Published in: T. Friesz (ed.) (2007) *Network Science, Nonlinear Science and Infrastructure Systems*.
Springer, New York, pp. 165–83
The final publication is available at www.springerlink.com

AN APPLICATION OF COMPLEX NETWORK THEORY TO GERMAN COMMUTING PATTERNS

**Sean P. Gorman,[1] Roberto Patuelli,[2] Aura Reggiani,[3]
Peter Nijkamp,[2] Rajendra Kulkarni,[1] Günter Haag[4]**

[1]George Mason University
Fairfax, VA, USA
sgorman1@gmu.edu; rkulkarn@gmu.edu

[2]Free University
Amsterdam, The Netherlands
rpatuelli@feweb.vu.nl; pnijkamp@feweb.vu.nl

[3]University of Bologna
Bologna, Italy
reggiani@economia.unibo.it

[4]Steinbeis Transfer Centre Applied System Analysis (STASA)
Stuttgart, Germany
haag@stasa.de

ABSTRACT

Simulating the structure and evolution of complex networks is an area that has recently received considerable attention. Most of this research has grown out of the physical sciences, but there is growing interest in their application to the social sciences, especially regional science and transportation. This paper presents a network structure simulation experiment utilizing a gravity model to identify interactions embodied in socio-economic processes. In our empirical case, we consider home-to-work commuting patterns among 439 German labour market districts. Specifically, the paper examines first the connectivity distribution of the German commuting network. The paper next develops a spatial interaction model to estimate the structure and flows in the network concerned. The focus of this paper is to examine how well the spatial interaction model replicates the structure of the German commuting network as compared to complex network models. Finally, the structure of the physical German road network is compared to the spatial flows of commuters across it for a tentative supply-demand comparison.

KEYWORDS: complex networks, commuting, infrastructure

1 INTRODUCTION

The separation of residential and job decisions has led to complex commuting patterns which have extended in geographical scale over the past decades. As a consequence, home-to-home trips have adopted multi-regional network configurations and have thus led to complex interactive networks. Commuting has become an important field of study in geography, transportation science and regional science (see Rouwendal and Nijkamp, 2004). Commuting has long been studied mostly in terms of forecasting and approximating flows (see, for example, Fotheringham, 1983; White, 1977, 1986). Recent works include the application of models such as the one developed at STASA (Haag et al., 2001). However, less efforts have been done in studying the structure and connectivity properties of commuting networks. A certain amount of literature is available, which studies commuting in a spatial framework. In a recent paper, Cörvers and Hensen (2003) used regional modelling in order to study functional relationships between regions. This approach was carried out in order to improve understanding of commuting behaviour. In particular, the authors' objective was to overcome the limitations of administrative regions in defining new areas that would maximize internal commuting.

A number of works have also been touching on the incorporation of the spatial configuration of commuting destinations, in particular with the works by Fotheringham (see mainly Fotheringham, 1983), who introduced competing destinations models. The competing destinations approach allowed to introduce in the Spatial Interaction Models (SIMs) an element representing the effects of the clustering of destinations, by means of particular accessibility measures. The judgement on this methodology is, however, not uniform. Network approaches to commuting have also been proposed, both at urban level (see, for example, Sheffi, 1985), and at zonal level (Thorsen et al., 1999). A graph theory approach has instead been proposed by Binder et al. (2003).

The above approaches, though, mainly revolve around the analysis of the effects of the road network on commuting. One further level of analysis is therefore necessary, which looks at the behaviour shown in terms of connections between the nodes of the commuting network. Several questions, in fact, need to be answered about the network. Is it highly centralised or decentralised? What are the efficiency and reliability implications of this and other of its connectivity properties? Complex network theory – if applied to commuting networks – can help answering these questions. A wide literature bloomed in recent years, studying the structural and performance implications – on transportation networks – of hypothetical natural disasters or terrorist attacks. The centralization or clustering levels of networks are therefore critical in such discussion.

In this paper, we intend to analyse the network properties of observed home-to-work commuting in Germany. These findings of this analysis are ultimately compared to the ones found by carrying out two simulation models: the first model is an unconstrained SIM, while the second model proposed is modelled according to the scale-free network theories recently made popular by the works of Barabasi and Albert (BA) (Barabasi, 2001; Barabasi and Albert, 1999), showing that networks with preferential attachment-based growth tend to be highly efficient and centralised. In addition to the above analysis, we also propose an analysis of the main German road network, by means of a shortest-path algorithm, and subsequently compare the structural properties found – for the road network – to the ones found for the real data and the simulation models.

The paper is therefore structured as follows: the next section briefly reviews complex network theories and their main implications to our discussion. Section 3 presents a some issues associated with SIMs, and introduces the model carried out for our experiments. Subsequently, Section 4 will first present the results of our empirical application. First, the data used for our analysis are presented, then the findings of the comparison between the properties of the observed commuting network and the simulation models are shown. Finally, a discussion of these findings in respect of the ones obtained by an analysis of the physical German road network is presented. Lastly, conclusions and future research directions are drafted in Section 5.

2 COMPLEX NETWORK THEORY: A BRIEF REVIEW

This section briefly reviews the main issues related to complex network theories, and in particular their implication for transportation networks. Contrary to the attention complex networks have been receiving in recent years, the study of such networks is not particularly new. Before Albert and Barabasi's groundbreaking discoveries, original research had in fact been carried out some forty years ago by Erdös and Renyi (ER) (1960), whose major assumption was an underlying random network structure. However, because of lacking computational power and suitable data, for the majority of the 20th century these theories were not adequately challenged and represented the basis for the most common methods of network simulation (Barabasi, 2001).

Finally, Albert and Barabasi (2002) found, in more recent times, that (large) complex networks were actually behaving according to three main characteristics:

1. Short average path length
2. High level of clustering
3. Power law and exponential degree distributions

In detail, short average path length indicates that any two nodes on the network can be reached with a limited number of hops. High clustering, instead, occurs because of nodes locating topologically close to each other in cliques that are well connected to each other. This property had been formalised by Watts and Strogatz (1998). Finally, the frequency distributions of node density (or, more generally, number of connections) are called degrees and can follow power law distributions. This third property implies connections that cut across the graph, directly linking different clusters of vertices. These direct links between clusters bring an increased level of efficiency – in terms of number of hops – to the network. This result shows the limits of the ER models, in which the exponential decay of the degree distribution did not imply a higher number of connections available to the most important nodes.

The novelty in the AB approach was in fact incorporating an additional component: network growth. Consequently, not only the number of nodes in the network can increase, but new nodes are found to have a higher probability of connecting to other nodes that are already well-connected. Formally, the mechanisms that govern network growth towards a power law degree distribution are (Chen et al., 2001):

a. Incremental growth – As observed above, the number of nodes in the AB models is allowed to grow and add nodes.
b. Preferential connectivity – Preferential connectivity expresses the frequently encountered phenomenon that new nodes have a higher probability to connect (or reconnect) to an existing node that already has a large number of connections (i.e. high vertex degree).

c. Re-wiring – Re-wiring can be considered as a consequence of the previous principle, as some links can be removed and re-connected in the network, though pointing at new nodes, on the base of preferential connectivity.

Still, after the recent developments described, there is a debate on how complex networks should be classified. Different ideas, for example, are proposed by Albert and Barabasi (2002) and Amaral et al. (2000) (see Schintler et al., 2005). Generally, the cause-and-effect relationships underlying large complex networks are still not exactly clear. Furthermore, the measures according to which networks should be measured are also in discussion. For example, Li et al. (2004, p. 11) suggest that – in the case of engineered networks – robustness should be "defined in terms of network performance" and be "consistent with the various economic and technological constraints at work." Remarkably, Li et al. employ a gravity model in generating their network's maximum throughput.

A certain amount of literature is now available on the analysis of transportation networks in terms of complex theory. Because of their short average path length, airline networks have been considered by Amaral et al. (2000) as a *small-world* network, referring to the model presented by Watts and Strogatz 1998). On the other hand, the same authors note that structural limitation of airline networks, such as limited space available in the airports, may hinder the emergence of scale-free properties. Other authors found similar results. Latora and Marchiori (2002) analysed the Boston subway network, while Schintler and Kulkarni (2000) observed congested road networks. Both articles found small-world network properties in the analysed networks.

Generally, one might argue that transportation networks are less prone to evolve into a scale-free structure over time given the fact that they tend to be planar. In fact, in planar networks, the maximum number of connections for a single node can be limited by the physical space available to connect it to other nodes, and it is this fact that makes the large number of connections needed for finding a power law distribution more difficult to obtain. Further, it may be observed that highly centralized transportation networks can be subject to threats to viability, in case of destruction of large hubs (Kwan et al., 2003). Scale free networks have many implications, but a far-reaching consequence of their unique hub structure is that they are very fault tolerant, while also susceptible to attack (Albert et al., 2000). Specifically, a scale-free network model remains connected when up to the 80% of nodes are randomly removed from the network, but when the most connected nodes are removed, the average path length of the network increases rapidly, doubling its original value when the top 5% of nodes are removed (Albert et al., 2000). In short, targeting the most connected nodes can cause significant damage to a scale-free network, making it highly susceptible to a coordinated and targeted attack. Further, these numbers and findings were highly similar to the ones found when real world networks were tested, including the Internet at the autonomous system level, and the WWW. When the most connected networks and web pages were attacked, the network rapidly failed. Albert et al.'s work was complimented by the analysis of Callaway et al. (2000), modeling network robustness and fragility as a percolation, and by Cohen et al. (2001), who used related methodologies. Preliminary analyses of these models on spatial network data have shown similar results when cities are the nodes and fiber connections between them are the links. Utilizing a model of node connectivity and path availability, Grubesic et al. (2003) found that the disconnection of a major hub cities can cause the disconnection of peripheral cities from the network. Spatial analysis of network failure has also been done for airline networks, finding similar results for the Indian airline network (Cliff et al., 1979).

Starting from these considerations, the next section will present the SIM that was modelled as an approximation of preferential attachment, in order to be compared to a scale-free model inspired by the theories described above.

3 SPATIAL INTERACTION MODELS: AN APPROXIMATION TOOL FOR PREFERENTIAL ATTACHMENT

3.1 Spatial Interaction Models for identifying Commuter Flows in the German Labour Market Network

Spatial interaction models are arguably one of the most common methods employed and studied for estimating commuting flows (see, recently, Johansson et al., 2003; Jörnsten et al., 2004; Thorsen and Gitlesen, 1998). Generally, SIMs have long been a popular technique for describing and explaining behavioural, demographic and economic phenomena in space (see Sen and Smith, 1995, for an extensive presentation of the family of methods). The main reason for the widespread utilization of SIMs is their simple mathematical form, in addition to the intuitive assumptions underlying the approach. It should be remembered that the most common specification of SIM had its origins in a resemblance to Isaac Newton's law of universal gravitation. The idea of utilizing models derived from this theory had already been introduced, in the 19^{th} century, in the field of social sciences by Carey (1858) and Ravenstein (1885), and subsequently mathematically formalised by Stewart (1941). Remarkably, SIMs have been shown to have theoretical justification in the entropy theory and in utility maximization/cost minimisation (see, for example, Nijkamp, 1975; Nijkamp and Reggiani, 1992). While Isard (1960) first suggested the use of SIMs in regional science, the entropy root of SIMs introduced by Wilson (1967; 1970), and subsequently the micro-economic derivation introduced by McFadden (1974; 1979) contributed to make SIMs more suitable to interpret spatial-economic phenomena.

The common form of a SIM (here presented as double-constrained) is as follows:

$$T_{ij} = A_i B_j O_i D_j f(-\beta c_{ij}) \quad \text{for } i = 1,...,I; j = 1,...,J \, , \tag{1}$$

where:

$$A_i = 1 \Big/ \sum_j B_j D_j \exp(-\beta c_{ij}) \, ; \tag{2}$$

$$B_j = 1 \Big/ \sum_i A_i O_i \exp(-\beta c_{ij}) \, ; \tag{3}$$

T_{ij} measures the flow of interaction between the origin i and the destination j, depending on the stock variables O_i and D_j, as well as on the deterrence function $f(-\beta c_{ij})$, and on the balancing factors A_i and B_j (see Reggiani, 2004). Generally, Fotheringham and O'Kelly (1989, p. 10) formulate a SIM in the general framework of the Alonso model (1978) as follows:

$$T_{ij} = f(\mu_1 v_1^1, \mu_2 v_1^2, ..., \mu_p v_1^p; \alpha_1 w_j^1, \alpha_2 w_j^2, ..., \alpha_q w_j^q; \beta c_{ij}) \, , \tag{4}$$

where v_i and w_j are measures of the "propulsiveness" and the "attractiveness" of i and j, respectively. Parameters μ, α and β link the above variables to the flows T_{ij}.

The deterrence function in (1) is depending on the deterrence factor β and the interaction costs c_{ij}. c_{ij} might also be considered as generalised costs. In our experiment, distances were used as a proxy of the interaction costs, since the analysis was carried out at the German district level (*kreise*). The functional form of the deterrence function is also a relevant issue. While in its first formulations the distance deterrence function was shaped as a power law function – as used in the Newtonian formula – Kulldorf (1955) showed that an exponential deterrence function seemed to better fit migration phenomena. Subsequently, the exponential deterrence form emerged mathematically from the entropy maximisation approach developed by Wilson (1967). The power form outlines a larger amount of flows – with respect to the exponential form – in the presence of long distances or travel times. In our analysis, the power law specification was used, as it showed to fit the data better. In addition to the shape of the deterrence function, the value of the β deterrence factor was researched. In the experiments conducted here, a value of 1.5 was chosen for the β deterrence factor, on the basis of a calibration procedure carried out on the available data expressed in the form of an unconstrained SIM. In particular, the unconstrained SIM used in our experiments is specified as follows:

$$T_{ij} = KE_i E_j d_{ij}^{-\beta} \tag{5}$$

In model (5), the flows T_{ij} are the employees commuting from the origin district i to the destination district j. They are a function of the number of persons E_i and E_j employed in the two districts, as well as of the distance d_{ij} between the two, in addition to a scaling factor K. The model that we propose is of course overly simple. However, what is relevant for our experiments is not the correct estimation of the German commuting flows, but instead the connectivity and structure of the commuting network (see Section 3.2).

When employing a SIM for estimating inter-urban commuting flows, additional issues should be cited. One of them is the treatment of internal commuting. In particular, the distance between the working and living areas is, by definition, null (although travel time or costs would not necessarily be). This issue is at times solved by assigning an arbitrary value to the distance for internal commuting. Alternatively, the flows assigned to internal commuting can be omitted in the analyses. A number of additional ways to treat internal commuting are available in the literature. The method suggested by Thorsen and Gitlesen (1998) starts from the consideration that intra-commuting might imply different transportation means, such as biking or walking. Thorsen and Gitlesen suggest an additional component to be added to the deterrence function exponent. This component would represent – depending on the case – either a start-up (generalized) cost for commuting between different zones, or a premium, interpreting the benefit of intra-commuting. An example model with these characteristics, reminiscent of the Champernowne deterrence function (see, for example, Sen and Smith, 1995), is presented by Thorsen and Gitlesen (1998, p. 279) for a double-constrained specification. Alternatively, the authors suggest that labour market characteristics might be used to influence the elements on the diagonal of the O/D matrix.

In our case, the elements of the diagonal are omitted from the analysis. This choice was taken mainly due to the network approach to commuting identified in the paper. As we analyse the connectivity and structural properties of the German commuting network, the measure of the number of commuters within a certain district would not add additional information about the network, apart from the "socio/economic weight" of a certain node. On the other hand, the number of fulltime employees in each district already grasps this aspect.

3.2 Interpretation of Spatial Interaction Behaviour as Preferential Attachment

The usual practice in the use of SIMs, when dealing with commuting flows, is to employ the models in forecasting future flows, given certain conditions. In our experiments, we propose the utilization of the simple SIM shown in Equation (5) as a tool for approximating the connectivity and structural properties of a commuting network. In particular, we want to verify if a SIM can allow for preferential attachment behaviour. As seen in Section 2, in the models introduced by Barabasi and Albert, nodes have a higher probability of connecting to other nodes that are already well-connected. The hypothesis that we will test in the next section is that commuting networks follow a similar preferential attachment-based behaviour in terms of connectivity and structure. They would not be the first transportation network to be referred to in this terms. In fact, hub-n-spoke networks operated by airlines are, as seen in Section 2, a well-known example of preferential attachment behaviour (see, for example, Bowen, 2002, and most importantly Wojahn, 2001).

An additional reason for the consideration of commuting networks in such a framework can be found if we think of preferential attachment as a maximization of utility levels. The idea is that utility is maximized by connecting to the most connected nodes of the network, as they give access to other points in the network by a minimal number of hops (therefore minimizing generalized costs). If so, this hypothesis would be consistent with the theoretical basis of utility maximization that justifies the use of SIMs. In particular, the hub-n-spoke network might – conceptually – be interpreted as a network tree consistent with a nested logit/hierarchical SIM structure (for the compatibility between nested logit and double constrained SIM, see Nijkamp and Reggiani, 1992).

The next section will first describe the data available for the experiment (see Section 4.1). We will then test the hypothesis of a SIM as an approximation of preferential attachment, by comparing the network properties observed for our naïve SIM, for a scale-free network, and for the real commuting network (Section 4.2). Subsequently, in Section 4.3 we will present the results of a structural analysis of the German road network (the physical network on which commuting is actually performed), and compare the properties found with the results previously obtained.

4 THE EMPIRICAL ANALYSIS

4.1 The Data Available

As shown in Section 3.1, the SIM estimated for our experiment formally employs two types of data:

a. The number of employees working in each German district. These are fulltime employees, for which the data were collected – as part of a yearly survey – in the year 2002.
b. The distance between each couple of origins and destinations. This is expressed in Kilometres, and acts as a proxy for more effective measurements, such as cost, or travel time.

Moreover, additional data are available for our experiment. In order to calibrate our SIM shown in model (5) (see Section 3.1), information on the German commuting flows has been used. The data consists, for each origin-destination couple (i,j), of the number of employees living in district i, and working in district j. These are therefore home-to-work data, which are available for the year 2002. The distribution of the number of commuters for each origin-destination couple, ranked in descendent order, is shown in Figure 1.

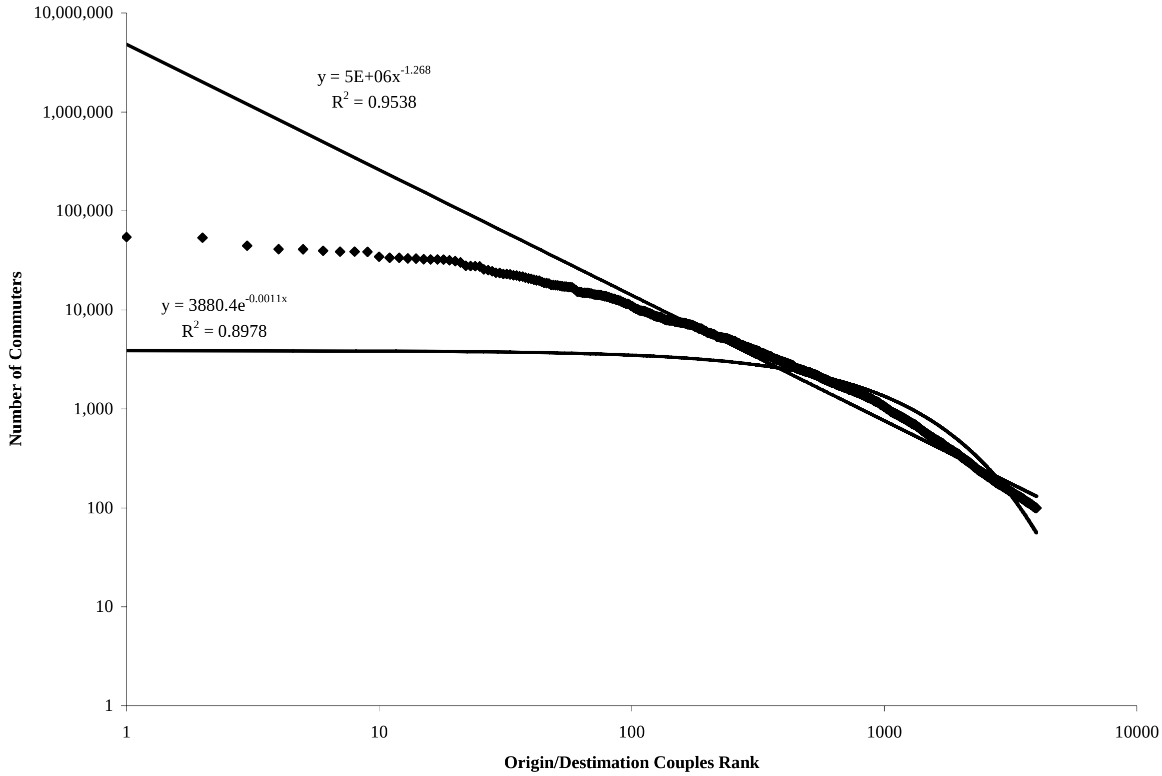

Figure 1 – The distribution of commuters flows over the couples of origins and destinations

Figure 1, which is adapted to a log-log scale, shows the decrease in the number of commuters for the (descendent) rank of origin-destination couples. The curve seems to better fit a power law distribution rather than an exponential one. However, the highest values in the data – which are the first values in the descending order rank – do not reach the levels expected for a power law fitting. This could reasonably be due to physical constraints given by city size and congestion issues. Nevertheless, the distribution shown in Figure 1 does not tell about the intrinsic properties of the network. The next section will therefore analyze the connectivity properties of the observed commuting network, and of the ones developed by the models presented in our exercise.

4.2 Analysis of Network Structure and Connectivity Properties in Germany

The goal for this experiment is to discover how well a spatial interaction model can be used to accurately model preferential attachment and the resulting structure of the German commuting network. The use of the BA model as structural model for network connectivity has received considerable attention, critique and extension. This section of the paper strives to refine the BA model for application to spatial economic networks, specifically the patterns of connectivity found in German commuting to work. The focus of the analysis is how well the spatial interaction model builds the topology on the real world network and not just its connectivity distribution. Li et al. (2004) have found that the power law distributions created by many network generation models can result in a wide range of actual topologies.

The first step of the analysis was to specify an accurate SIM for the German commuting network. Many economic and social forces shape the contours of commuting patterns. A simple supply and demand model based on employment and distance was used in this initial case. The SIM was formalised in Section 3.1, model (5). Once the spatial interaction for all possible pairs of German labour markets is calculated, a threshold for connectivity is determined. In this case, a threshold of 100 commuters was determined as the cut off for what would be considered an adequate economic flow for there to be a connection in the network between two (labour) districts. The resulting edge list of pairwise connections between districts was then used for comparison to the BA model and real world data. The BA model created for comparison was based on a 439-node network with a 0.3 connectivity probability, an alpha parameter of 0.3, and an initial three districts to connect to. Once both networks are modelled, they are compared to each other and the real commuting network structure. The resulting network topologies can be seen in Figure 2.

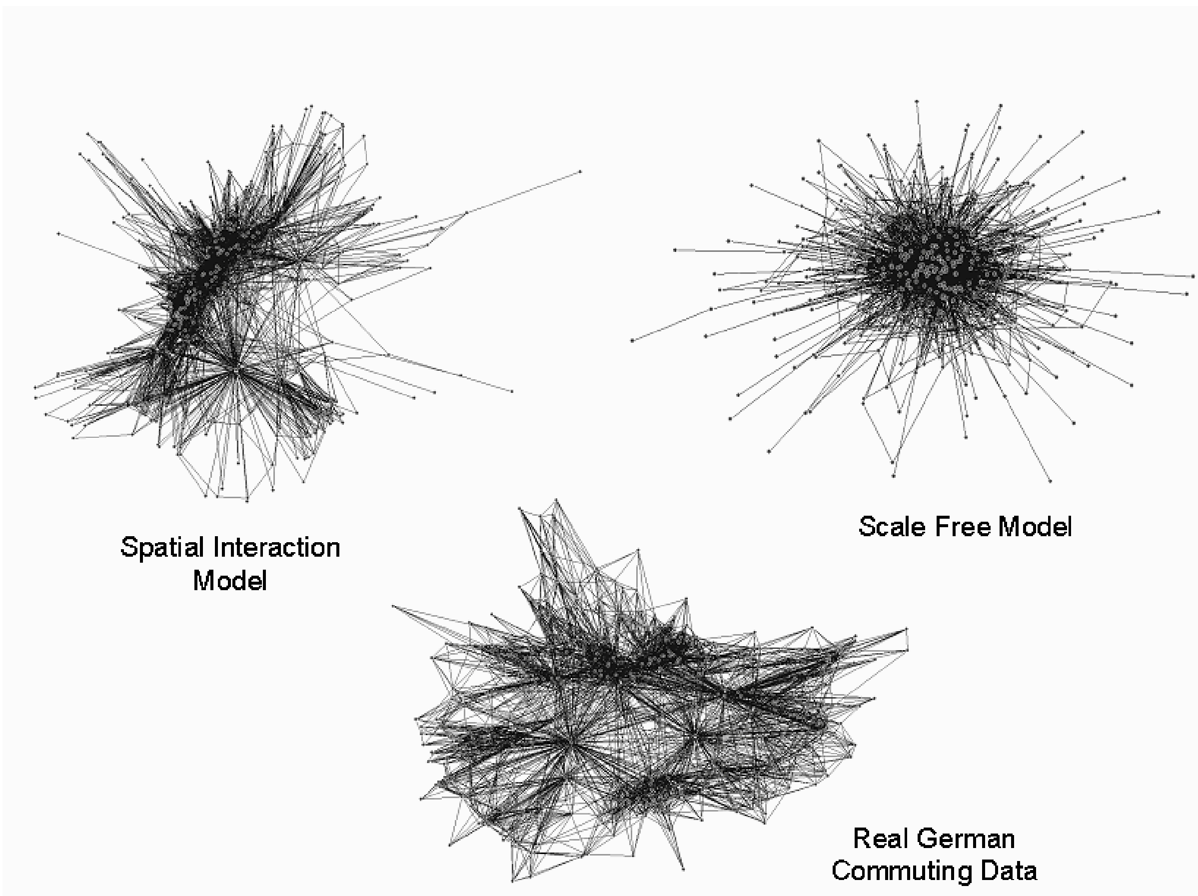

Figure 2 – Network Visualization of the Network BA Model, Spatial Interaction Network Model, and German Commuting Network

From a visual inspection of the three network visualizations, the spatial interaction network model comes closest to replicating the German commuting network, although it lacks the same level of interconnectivity seen in the real data. The BA model illustrates even less interconnectivity that the spatial interaction model, with most connections going directly to the hubs, with little of the connectivity between spokes seen in the real commuting data.

In order to provide the next level of analysis the connectivity distribution of each network is calculated as a log-log plot. The results of all three network connectivity distributions are plotted in Figure 3. The number of connections for each district is ranked in descendent order.

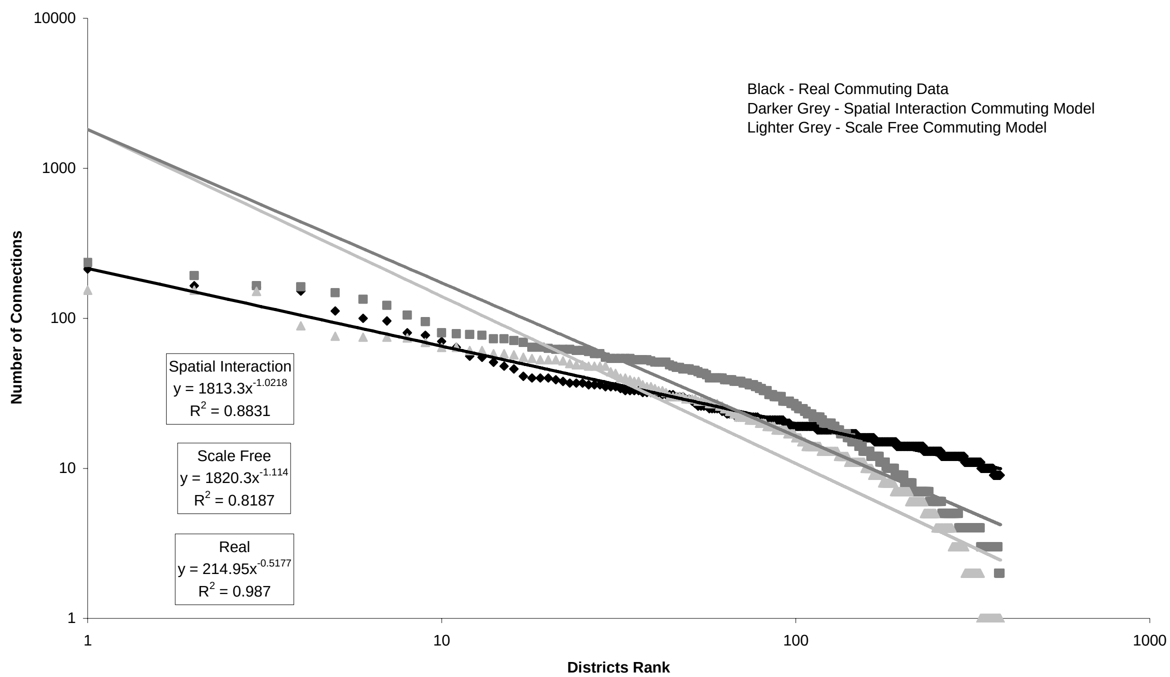

Figure 3 – The resulting connectivity distributions of the BA network model, spatial interaction network model, and the German commuting network

Interestingly, neither of the models recreated the power law connectivity distributions seen in the real data. They have instead exponential distributions of connectivity. The two models also have much steeper slopes, each with an exponent over one while the real data is just over 0.5. Of the two, the SIM is modestly closer in slope and distribution to the real data, but not enough to be of consequence. Overall, the three distributions are highly similar despite the slight numeric differences, while the topology visualizations seen in Figure 2 are vastly different. This is most prominent in the SIM and the BA model, which have very similar connectivity distribution, but entirely different topology visualizations. To examine these topologic differences, a variety of structural indicators are calculated.

For each network, a series of indicators is calculated, which provide a comparison of the structure of each network generated (see Table 1). The first indicator presented in Table 1 is the clustering coefficient, which provides the first insight into what delineates the structural differences in the three networks. The lack of clustering seen visually in the BA model is found in the statistics, with a far lower coefficient than that of the SIM or of the real world data. The SIM overestimates the level of clustering seen in the real data, but is considerably closer than the BA model. The diameter parameter calculates the longest shortest-path in the network, and provides a measure of the efficiency of the structure. All three networks have relatively small diameters, and the real world commuting network is the most efficient. Again, the diameter of the SIM is closer than the one of the BA model, off by only one. The subsequent set of statistics deals with the number of connections nodes in the network have. These are the statistics that are more of local connectivity indicators, and not global structural indicators of the networks. From a connectivity standpoint, the SIM does a nice job of accurately capturing the average, maximum, and minimum of the degree connectivity seen in the real commuting data. It falls short in capturing the standard deviation of average connectivity, where the BA model is closer, but the average connectivity is still six degrees off of the real data. The centralization parameter measures the amount of core connectivity in the network. The SIM overestimates the amount of centralization, while the BA model underestimates it. However, the SIM is closer to the real data values. Lastly, a measure of betweenness is offered for each network. Betweenness is a measure of routing frequency, where all shortest paths across a network, and then the number of times each node is used in all paths, are calculated. This provides a convenient measure of the global structure of the network, since the indicator samples paths across all segments of the network. The average and standard deviation of betweenness for the SIM and the real data are very close. This confirms what was seen previously in the network visualization, which illustrated a similar topological structure between the two networks. The significant differences between the real network and the BA model, according to betweenness measures, confirm the distinct differences between the two also seen in the visualization.

Table 1 – Network properties of the BA network model, spatial interaction network model, and the German commuting network

	Real Data	*BA Model*	*SIM*
Clustering Coeff.	0.659	0.29	0.803
Diameter	4	6	5
Average Degree	18.21	12.469	18.462
Std. Dev. Degree	19.268	19.174	26.805
Max Degree	213	154	235
Min Degree	5	0	1
Centralization	44.68%	32.46%	50.36%
Betweeness Mean	313.125	251.866	297.083
Betweeness Std. Dev.	1929.005	935.554	1997.69

4.3 A Structural Analysis of the Physical Commuting Network by Means of Shortest Path Algorithm

In addition to studying the flows of commuters between cities, it is also possible to study the structure of the infrastructure they utilize. The economic flows of commuters and the physical links of infrastructure are intrinsically connected, but belong to two very different network structures. Commuting flows are belong to logical networks, which are non-planar in nature, since the fact that two links intersect does not mean a node actually exists at their intersection. A flow in the commuter network could therefore be between Frankfurt and Munich with only two nodes and one link, even though the physical path goes through Stuttgart. The physical network, on the other hand, is planar; the intersection of two links creates a navigable intersection. In order to travel from Munich to Frankfurt, several intermediate nodes have to be traversed. Commuting data represent the flows across the physical network, but the two networks are quite different in nature and structure.

To begin to address the relationship and differences of these networks, the physical road network of Germany is analysed. Unlike the commuting flow network, it is straightforward to visualize what the road network looks like with a simple map. Figure 4 provides a map of the German road network.

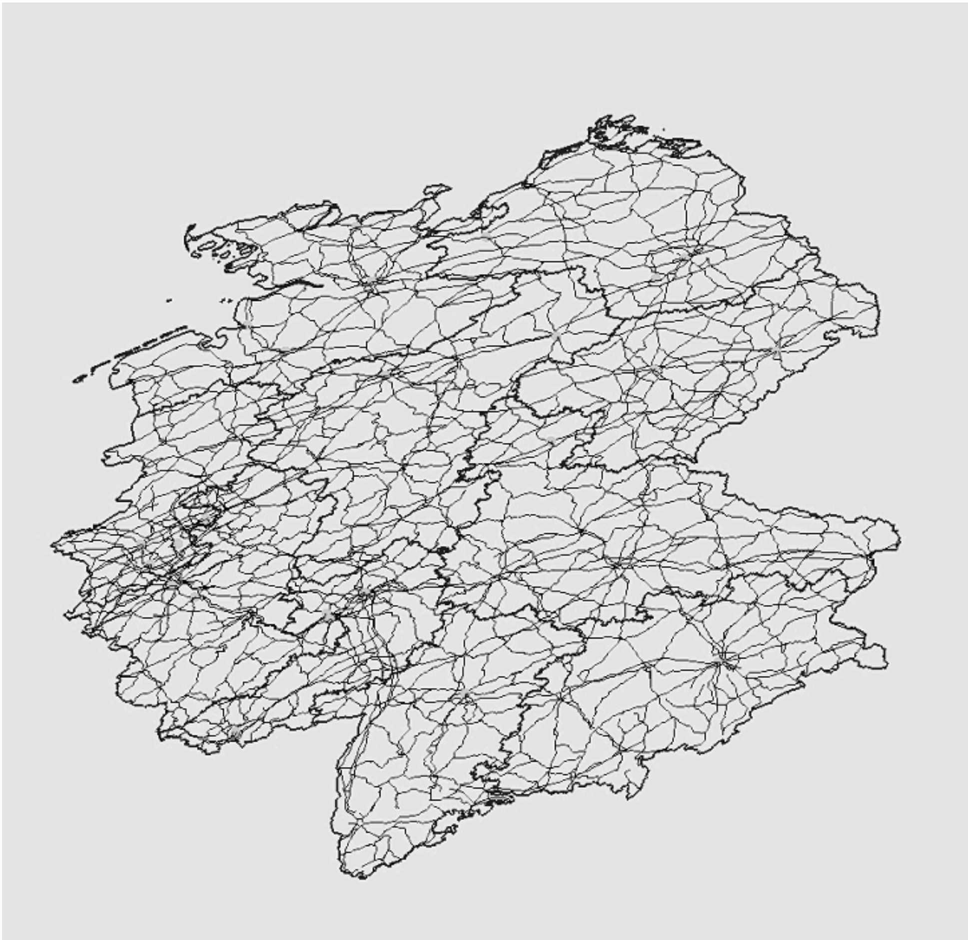

Figure 4 – The German road network

In Figure 4, the white lines are major roads in Germany, while the dots represent major cities connected by the roads. While the map does illustrate the layout of the road network, it does not give much insight into its structural properties. In order to gain some perspective on the structure of the road network, a routing frequency analysis was performed. The road network was first partitioned into nodes and links, then shortest-paths were calculated to and from all nodes in the networks. Links were then assigned a frequency count, based on the number of times the link was utilized in all possible link combinations. This provided a structural analysis of which links, in the German road network, are most critical and heavily utilized in all possible travel combinations. To visualize these results Figure 5 was built.

In Figure 5, the height and colour (lighter to darker) of the peaks is determined by the number of routes that use a particular link in the road network. The higher the frequency of routes, the higher the peak. The routes that connect through the middle of the country are particularly well utilized, especially the routes connecting Berlin to Frankfurt and Stuttgart. The routes connecting Berlin to Munich and Hamburg are also prevalent. In general, routes in Western Germany have a higher frequency than Eastern Germany routes. It should be noted that this analysis is simply based on shortest-path frequency and does not account at all for socio-economic dimensions like population or employment as with the previous gravity model.

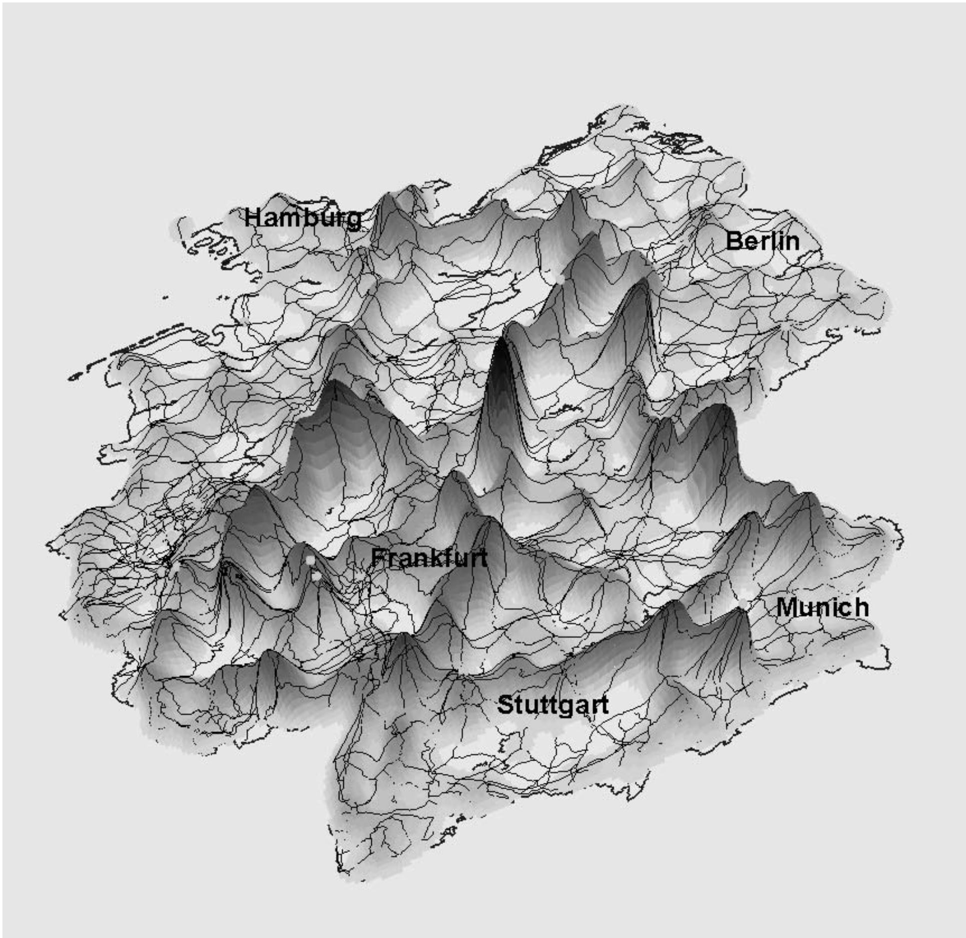

Figure 5 – German road network routing frequency analysis

4.4 Concluding Remarks

The present analysis does provide preliminary insights into the German infrastructure network, which underpins the economic commuter network flowing on top of it. The intuitive next direction is to combine the two networks and examine their interdependencies. It would be useful to investigate the relationship between the economic flows of commuters and the physical network structure of the roads they use. Are the highest flows of commuters also utilizing the highest-frequented structural links? How well does the physical structure of the network match the economic flows across it? If there is failure in the physical network, how will it impact the economic flow of commuters or, more importantly, logistics and supply chain networks? Unfortunately, it was not possible, at this stage, to obtain a geo-referenced map of the German districts to map the commuter flows on top of the physical routing frequency. Such a possibility would at least provide a first-cut comparison of commuting links in a labour district to the number of routing paths. These are all possible future directions for the research.

5 CONCLUSIONS

The aim of this paper was to provide a comparative analysis among different approaches on the theme of commuting from a network perspective. In particular, we were interested in studying the structure and properties of the German commuting network. In order to better understand this real-world commuting network, we needed to compare it to different network models that could approximate it. Two models (a SIM and a scale-free model) were developed, based on widely different assumptions. However, both models aimed at simulating a “preferential connectivity” behaviour, according to which, well-connected nodes in the network have a higher probability of attracting connections from other nodes over the network.

A comparison of the real (network) commuting data with the ones generated by our two models was carried out in two ways: a) a visual comparison of the network structures; b) a comparison of the network properties calculated for the three networks. The visual comparison showed that the SIM seemed to better approximate the decentralized configuration of the commuting network. The scale-free network showed instead a highly centralized structure. These observations were then reinforced by the comparison of the network properties parameters. Although showing similar values for some network parameters, the SIM, rather than the scale-free model, provided values that were more consistent with the ones calculated for the real data.

An additional analysis was subsequently carried out, examining the German physical road network. This analysis on the road infrastructure visually showed which points, according to a shortest-path routing algorithm, are (theoretical) critical points in the German road network, as they are central to the routes calculated over the network.

Summarizing, our experiments made a first attempt at interpreting commuting networks from a complex network perspective. More detailed experiments might be carried out, by developing more refined SIMs (like double-constrained SIMs) and scale-free models, which should include parameters that better suit the type of network that is being approximated. A further experiment might be to weigh the number of connections – in a scale-free approach – with the volume of commuters. Also, a set of new experiments could be carried out, by integrating commuting flows (and maybe other economic factors) with the physical road network, as suggested in Section 4.4. For example, it could be interesting to observe how changes in the physical road network would influence the results of a SIM, where the distance deterrence factor is not expected to vary (Jörnsten et al., 2004). Also, the use of double-constrained SIMs is necessary, in order to account for spatial characteristics. General equilibrium models, employing variables such as wages or migration might be used for comparison sake.

REFERENCES

Albert, R., Barabasi, A.-L. (2002). Statistical Mechanics of Complex Networks. *Review of Modern Physics,* 74, 47

Albert, R., Jeong, H., Barabási, A.-L. (2000). Error and Attack Tolerance in Complex Networks. *Nature,* 406, 378-382

Alonso, W. (1978). A Theory of Movements. In N. M. Hansen (Ed.), *Human Settlement Systems*. Cambridge, Mass.: Ballinger

Amaral, L. A. N., Scala, A., Barthélémy, M., Stanley, H. E. (2000). Classes of Small-World Networks. *Proc. Nat. Acad. Sci. USA,* 97, 11149-11152

Barabasi, A.-L. (2001). The Physics of the Web. *Physics World,* July 2001

Barabasi, A.-L., Albert, R. (1999). Emergence of Scaling in Random Networks. *Science,* 286, 509-512

Binder, J., Haag, G., Rabino, G. (2003). Analysis and Modelling of Commuter Flows: Application to the Regions of Stuttgart and Turin. *Jahrbuch der Regionalwissenschaften,* 23, 117-139

Bowen, J. (2002). Network Change, Deregulation, and Access in the Global Airline Industry. *Economic Geography,* 78(4), 425-439

Callaway, D. S., Newman, M. E. J., Strogatz, S. H., Watts, D. J. (2000). Network Robustness and Fragility: Percolation on Random Graphs. *Physical Review Letters,* 85(25), 5468-5471

Carey, H. C. (1858). *Principles of Social Science*. Philadelphia: J. Lippincott

Chen, Q., Hyunseok, C., Govindan, R., Sugih, J., Schenker, S., Willinger, W. (2001). *The Origin of Power Laws in Internet Topologies Revisited.* Paper presented at the IEEE Infocom, 2002

Cliff, A., Haggett, P., Ord, H. (1979). Graph theory and geography. In R. Wilson, L. Beineke (Eds.), *Applications of Graph Theory* (Vol. 293-326). London: Academic Press

Cohen, R., Erez, K., Ben-Avraham, D., Havlin, S. (2001). Breakdown of the Internet Under Intentional Attack. *Physical Review Letters,* 86(16), 3682-3685

Cörvers, F., Hensen, M. (2003). *The Regionalization of Labour Markets by Modelling Commuting Behaviour.* Paper presented at the 43rd ERSA, Jyvaeskylae

Erdos, P., Renyi, A. (1960). *On the Evolution of Random Graphs* (Vol. 5): Publication of the Mathematical Institute of the Hungarian Academy of Science

Fotheringham, A. S. (1983). A New Set of Spatial-Interaction Models: The Theory of Competing Destinations. *Environment and Planning A,* 15, 15-36

Fotheringham, A. S., O'Kelly, M. E. (1989). *Spatial Interaction Models: Formulations and Applications*. Dordrecht: Kluwer Academic Publishers

Grubesic, T. H., O'Kelly, M. E., Murray, A. T. (2003). A Geographic Perspective on Telecommunication Network Survivability. *Telematics and Informatics,* 20(1), 51-69

Haag, G., Binder, J., Koller, M. (2001). *Modellgestützte Analyse zur Regionalen Entwicklung von Beschäftigungsvolumen, Lohnsummen und Beitragseinnahmen*. Nuernberg and Stuttgart: Institute of Labour and Employment Research (IAB) and STASA

Isard, W. (1960). *Methods of Regional Analysis*. Cambridge: MIT Press

Johansson, B., Klaesson, J., Olsson, M. (2003). Commuters' Non-linear Response to Time Distances. *Journal of Geographical Systems,* 5(3), 315-329

Jörnsten, K., Thorsen, I., Ubøe, J. (2004). Replication/Prediction Problems in the Journey to Work. *Environment and Planning A,* 36(2), 347-364

Kulldorf, G. (1955). Migration Probabilities, *Lund Studies in Geography, Series B: No. 14*. Lund, Sweden: Department of Geography, Lund University

Kwan, M.-P., Murray, A. T., O'Kelly, M. E., Tiefelsdorf, M. (2003). Recent Advances in Accessibility Research: Representation, Methodology and Applications. *Journal of Geographical Systems,* 5, 129-138

Latora, V., Marchiori, M. (2002). Is the Boston subway a small-world network? *Physica A,* 314, 109-113

Li, L., Alderson, D., Willinger, W., Doyle, J. (2004). *A First-Principles Approach to Understanding the Internet's Router-Level Topology.* Paper presented at the SIGCOMM'04, Aug. 30-Sept. 3, 2004, Portland, Oregon, USA

McFadden, D. (1974). Conditional Logit Analysis of Qualitative Choice Behavior. In P. Zarembka (Ed.), *Frontiers in Econometrics* (pp. 105-142). New York: Academic Press

McFadden, D. (1979). Qualitative Methods for Analysing Travel Behaviour of Individuals: Some Recent Developments. In D. A. Hensher, P. R. Stopher (Eds.), *Behaviour Travel Modelling* (pp. 279-318). London: Croom Helm

Nijkamp, P. (1975). Reflections on Gravity and Entropy Models. *Regional Science and Urban Economics,* 5, 203-225

Nijkamp, P., Reggiani, A. (1992). *Interaction, Evolution and Chaos in Space*. Berlin Heidelberg: Springer-Verlag

Ravenstein, E. G. (1885). The Laws of Migration. *Journal of the Royal Statistical Society,* 48(2), 167-235

Reggiani, A. (2004). Evolutionary Approaches to Transport and Spatial Systems. In P. R. Stopher, K. Button, K. Haynes, D. A. Hensher (Eds.), *Handbook of Transport Geography and Spatial Systems* (pp. 237-253). Amsterdam: Elsevier

Rouwendal, J., Nijkamp, P. (2004). Living in Two Worlds: A Review of Home-to-Work Decisions. *Growth and Change,* 35, 287-303

Schintler, L. A., Gorman, S. P., Reggiani, A., Patuelli, R., Nijkamp, P. (2005). Scale-Free Phenomena in Communications Networks: A Cross-Atlantic Comparison. In A. Reggiani, L. A. Schintler (Eds.), *Methods and Models in Transport and Communications. Cross Atlantic Perspectives* (pp. 201-220). Berlin: Springer-Verlag

Schintler, L. A., Kulkarni, R. G. (2000). The Emergence of Small World Phenomenon in Urban Transportation Networks: An Exploratory Analysis. In A. Reggiani (Ed.), *Spatial Economic Science: New Frontiers in Theory and Methodology* (pp. 419-434). Berlin: Springer

Sen, A., Smith, T. E. (1995). *Gravity Models of Spatial Interaction Behavior*. Heidelberg and New York: Springer

Sheffi, Y. (1985). *Urban Transportation Networks*. Englewood Cliffs, NJ: Prentice-Hall

Stewart, J. Q. (1941). In Inverse Distance Variation for Certain Social Influences. *Science,* 93, 89-90
Thorsen, I., Gitlesen, J. P. (1998). Empirical Evaluation of Alternative Model Specifications to Predict Commuting Flows. *Journal of Regional Science,* 38(2), 273-292
Thorsen, I., Uboe, J., Naelig;vdal, G. (1999). A Network Approach to Commuting. *J Regional Sci,* 39(1), 73-101
Watts, D. J., Strogatz, S. H. (1998). Collective Dynamics of Small-World Networks. *Nature,* 363, 202-204
White, M. J. (1977). A Model of Residential Location Choice and Commuting by Men and Women Workers. *Journal of Regional Science,* 17(1), 41-52
White, M. J. (1986). Sex Differences in Urban Commuting Patterns. *The American Economic Review,* 76, 368-372
Wilson, A. G. (1967). A Statistical Theory of Spatial Distribution Models. *Transportation Research,* 1, 253-269
Wilson, A. G. (1970). *Entropy in Urban and Regional Modelling*. London: Pion
Wojahn, O. W. (2001). Airline Network Structure and the Gravity Model. *Transportation Research: Part E: Logistics and Transportation Review,* 37(4), 267-279