degree *k*. The latter is confirmed also by the ranking of the ten top scorer towns according to their BC reported in table 2.

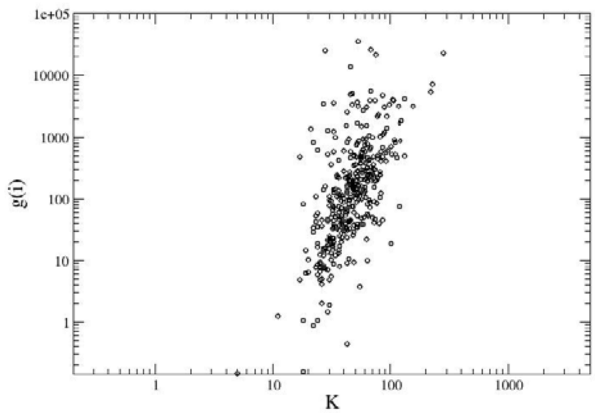

Figure 5 Plot of the betweeness centrality BC versus the degree for the SiMCN. A positive correlation is clearly detected.

Table 2 Ranking of the ten top scorer towns in the SiMCN for inter-centrality values.

Rank	Municipalities centers	Betweeness Centrality
1	Alimena	35377
2	Motta Sant'Anastasia	26401
3	Busetto Palizzolo	25348
4	Palermo	22824
5	Augusta	21314
6	Castellamare del Golfo	13939
7	Catania	7078
8	Corleone	5585
9	Messina	5354
10	Raccuja	5085

3.3 *Weighted network approach to the SiMCN: analysis and interpretation*

The dataset adopted for this investigation, the Origin Destination Table (ODT), conveys information related to the number of commuters that preferentially move from residential origin municipal towns to habitual municipal destinations. Therefore, this information can be adopted as a base to construct the weighted adjacency matrix [*W*] of the SiMCN, where a generic rectangular element w_{ij} is equal to the sum of the number of commuters moving from the town *i* to the town *j* and vice versa, and a generic quadratic element w_{ii} is equal to zero. In this case, the symmetric weighted adjacency matrix [*W*] stands as the standard mathematical representation of the SiMCN now conceived of as an undirected weighted network.

The analysis of the probability distribution of the weights *P*(*w*) reveals a large heterogeneity of values: w_{max} is equal to 10233, three orders higher than its average value $<w>=37.6$. In figure 6, the trend of the cumulative probability distribution of the weights is pictured revealing a power law behaviour over a wide range of weight values ($P(w) \approx w^{-\beta}$, with exponent $\beta = 2.14$).

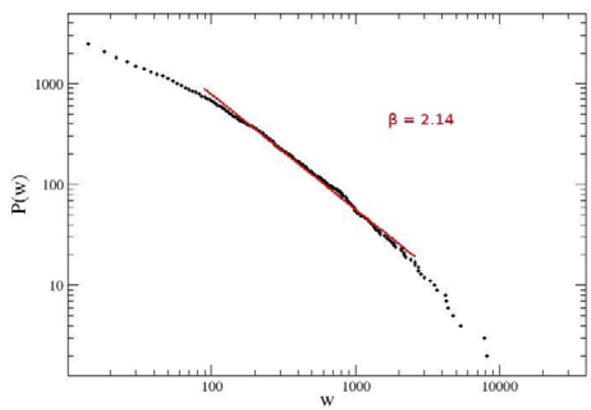

Figure 6 Log log plot of the cumulative probability distribution of the weights. A power law regime emerges over a wide range of w values with a slope exponent equal to 2.14.

The distribution of the weights characterizes the SiMCN, since the relevant connections can be analysed among the main towns of the island. In table 3, the five most important inter-municipal “dorsal links” are ranked by the correspondent weight.

Table 3 The five top scorer connections between municipalities in the SiMCN ranked according to their weight.

Rank	Pairs of connected municipal centers	Weight *w*
1	Misterbianco - Catania	10233
2	Erice - Trapani	8127
3	Gravina Di Catania - Trapani	7857
4	Tremestri Etneo - Trapani	5400
5	Trapani - Aci Castello	4724

In this weighted network approach to the analysis of the SiMCN, it is useful to adopt a centrality measure obtained through the strength *s*, by generalising the degree *k*, defined with the relation

$$s_i = \sum_{j \in v(i)} w_{ij} \qquad (7)$$

The strength offers another proxy indication for the centrality of a node in a network. In this case, the strength can be interpreted as a measure of the capacity of a town to attract commuters from first neighbour municipalities. In table 4, a ranking of the ten most "attractive" towns is reported.

Table 4 Ranking of municipalities in the SiMCN by their value of strength.

Rank	Municipalities centers	Strength *s*
1	Catania	80326
2	Palermo	49500
3	Siracusa	17800
4	Messina	16987
5	Trapani	16578
6	Misterbianco	13801
7	Agrigento	13323
8	Gravina di Catania	10644
9	Erice	9794
10	San Giovanni La Punta	9099

In figure 7, the trend of the cumulative probability distribution of the strength is described: it is possible to appreciate a power law behaviour ($P(s) \approx s^{-\gamma}$, with exponent $\gamma = 2.1$) and a relatively heavy tail of the curve, which are a signature of the emergence of a not negligible number of hubs in the SiMCN, when conceived of as a weighted network. In this case, no characteristic value can be indicated for the probability distribution and the slope exponent is a quantitative indication of the level of heterogeneity in the system.

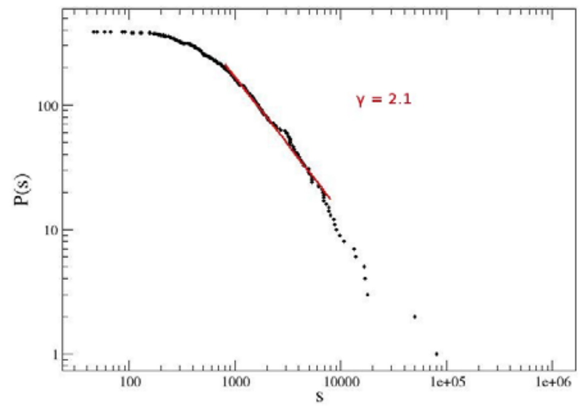

Figure 7 Log log plot of the cumulative probability distribution of the stength P(s). The line fits a power law trend with slope exponent equal to 2.1.

These results open a novel perspective to the analysis developed so far. While the SiMCN can be classified, in the light of a pure topological approach, as a small world random network, under the lenses of the analysis of the commuters' dynamics, the SiMCN seems to belong to the class of weighted scale free networks.

An important property of weighted networks is the patter of flows distribution among the different connection to a node. In the case at hand, this quantity is able to evaluate the existence of dominant connections among those directed to a town. According to Barthélemy et al (2005), such a measure can be appreciated though the disparity, defined as

$$Y_2(i) = \sum_j \left(\frac{w_{ij}}{s_i} \right)^2 \qquad (8).$$

For the case of the hubs (k>>1), this quantity allows distinguishing situations where the weights are of the same order ($Y_2\ 1/k$ <<1) from situations where a few connections dominate and the rest of the traffic is distributed among the remaining connections (Y_2 is of order 1/*n* and *n*<<*k*). Figure 8 refers to the analysis of the spectrum $Y_2(k)$, a quantity obtained by averaging the disparity of the nodes for each of the degree *k* values. It is possible to observe an average behaviour of the form $kY_2 \approx k^{1-\vartheta}$, with $\vartheta = 0.31$. In this situation, it is possible to observe the emergence of a small number of connections that carry a very large amount of traffic to hubs, while a negligible amount of commuters' flow is distributed among the many remaining minor order connections. In other words, in the SiMCN inhabitants commute preferentially to hub towns travelling from a few towns, while a contained number of commuters move habitually from a large number of the remaining towns.

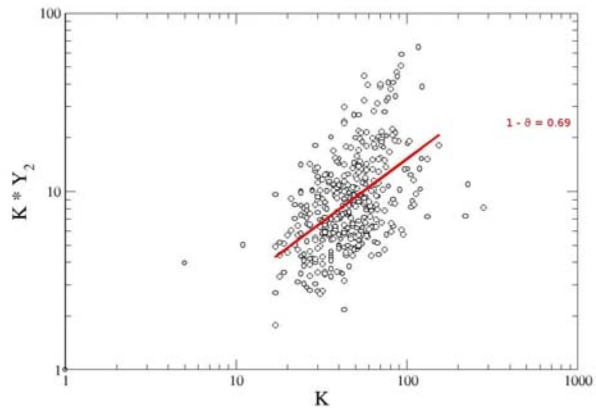

Figure 8 Log log plot of the spectrum of the disparity versus the degree. The average behavior is a signature of the emergence of a small number of connections that dominate in the SiMCN.

3.4 Analysis of the interplay between dynamics and topology

In this section an analysis of the interplay between dynamic, demographic and topological properties of the SIMCN is developed.
In order to inspect the relation between the strength of nodes and their degree, in figure 9 the spectrum of the average value of *s* for each degree *k* of the nodes is reported.

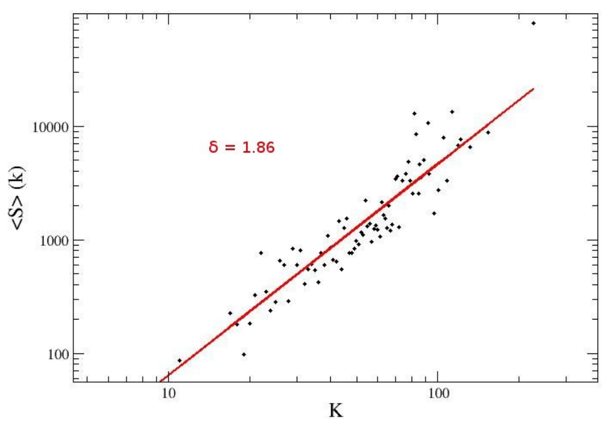

Figure 9 Log log plot of the spectrum of the strength s versus the degree k. A super linear correlation is evident with a slope exponent $\delta = 1.86$*.*

It is possible to observe a positive correlation between these two quantities and a power law regime over the whole range of degree values ($s(k) \approx k^{\delta}$, with an exponent $\delta = 1.86$). This implies that the strength *s* of a given node, on average, scales twice as fast as its degree *k*: the higher the degree of a node (the number of first neighbour towns whose residents commute to the town at hand) the much higher the strength (corresponding to the total amount of commuters' traffic attracted). This evidence stands as a confirmation of a similar phenomenon occurring also in the Sardinian inter-Municipal Commuting Network (SMCN). In both cases, the traffic per connection increases when the number of connections (degree *k*) increases: this super-linear behaviour leads to the need to hypothesize the existence of some hidden economies of scale.

The choice of taking into account the pattern of the weights over the network implies a mathematical update on the measures of local structural properties provided in section 3.2 by the clustering coefficient and by the average degree of the neighbours.
A modification of the clustering coefficient able to take into account the distribution of the weights on the relevant edge of the local systems of cliques has been proposed by Barrat et al (2004) with the expression

$$C^W(i) = \frac{1}{s_i(k_i - 1)} \sum_{j,h} \frac{(w_{ij} + w_{ih})}{2} a_{ij} a_{ih} a_{jh} \qquad (9),$$

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]. As for the topological case it is possible to average over all nodes of same degree *k* obtaining the clustering spectrum $C^W(k)$.
This weighted clustering coefficient 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, it is possible to consider not only the number of closed triangles but also their total relative weight with respect to the vertex' strength. In the case of random networks $C^W = C$; by contrast in real weighted networks, two different situations may arise. 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; while if $C^W > C$, the network studied is on average constituted by interconnected triples more likely formed by edges with larger weight.

As figure 10 shows, while the spectrum of the weighted clustering coefficient with respect to the degree *k* displays a slightly downward sloping trend, the relative difference between the weighted and the topologic clustering coefficient, despite the inevitable fluctuations, increases with respect to the same quantity. This is a clear signature that the SiMCN is constituted by interconnected triples where the vertices are preferentially connected through edges whose weight is relatively large. This results lead to hypothesize the emergence of a rich club phenomenon: hub towns tend to have a high level of local cohesiveness especially when in their sub systems a large traffic flows through the connections. The latter evidence in part contradicts and enriches the information provided by the inspection of the sole topologic clustering coefficient. Regarding this property, the SiMCN is very similar to the SMCN.

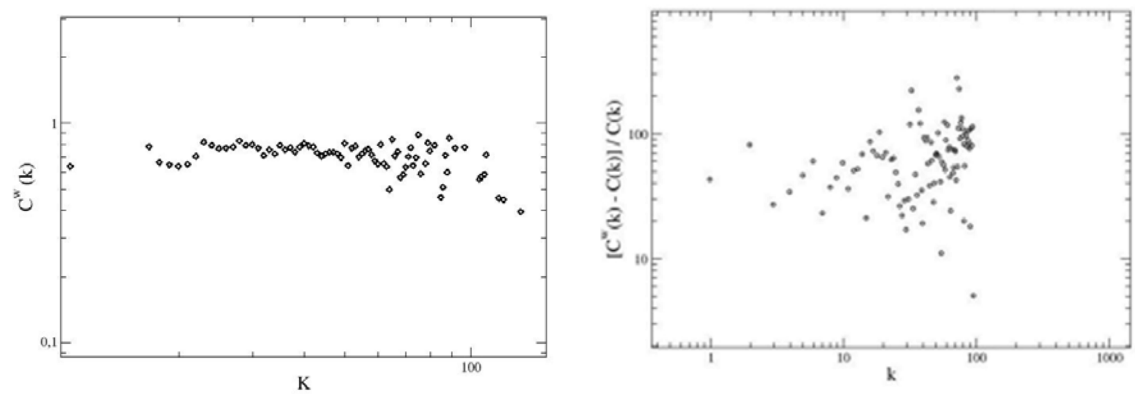

Figure 10 On the left: the log log plot of the weighted clustering coefficient versus the degree k displays a slight downward sloping behaviour. On the right: the log log plot of the relative difference between weighted and topologic clustering coefficients shows an upward sloping trend.

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

$$k^{w}{}_{nn}(i) = \frac{1}{s_i} \sum_{j \in V(i)} w_{ij} k_j \qquad (9).$$

The latter refers to a local weighted average of the nearest neighbor degree according to the normalized weight of the connecting edges. As usual, it is also possible to average this quantity over the nodes with a given degree k value obtaining the spectrum $k^{w}_{nn}(k)$.

The inspection of the difference of the weighted and un-weighted measures implies that if $k^{w}{}_{nn} > k_{nn}$ the edges with the larger weights are pointing to the neighbors with larger degree, while the opposite holds if $k^{w}{}_{nn} < k_{nn}$. The $k^{w}{}_{nn}$ thus measures effective affinity to connect with high or low degree neighbors according to the magnitude of the actual interactions.

Figure 11 reports on this analysis, by indicating that both the weighted average degree of the neighbours and its differential expression clearly increase with respect to the degree *k*. This implies that taking into account the distribution of the weights leads to a completely different representation of the local structure of the network. The SiMCN is constituted by nodes that are preferentially connected to nodes that have at least the same degree k: thus, this systems can be classified as a assortative mixed weighted network.

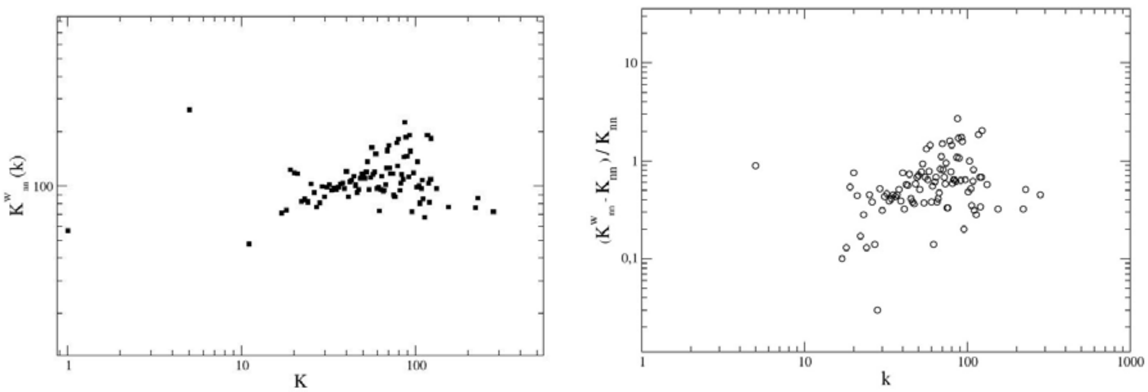

Figure 11 On the left: the log log plot of the weighted average degree of the neighbours versus the degree k displays an upward sloping trend. On the right: the log log plot of the relative difference between weighted and topologic average degree of the neighbours shows an upward sloping behaviour. The SiMCN belongs to the class of assortative weighted networks.

3.5 Relating topology and traffic to socio-economic phenomena

The degree k and the strength s provide the analyst with first indications on the level of centrality of each municipal centre within the SiMCN (i.e. of each vertex of the network). In this sense, the performance of each town increases when its relations increase.
It is interesting to compare the latter network centrality indexes with one of the main indicator of the endowment of local resources, which is the number of resident inhabitants in the municipal districts of Sicily (Istat, 2001). This can be considered indeed as an indicator of demographic centrality of each Sicilian town.
Geographic analyses of the quantities above are reported in figure 12.

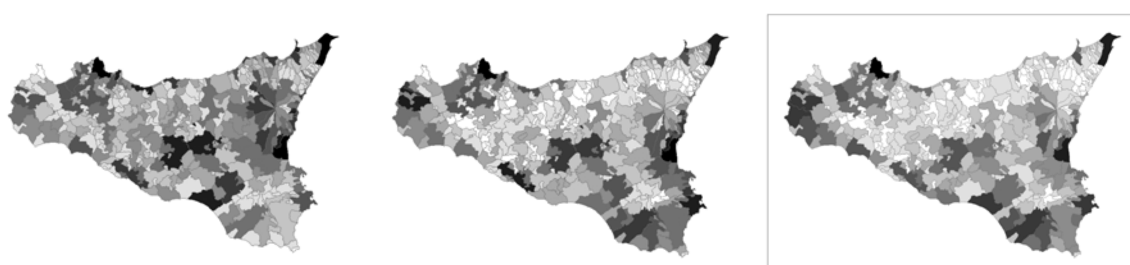

Figure 12 Analysis of the geographical distribution among municipal centres of Sicily of the degree k (on the left), of the strength s (on the centre), and of the resident population pop (on the right). Scale: the darker the shade, the higher the value.

As figure 13 demonstrates, population is found to scale proportionally both to the degree *k* and the strength *s*. While in the latter case the value of the slope coefficient signs an almost neutral proportionality between pop and s, in the other case population increases super linearly with respect to the degree k. These is of course due to the correlation between s and k detected previously.

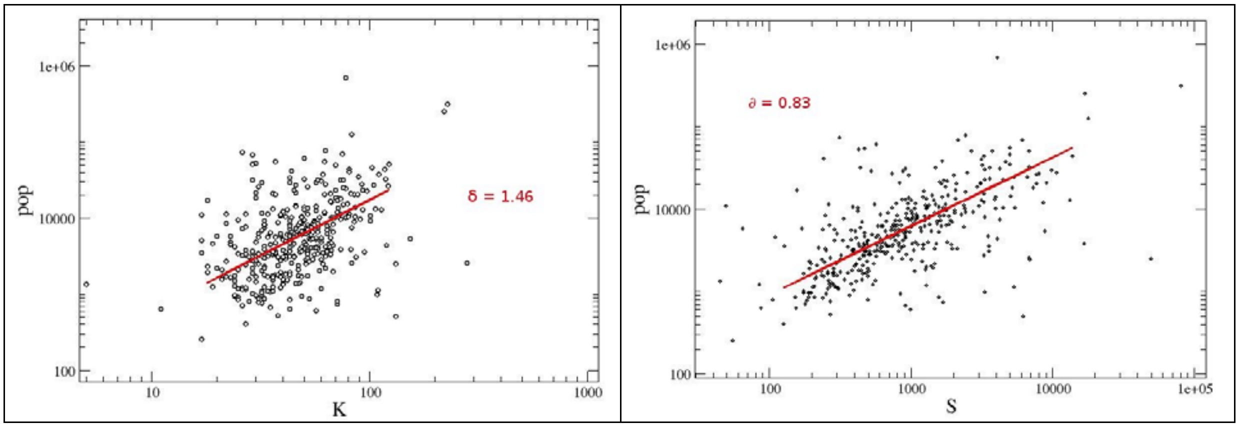

Figure 13 On the left: log log plot of population over the degree k, slope coefficient equal to 1.46. On the right: log log plot of population over the strength s, slope coefficient equal to 0.83.

This result, which has in part been detected also for the SMCN, implies that in the SiMCN the higher the topological centrality of a town the much higher its stabile population.

4 Discussion: comparing SMCN and SiMCN

In the latter section, the SiMCN has been analysed and its most relevant statistical properties have been reported. Also a brief interpretation of those results has been proposed, as it introduces to the detection of emerging phenomena that characterize that commuting system. In many cases, similar phenomena have been found to characterize also the corresponding Sardinian commuting network (SMCN).
In order to inspect in more details how far it is possible to state that Sardinia and Sicily have similar commuting networks, in this section a synthetic comparative framework is constructed and commented.
With respect to the topological properties summarized in table 5 and 6, it is possible to classify both networks as small worlds, since they display a very low value of the average path length $<l>$ comparing to the number of nodes. This implies that information is able to flow very efficiently in those networks and by contrast also epidemic spread may be favoured in the same system, conceived as the site of social interaction.

Table 5 Comparative overview of topological properties of the SMCN and SiMCN

	N	E	k_{min}	k_{max}	$<k>$	$<l>$	l_{max}	$<C>$
SMNC	375	8124	8	279	40	2.0	3	0.26
SiMCN	391	9993	1	280	51	1.98	4	0.52

The behaviour of the distribution $P(k)$ displays a bell shape both for Sardinia and for Sicily with a characteristic and defined mean value: this leads to conclude that those networks belong to the class of random graphs. A similar trend of the clustering coefficient confirms a common property of the local structure in many infrastructure networks: hub towns tend to connect otherwise disconnected regions, while small degree k towns are locally very densely connected. The average degree of the neighbours for both networks has a spectrum with a downward sloping behaviour: the SMCN and the SiMCN belongs to the class of disassortative mixed networks, where hub towns preferentially exchange commuters with lower degree towns.

Table 6 Comparative overview of topological properties of the SMCN and SiMCN, part 2.

	$P(k)$	$C(k)$	$K_{nn}(k)$
SMNC	Exponential behaviour	Downward sloping	Disassortative mixed
SiMNC	Exponential behaviour	Downward sloping	Slightly disassortative mixed

With respect to the traffic properties outlined in table 7, it is possible to see that the weights in both cases are broadly distributed along a power law behaviour of their probability. The behaviour of the cumulative probability distribution of the strength s fits clearly a power law line with a slope exponent close to 2 in both the SMCN and the SiMCN: in this sense those systems can be classified and scale free weighted networks.

Table 7 Comparative overview of the traffic properties of the SMCN and the SiMCN

	$<w>$	w_{max}	$P(w)$	$P(s)$	Y
SMNC	27	13953	Power law with exp 1.8	Power law with exp 2.0	
SiMNC	37.6	10233	Power law with exp 2.14	Power law with exp 2.24	0.69

Regarding to the properties described in table 8 on the analysis of the interplay between traffic, topologic and demographic characteristics, it is possible again to detect similarities between the SMCN and the SiMCN. In both cases a superlinear behaviour in the spectrum of the strength is found with respect to the degree k: in these networks the traffic per connection increases when the degree k increases. In other terms, the higher the topologic centrality of a town, the much (almost twice as much in both systems) higher its traffic centrality. The inspection of the local properties of the weighted networks reveals a common attitude also with respect to the local cohesiveness (clustering coefficient) and the assortativity (average degree of the neighbours). Both networks display the emergence of the rich club phenomenon: two destination towns are preferentially connected to a hub town if the traffic over the connection is large. The common increasing trend of the differential average degree of the neighbour lead to conclude also that the SMCN and the SiMCN can be classified as assortative mixed networks.

Table 8 Comparative overview of interplay properties of the SMCN and SiMCN

	<s>(k)	C^W-C	$K_{nn}{}^w$-K_{nn}	Pop (k)	Pop(S)
SMNC	Upward sloping with exp 1.9	Rich club phenomenon	Assortative mixed	Upward sloping with exp 1.70	Upward sloping with exp 0.90
SiMNC	Upward sloping with exp 1.86	Rich club phenomenon	Assortative mixed	Upward sloping with exp 1.46	Upward sloping with exp 0.83

5 Brief conclusions and outlook remarks

In this paper the authors have developed a framework for a comparative weighted network approach to the commuting system of the island of Sicily, Italy, with respect to an analysis on the correspondent system already performed for the other Italian island of Sardinia.
As a preliminary result, it is possible to state that the Sardinian and the Sicilian inter-municipal commuting networks display similar general and local statistical properties.
This implies that, as far as this study is concerned, similar geographical settings lead to common relevant network properties.
Further study are in progress to extend this approach to the case of different geographical setting, in order to check whether the resulting statistics present the hypothesised divergence.

References

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

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

Barabàsi AL, Albert R, 1999, "Emergence of scaling in random networks", *Science* **286**, 509-512

Barrat A, Barthélemy M, Pastor-Satorras R, 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, 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

Campagna M, Caschili S, Chessa A, De Montis A, Deplano G 2007. Inspecting the influence of space on a complex real network, paper presented at Net2007, Urbino, Italy, May 18-19 2007.

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

Crucitti, P., Latora, V. and Porta , S., 2006, "Centrality measures in spatial networks of urban streets" *Physical Review E* **73**, 036125.

Erdős, P. and Rényi, A. (1959). On random graphs. *Publicationes Mathematicae Debrecen*, 6, 290-297.
Erdős, P. and Rényi, A. 1960, "On the evolution of random graphs" *Publ. Math. Inst. Hung. Acad.* 5, 17-60
Gastner MT, Newman MEJ, 2004, "The spatial structure of networks" *Cond-mat* **0407680**
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, Amaral LAN, 2003, "Structure and efficiency of the World-wide Airport network" *Cond-mat* **0312535**
Italian National Institute of Statistics (Istat), 1991a, 13° Censimento generale della popolazione e delle abitazioni, Matrice origine destinazione degli spostamenti pendolari della Sicilia (13° General census of population and houses, Origin destination matrix of the commuting movements of Sicily).
Italian National Institute of Statistics (Istat), 2001, 14° Censimento generale della popolazione e delle abitazioni, (14° General census of population and houses).
Jiang B, Claramunt C, 2004, "Topological analysis of urban street networks" *Environment and Planning B: Planning and Design* **31** 151-162
Latora V, Marchiori M, 2003 "Economic small-world behavior in weighted networks" *The European Physical Journal B* **32**, 249-263
Latora V, Marchiori M, 2002, "Is the Boston subway a small-world network?" *Physica A* **314**, 109-113.
Newman MEJ, 2003 "Structure and function of complex networks" SIAM review **45**, 167-256.
O'Kelly ME, 1998 "A geographer's analysis of hubs-and-spoke networks", *Journal of Transport Geography* **6,** 171-186.
Pastor-Satorras R, Vespignani A, 2004, "*Evolution and Structure of the Internet*" Cambridge University Press, Cambridge, USA
Reggiani A, Signoretti S, Nijkamp P, Cento A (in press) Connectivity measurement for network models in air transport: a case study of Lufthansa. Proceedigs of Net2006, Springer Verlag.
Schintler LA, Gorman SP, Reggiani A, Patuelli R, Gillespie A, Nijkamp P, Rutherford J, 2005, "Complex Network Phenomena in Telecommunication Systems" *Networks and Spatial Economics* **4**, 351-370.
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
Watts DJ, Strogatz SH, 1998, "Collective dynamics of 'small-world' networks" *Nature*, 393, 440-442