

TI 2013-172/VIII
Tinbergen Institute Discussion Paper



Digital Infrastructure and Physical Proximity

Emmanouil Tranos

*Peter Nijkamp**

Faculty of Economics and Business Administration, VU University Amsterdam.

** Tinbergen Institute*

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and VU University Amsterdam.

More TI discussion papers can be downloaded at <http://www.tinbergen.nl>

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 1600

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Duisenberg school of finance is a collaboration of the Dutch financial sector and universities, with the ambition to support innovative research and offer top quality academic education in core areas of finance.

DSF research papers can be downloaded at: <http://www.dsf.nl/>

Duisenberg school of finance
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 8579

Digital infrastructure and physical proximity

Prepared for 'Regional development and proximity relations' (forthcoming), edited by A. Torre and F.

Wallet. Northampton, MA, Edward Elgar.

Emmanouil Tranos, Peter Nijkamp

Dept. of Spatial Economics, VU University

De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands, e.tranos@vu.nl

1. Introduction

¹Some 2000 years ago, the average annual distance a person would normally travel, was approximately 500 km. The action radius of most people remained rather stable, but it rose gradually after the industrial revolution to some 1820 km (by car, bus, railway or aircraft) in the year 1960. Then, a period of rapid increase started, with almost 4390 km per year in 1990. Clearly, air transport, but also technological advances and changing lifestyles formed the background of this megatrend (see Chaturin 1988, Schafer 1988). Accessibility and proximity have become keywords in understanding the geographical pattern of the 'homo mobilis'. The emerging question is if and how this pattern of physical movement will be affected by the digital revolution.

Virtual connectivity through cyberspace has been a source of further unprecedented rise in the action radius of modern man. There is an increasing awareness that cyberspace and physical space are not substitutes. On the contrary these two domains are related with a complex relation, with cyber-place (CP) being the intermediate layer. The latter is essentially an integral part of Batty's (Batty 1997) concept of virtual geography, which is identified as the projection of the infrastructural layer of cyberspace on traditional space. CP is defined in accordance with cyberspace, the function of which is supported by CP. Just like other elements of the Internet jargon such as the Internet superhighways, virtual communities, web-surfing, telecommuting, etc., CP and cyberspace are geographic metaphors. Apart from a way to understand the Internet structure, such metaphors expose its strong spatial foundation (Graham 1998). Just like any other social and economic activity, which is "inscribed in space and takes place" (Swyngedouw 1993, p. 305), the Internet, as a platform for virtual interaction among individuals and organizations, has a geographical component. Following Batty's conceptualization, the virtual geography's element which is mostly responsible for the Internet's spatiality is the CP. Malecki (2002), in

¹ This book chapter draws upon our ongoing research on Cyber-place and proximities (Tranos and Nijkamp forthcoming)

his work about CP, recognizes that this element of the virtual geography fits best with the research questions that economic geography focuses on, as the cyberspace relies on the CP's real world fixities (Kitchin 1998a, Kitchin 1998b).

Despite the interpretation of early commentators that the Internet has an anti-spatial nature (Mitchell 1995), the above discussion highlights the necessity to further explore the spatial dimension of such a system. Indeed, the rapid Internet penetration resulted in deterministic views about the impact of information and communication technologies (ICTs), declaring the death of cities (Gilder 1995, Drucker 1998, Kolko 1999), the death of distance (Cairncross 2001), the emergence of tele-cottages (Toffler 1980), and in general the end of geography, because of the widespread penetration of ICTs. However, such narratives have not been accompanied by hard evidence and empirical investigation. Although we know that "cities are well and alive" (Malecki 2002, p. 419), and that ICTs did not generate such one-way dramatic impacts, there is not yet sufficient empirical knowledge about the relation between physical space and the Internet, especially at the macro – aggregated – level of geographic analysis. Indeed, the complexity of CP and cyberspace needs caution in the a priori adoption of (over-) simplistic approaches such as the cartoonish 'shrinking world' metaphor, which encompasses the complex interrelations between capital, space and technology and the subsequent recasting rather than shrinking impacts on space (Kirsch 1995).

On the basis of the above, this chapter aims to achieve two goals. Firstly, a substantial effort is made in shedding more light on the complex nature of the CP adopting a spatial network perspective. The infrastructural layer of the Internet has developed as a network of myriad different networks. Although the topology of the Internet has been extensively studied from a complex system perspective (e.g. Faloutsos *et al.* 1999, Adamic and Huberman 2002, Pastor-Satorras and Vespignani 2004) and was one of the main test-beds for conceptual network models (Barabási and Albert 1999, Albert and Barabási

2002), to date insufficient effort has been made to approach this complexity from a spatial perspectiveⁱ. Using this analysis, we also draw the big picture of the new – digital – urban hierarchies as they derived by their IP connectivities.

The results of this first analysis will feed into the following, second objective. Although we know that spatial configuration and the importance of agglomeration for social and economic activities is valid in the frame of the digital economy, we still do not know whether and how the Internet is affected by the tyranny of distance. While it is established that the Internet is a friction-reducing technology (Cohen *et al.* 2002, Cohen-Blankshtain and Nijkamp 2004), the effect of distance and proximity on its structure is vague as we do not know whether its infrastructural layer is affected by centripetal or centrifugal forces. Put simply, we do not know if the cost of physical distance affects the structure of a system to such an extent that it diminishes the cost of distance in digital communications.

The novelty of our contribution lies in the fact that although the geographic analysis of the Internet already has a short history of almost 15 years (Moss and Townsend 1997, Wheeler and O'Kelly 1999, Malecki and Gorman 2001, Malecki 2004), the impact of distance on the formation of the Internet – and most specifically on the formation of CP, which is the most well-defined Internet element in a spatial context – has not yet been empirically tested. The above is not surprising, as it reflects the overall disregard of the Internet by the spatial sciences because of its intangible, elusive, and complex technical nature (Bakis 1981, Hepworth 1989, Kellerman 1993). After all, telecommunications infrastructure only becomes visible when it stops working (Star 1999).

To empirically test the above research questions, we utilize here an extensive aggregated data set for the European CP, which, as far as we are aware, has never been used before in a spatial context. Complex network analysis and gravity models which utilize panel data specifications will be employed in our study to quantitatively approach these research questions. From a spatial perspective, two different

levels are used as a sensitivity analysis: a pan-European one using NUTS-3 regions and a local one, using Dutch municipalities.

The structure of the present chapter is as follows. In Section 2, the CP database is described. Then, in Section 3, the network structure of CP is explored using complex network analysis methods. Based on these results, Section 5 explores the urban geography of CP. Section 5 presents the modeling exercise on the impact of distance and proximity on the formation of CP. Finally, this chapter ends in Section 6 with some concluding remarks.

2. Database Description

The main data used for this study is the output of the DIMES project. This is “a distributed scientific research project, aimed to study the structure and topology of the Internet, with the help of a volunteer community” (DIMES 2010). It is based on 3–6 million tracerouteⁱⁱ measurements made daily by a global network of more than 10,000 agents, who are voluntarily participating in this research project (for a description of the DIMES project, see also Shavitt and Shir 2005, Carmi *et al.* 2007). The final outcome of the DIMES project is derived after the triangulation and geo-location of the IP (Internet Protocol) links discovered by the DIMES volunteers, and it contains all the IP links between any two cities discovered by the DIMES agents. This is an infrastructural measure, as the IP links represent physical inter-urban data links, which comply with the IP protocolⁱⁱⁱ.

Nonetheless, a few points need to be highlighted here. Firstly, it needs to be said that this is only a very small fraction of the Internet. Indeed, the DIMES project only includes the IP links which have been captured by the DIMES agents. Because of the existence of the TCP/IP (Transmission Control Protocol / Internet Protocol) protocols, different routes are formed containing different city-to-city IP links to establish a logical link between a data-packet origin and destination. By sending data packets from the

agents' locations to known destinations, DIMES researchers record the different IP links used by its agents, completing in this way the largest available data source for geo-coded IP links. Although different data sets have been used to study the structure of the Internet from a non-spatial topological perspective, to date no other data set with geographic reference is available, at least not to the same extent as the DIMES data set.

In addition, an inherent limitation of the geographic analysis of the Internet is its topological rather than geographic basis. Indeed, the Internet was designed as a logical network, the links of which are defined in topological terms, not using geographic coordinates. Thus, the Internet architecture of topological destinations (IP addresses) has little to do with physical locations (Dodge and Zook 2009). In order to understand the above structure from a geographic perspective, an indirect approach is adopted and effort is spent to geo-code the different IP addresses using IP registration tables. This task is part of the DIMES project. We need to highlight here a potential accuracy issue due to the geo-coding process. It is common that IP addresses are owned by specific firms, which lease these IP addresses to the content providers (Dodge and Zook 2009). The outcome of this process is that usually the physical location of the IP address, which is derived by the geo-coding process, does not match with the location of the content. However, this does not create any bias in this study, as the focus of this chapter is the CP and the physical infrastructure of the Internet.

For the purposes of our analysis, an aggregation process has carried out. Initially, the IP links provided by DIMES were geo-coded at the city level. In order to homogenize and standardize the data, the IP links among European cities were aggregated at the regional NUTS^{iv}-3 level in such a way that the city-to-city links were aggregated to region-to-region links. It needs to be noted here that the intra-regional links derived from the aggregation process were also included in the subsequent econometric analysis. In addition, the IP links were also aggregated to the Dutch municipality level.

In general, although specific limitations exist in studying the Internet from a geographic perspective, the DIMES data set appears to be the richest available secondary data source with a geographic projection of the Internet infrastructure. Despite the above-mentioned limitations, the size of the DIMES experiment and the wide-spread locations of the DIMES agents enable us to safely use this data set, especially considering the general lack of geographic data on the Internet and the Internet infrastructure.

3. The Network Structure of CP

The first phase of the analysis focuses on exploring the network topology of the IP network. To support this analysis concepts and methods from the complex network analysis (CNA) field are utilized here. To provide a brief introduction, the ideas which underpin this section are derived from the *new science of networks* (Barabási 2002, Buchanan 2002, Watts 2003, Watts 2004), an analytical field which focuses on large-scale real-world networks and their universal, structural, and statistical properties (Newman 2003). Despite the strong focus of CNA on statistical physics, strong parallels exist between CNA and regional science, as the latter has a strong interest in networks (Cornell University 2010): while spatial economics and regional science focus on spatial structure, network analysis focuses on topological structure; and, while the former highlights the economic meaning of functional forms, the latter stresses the connectivity patterns of functional forms (Reggiani 2009, Reggiani and Nijkamp 2009). Using as a basis this conceptual parallel, CNA is used here as a tool to explore connectivity patterns in the topological configuration of the CP. The understanding of the latter is a vital step for continuing with the second part of our analysis, where the relation between physical distance and the topology of CP will be modeled.

Table 1 presents some basic network statistics for the CP for the years 2005 and 2008. Although it seems that the size of the CP has decreased over time, this refers to the part of the IP space captured by the DIMES project and not to the overall Internet. Thus, we cannot draw conclusions about the change of the size of the CP. Nonetheless, a change in the topology of the network is observable. Although less IP links were captured in 2008, an increase in the average and maximum *degree centrality* is observed, with the latter being almost doubled during the study period. Degree centrality is a connectivity measure and, in this case, is defined as the number of the accumulated IP links in each region discovered by the DIMES agents during the course of one year. Such a measure reflects the topology of the network. The vast increase in the maximum degree centrality is not reflected in the Gini coefficient, as the degree distribution was only slightly more uneven in 2008. Nonetheless, degree centrality appears to follow a highly uneven distribution in both years, while some form of hierarchy can be observed, with some regions performing network hub roles.

The outcome of the uneven distribution of the IP links among the European regions is an *efficient* CP. Indeed, the average network distance is remarkably short despite the very low density of the CP. In a network framework, distance does not refer to Euclidean distance, but to the number of nodes that separate any two nodes^v. For the case of CP, any two regions are separated on average by one intermediate node, which results in a network distance a little higher than 2. The latter is an indication of efficiency, as it reflects the ability of the network to transfer data flows with minimal routing, which involves less cost from the providers' point of view as well as faster and more secure communications for the Internet users.

The above qualities can be attributed to the *small world* (SW) characteristics of the CP. The latter refers to a widely used network model, whose main characteristic is the existence of highly-connected cliques, which gain global connectivity via a few links that span the entire network, linking distant clusters (Watts

and Strogatz 1998). This theoretical network model became popular because of its real-world applications. The CP resembles SW networks because of the short average distance – shorter than that observed in same size random networks (RN) – and the high clustering coefficient^{vi} – higher than that observed in same size RN.

Another crucial element of the SW networks is the degree distribution, which distinguishes this network type from another widely used type of network model known as *scale free* (SF). SF networks share the above characteristics with SW networks, but the degree distribution of their nodes follows power laws, contrary to the exponential laws which characterize SW networks. The different distributions reflect the difference between these two types of networks in terms of the nodes' heterogeneity: while the power-law degree distribution of the SF networks reflects the existence of a very few super-connected hubs and a vast majority of less-connected vertices (Barabási and Albert 1999), the exponential-degree distribution of SW networks resembles highly-connected cliques and less heterogeneous nodes^{vii}. Following Newman (2005), the estimation of the degree distribution curve is based on the *cumulative degree function* (CDF) derived from an inverse rank-plot graph. Figure 1 presents the CDFs for the years 2005 and 2008.

Insert Figure 1

The scatter plots reveal the existence of two different curves for both years: a straight line indicating a power law for the most-connected nodes of the IP network and a curve suggesting an exponential law for the least-connected nodes. This 'dual' character of the CDF exposes a *power law with a cutoff*, since the power law does not fit the overall distribution, but only the most-connected nodes. The above visual observation is supported statistically by curve estimations based on OLS and the relevant log-log transformations (Faloutsos *et al.* 1999, Gorman and Kulkarni 2004, Schintler *et al.* 2004, Patuelli *et al.*

2007, Reggiani *et al.* 2010, Tranos 2011). The results of the OLS are presented in Table 1, where three different laws are tested: exponential, power, and power with cutoff (Tanner function), accordingly:

$$, \tag{1}$$

$$, \tag{2}$$

$$\tag{3}$$

Insert Table 1

Indeed, the OLS results confirm the visual observation that Tanner functions better fit better the overall distribution for both years. So, at an aggregated NUTS-3 level, the European IP network does not form a clear SF structure. From a spatial perspective, this can be interpreted as an agglomeration effect of IP connectivity in a limited number of regions which act as hubs. At the same time, the exponential tail denoted the existence of a cluster of less-connected regions, which is more homogeneous in terms of IP connectivity than if a hierarchical and clear SF topology were present. Moving a step forward, the next section will explore the impact of physical and relational proximities on the formation of this complex network.

4. The Urban Geography of CP

Despite the fact that CP fails to form a clear SF topology – at least at this level of regional aggregation – there are still cities which perform very important hub roles for the function of this system. Table 2 presents the 25 most connected NUTS 3 regions in 2005 and 2008 according to the total number of IP links (weighted centrality). The results are not surprising. From one hand we can identify the main agglomerations of the European urban system in the top of the IP connectivity hierarchy: London, Amsterdam, Paris, Rome, Frankfurt, and Madrid. Some of these cities have been identified elsewhere as

the 'golden diamond' of the European Internet infrastructure (Tranos and Gillespie 2011). Indeed, London, Paris, Amsterdam and Frankfurt always appear to be in the top of the relevant European hierarchies. Stockholm is a very interesting case as it is consistently positioned second after London. This reflects its role as a communication hub for the IP network in Scandinavia and in Europe in general. In addition, it also reflects the importance of the knowledge economy – to the extent that the Internet infrastructure is related with this – in Scandinavia (e.g. EC 1999, Paci and Usai 2000, Cutrini 2009, Rodríguez-Pose and Tselios 2009b). Apart from these main European agglomerations, we also find in the group of the 25 most connected regions, cities such as Warsaw, Zurich, Brussels, Bern, Milan, Copenhagen and Wien. During the 4 year period we can observe differences as cities such as Berlin, Stuttgart and Helsinki have been replaced by Bucharest, Gothenburg and Outer London. In regard to the Romanian capital, this reflects the rapid increase of the digital economy in South-Eastern Europe. Athens and Sofia are only a few positions below in 2005 (32nd and 35th, respectively) gaining extensively in terms of IP connectivity during the study period (58th and 96th in 2005, respectively).

Insert Table 2

Furthermore, Outer London represents another phenomena: from one hand side we can observe a typical case of Modifiable Areal Unit Problem (MAUP - Openshaw 1984) which is related with the diverse size of the NUTS 3 regions in Europe and the non-systematic boundary building, and from the other side the spillover effect of the Internet connectivity. In regards to the former, the case of outer London indicates the artificial nature of the NUTS 3 boundaries as both Inner and Outer London regions refer to the same urban area. In general, the different size of NUTS 3 regions is a potential source of bias. Although such a limitation is inherent in any analysis adopting the NUTS classification, at the same time

this methodological choice is almost necessary for such a Pan-European analysis². Secondly, the above example highlights the spatial spillover effect of the IP connectivity. For the case of London, apart from Inner and Outer London, other adjacent regions of Cambridge, Hertfordshire and Surrey are also included in the list of the 25 most connected regions. The same applied for Paris (Hauts-de-Seine), Frankfurt (Offenbach, Landkreis), and for the polycentric regions of Randstadt (Amsterdam, The Hague and Utrecht). Apart from the MAUP, the above phenomenon is related with the location of this infrastructure at the micro-scale. Usually, the Internet physical links terminate in data-centers, which are buildings with specific requirements. These facilities are usually found in wider metropolitan areas, employing redundant buildings such as warehouses and department stores with high ceilings and high capacity power supply: they are found in locations which combine both access to high capacity backbone networks and closeness to customers in order for them to have physical access to their equipment. However, usually such facilities can neither afford the cost or find buildings with proper specifications in central locations (Evans-Cowley *et al.* 2002, Townsend 2003). Nowadays, it is also common to find collocation facilities in remote areas which combine access to backbone networks and low cost electric power. The discussion has also emerged for locating such facilities in areas where renewable energy is available or in areas with low temperatures to decrease the cooling cost (for this discussion see Arnaud 2009). An indication of the above spillover effect can be observed in the below maps (Figure 2). Using local indicators of spatial association (LISA) we identify clusters of regions with high IP connectivity. This observation is a first indication of the existence of spatial spillover effects in the installation of the digital infrastructure.

Insert Figure 2

² Alternative options include the aggregation at NUTS 2 level, but this would have resulted in much lower resolution analysis, and the use of Larger Urban Zones (Urban Audit in Eurostat 2011), which is not related with as extensive available statistical data as NUTS 3 regions.

5. CP and Proximity

In this section the impact of distance on the formation of CP is empirically tested. Distance is approached here both as a measure of cost and proximity. Starting from the cost perspective, as explained above, CP is a physical infrastructural layer and of course the installation of this infrastructure involves some cost. From the Internet Service Providers (ISPs – i.e. the owners of CP) point of view, the topology of their networks is the outcome of various factors. From one hand they try to minimize installation cost, maximize network efficiency and resilience and adjust the above to the individual business models and the geographic scope of each ISP. Such constraints include: (a) minimisation of the length of the installed fibre optics and the number of network edges because of the fibre installation cost; (b) creation of hubs in order to achieve economies of scope and scale; (c) avoidance of clear hub-and-spoke structures because of the increased vulnerability; (d) reduction of the use of switching points in order to increase network efficiency (minimize network latency); and (e) increase of presence in popular Internet Exchange Points (IXPs) in order to peer data with other ISPs (Tranos 2011). On the other hand, the IP network locations represent – to some extent – the ISPs perspective for the demand for such services. Given the above constraints, ISPs will install more infrastructure (i.e. IP connectivity) between places that according to their market expectations, will generate more demand for bilateral IP communications.

The above factor is the joint point with the notion of proximity as the underlying assumption behind our empirical analysis is that demand for such services will be increased between neighboring places. To justify this assumption we use argumentation borrowed from the French School of Proximity and Evolutionary Economic Geography. The former is a group of industrial economists, the main research objective of which is to endogenize space in economic analysis, and, more specifically, to incorporate space and other territorial proximity elements in a research framework, which aims to better

understand the dynamics of innovation (for a review of the French School of proximity, see Torre and Gilly 2000). A second development in further decomposing and analyzing the different components of proximity is research related to innovation and territorial learning in the broader framework of evolutionary economic geography. In recent years we have experienced an increased interest in factors which explain how firms and regions interact as part of a 'collective learning process', since learning and knowledge creation are essential components of the firms' and regions' competitive advantage (Boschma 2005). The notion of proximity and its different expressions are linked with ideas about knowledge transfer and creation, tacit knowledge, and learning regions (Boschma 2004). The common basis of these approaches is the importance of proximity in innovation creation: more proximate actors will interact more intensively and this will lead to more innovation production. However, it needs to be highlighted here that proximity in this context is not defined only as physical proximity but also in relational terms. While Torre and Rallet (2005) refer also to *organized* proximity apart from the geographical one, Boschma (2005) introduces more proximity components including on top of the above, *institutional*, *cognitive*, and *social* proximity. Unlike physical proximity, organized proximity is a relational notion, and refers to the ability of an organization to enhance interaction between its members. The main point behind this concept is that members of the same organization will interact together more intensively than actors outside the organization (Torre and Gilly 2000). Cognitive proximity is defined as the level of similarity of the knowledge base of different organizations (Nooteboom 2000, Balland 2011). Organizations collaborate and form links and networks using as criteria for their choices the knowledge background of the potential partners, as people and organizations, which share the same knowledge background and expertise may learn from each other (Boschma 2005). Social proximity lies on the idea that economic relations are embedded in a social context and this is why economic outcome is affected by social relations. Similarly to the above, the learning performance of an organization is affected by the social ties of its actors (Boschma 2005).

Finally, following North's (1993) definition, institutions are the amalgamation of formal rules and informal constraints including behavioral and social standards, while organizations can be approached as a group of agents performing the same activity. Put simply, organizations define agents' practices and strategies in the overall context provided by the institutional ecosystem in which they are positioned (Kirat and Lung 1999).

Although the focus of this chapter is the relation between CP and physical distance or, in other terms, geographical proximity, we can 'borrow' the conceptual work on the different proximity dimensions to build our hypothesis: demand for IP communications and consequently the structure of CP, to the extent that the latter meets the former, is negatively related with distance. This hypothesis is empirically tested in this section. In order to do so, the intensity of the NUTS 3 IP links derived by the DIMES project are utilized in a gravity model with panel data specifications. Before we present the econometric specifications and the results of the modeling process, the quantification of the different proximity types is discussed.

The key variable in our modeling exercise is geographic proximity. Although it appears to be the most straightforward and easy to understand dimension of proximity, different conceptualizations can be utilized as physical proximity might be affected not only by Euclidean distance, but also by the transportation cost between two places and their accessibility. In our case, geographical proximity is represented by the Euclidean distance between the centroids of two NUTS-3 regions. In addition, other variables are also included in the analysis to capture other facets of proximity. Following Hoekman et al. (2009), institutional proximity is defined on the basis of whether or not two regions are part of the same country. The underlying assumption is that two regions, which share the same institutional-country characteristics, will be characterized by a higher level of virtual interaction than two regions from different countries. Thus, the digital infrastructure between nearby places in institutional terms is

expected to be more intensive than between distant places. For this reason, a dummy variable is built with 0 denoting an IP link between two regions from a different country and 1 indicating an intra-country IP link. In addition, the adopted level of analysis enables us to capture even more detailed institutional characteristics, as the adopted city-region spatial unit is characterised by a degree of functional integration. On that basis, we may assumed that the location of two cities in the same region will positively affect the intensity of the digital infrastructure between them. For this reason, a second dummy variable is built with 1 indicating intra-regional IP links, and 0 vice versa. Finally, a set of dummy variables is introduced in order to capture the potential effect of different geographical areas inside the European territory such as IP links inside the Nordic region, the South and the Central-East Europe. These variables can be approached as a proxy for social proximity given the cultural similarities among countries from these areas.

The dependent variable in our proposed models is the intensity of IP links between NUTS 3 regions. As mentioned before, this is an infrastructural indicator, which does not represent the capacity of the digital infrastructure, but rather the number of IP links between places, as they were captured by the DIMES project. Of course, it would have been very interesting to have either the capacity of the installed infrastructure or the flows that run through these fiber-optic cables, but such data is not available at such fine-grained scale. Nonetheless, the intensity of the IP links between any two places provides good insight about the overall digital infrastructure – including bandwidth – installed between any two places.

Empirical results

In order to investigate the impact of physical distance and other types of proximity on the intensity of CP links the following generalized gravity model is estimated:

(1)

The dependent variable here is the natural logarithm of the intensity of IP links between any i and j during the period 2005-2008. The right-hand side variables reflect geographical proximity () as well as institutional proximity (,) and social proximity (and). In addition, represents an array of dummy variables controlling for unobserved country-to-country effects. In total, 34 European countries are included in the analysis. Out of the 561 potential country-to-country combinations^{viii}, 508 different country-to-country pairs have been derived from the NUTS 3 to NUTS 3 region links. Furthermore, IP_ln_{it} and IP_ln_{jt} are the masses of the Newtonian formula, and in this case they represent the natural logarithm of the weighted degree centrality of the IP links. Simply put, these variables present the total number of IP links originating or terminating in i and j . It should be noted here that, for the calculation of these variables, the IP links with non-European cities were also included in the analysis. This choice was made in order to better reflect the overall importance of i and j in CP. Finally, yearly effects t_2-t_4 are also included in the regressions, α_0 is the effect common to all years and pairs of regions, and , with μ_{ij} a bilateral random effect and u_{ijt} the idiosyncratic error.

Instead of estimating the above model cross-sectionally, a panel data specification is preferred. Firstly, panel data improves the researchers' ability to control for missing or unobserved variables (Hsiao 2003). Such an omitted-variable bias as a result of unobserved heterogeneity is a common problem in cross-section models. In addition, potential selection bias in IP links due to the traceroute process can be addressed more efficiently with panel data. In a nutshell, a panel data specification reduces the risk of obtaining biased estimators (Baltagi 2001).

While panel data introduces methodological gains, there are also shortcomings that need to be addressed. According to the literature (Wooldridge 2003), the most widely used panel data models are the *fixed effects* (FE) and *random effects* (RE). As the main aim of this chapter is to estimate the impact

of the different proximity measures on the creation of CP, it is preferred to use an RE model rather than an FE model, as because of the first differentiation process, the latter will result in the elimination of most of the proximity measures, because they are measured as time-invariant variables (e.g. Brun *et al.* 2005, Etzo 2011).

Table 3 presents the results of the different regressions. The first column presents a simple gravity model which uses only physical distance and masses. Distance has a negative impact on the intensity of IP connectivity, which is consistent in all other specifications tested and analyzed below. In simple terms, neighboring regions in physical terms are characterized by higher IP connectivity. With regard to institutional proximity, a positive impact can be observed on the formation of the CP. The results of our analysis indicate that IP connectivity is higher for intra-country links. Or, from a different perspective, *border effects* can be observed in installing IP infrastructure. Despite the importance of transnational IP links in gaining global connectivity, it seems that the number of IP addresses is positively affected by institutional proximity. In the same vein, an institutional effect also emerges at the regional level, suggesting localization effects on the structure of the IP network³. The latter result confirms the argumentation provided in Section 3 about the SW attributes of the CP. Nonetheless, it needs to be highlighted here that the focus of the analysis is the IP connectivity and the number of IP addresses, and not the actual capacity of these links. Finally, the proxy variables for social proximity have also a positive sign, indicating higher intensity between these areas. However, their impact becomes insignificant when the country-to-country effects are also included in the analysis (column 5, Table 3).

Insert Table 3

To increase the robustness of our analysis, an attempt is also made to analyze the role of distance on the formation of CP at the very detailed scale of the municipal level. As mentioned above, MAUP might

³ For intra-regional IP links the distance used in the gravity model is equal to the diameter of a circle, the perimeter of which is similar to the perimeter of the NUTS 3 region.

create bias in the above analysis because of the NUTS 3 classification. In order to overcome the potential bias, and also as a sensitivity analysis of the above model, the IP data was re-aggregated using as the Dutch municipalities as the areal unit. Based on this, a new simple gravity model is built and the results are presented below in Table 4.

Insert Table 4

Again, the impact of physical distance is tested along with the Newtonian masses. As before, the latter are presented by total amount of IP links terminating in each municipality, including the links with non-Dutch destinations. As it can be seen in Table 4, column 1, physical distance has again a significant negative impact even at this very fine-grained scale. To further explore this relationship, we also included a dummy variable indicating intra-municipality IP links. Just like in the previous gravity model, for this case we used as distance the diameter of a circle, the perimeter of which is similar to the perimeter of the municipality. The importance of intra-municipal IP links overshadows the impact of distance, and the latter has a negative impact, which is not statistically significant. However, this is not an indication that distance is not relevant at such a small scale. On the contrary, distance is related with the location of IP links inside the borders of a municipality, because such links are always of short distance. Interestingly enough, the boarder effect is strong in this case than distance itself.

To summarize the above multi-scalar results, IP connectivity appears to be higher between neighboring. Border and localization effects become significant, even for the digital infrastructure.

6. Conclusions

Although spatiality is usually not an issue in the discussion around the Internet, our study aims to address this viewpoint, approaching the Internet infrastructure from a spatial perspective, despite the

technical and conceptual difficulties that such an exercise involves. The novelty of this chapter lies not only in the spatial perspective adopted in the analysis, but also in the effort to quantify these issues through an evidence-based modeling approach. While the discussion about the relation between the Internet and real-world geography is not new in the literature, hard evidence deriving from a modeling exercise has not yet been extensively utilized. For instance, while it is clear nowadays that the ‘end of geography’ discussion was not a valid proposition, convincing empirical evidence has still not yet been provided.

Our analysis reveals that Tobler’s First Law of Geography is valid in the CP. The intensity of IP connectivity is higher between neighboring regions indicating the role of physical distance in the formation of the CP. The latter can be approached both as an indication of cost in physical connectivity (Waxman 1988) and also as an indication of higher demand in IP communications between nearer destinations. Our results are in agreement with previous studies which focus on the interlink between physical distance and the digital infrastructure (D'Ignazio and Giovannetti 2007). In addition, in the present chapter the notion of distance is extended in relational terms. According to this, CP is not only affected by costs deriving from physical distance, but also by costs related with institutional proximity.

Drawing on the results of our analysis, specific spatial processes can be identified. Firstly, centripetal forces agglomerate IP links in specific locations, which act as the hubs of this digital infrastructure. During the four-year study period, the uneven distribution of IP connectivity has – marginally – increased. But, overall, CP appears to be strongly curved by agglomeration forces. In addition, spillover effects can be observed, as high IP connectivity can also be found around metropolitan NUTS 3 regions. On the other hand, centrifugal forces ‘protect’ the less-connected regions, securing a level of connectivity which would not be observed if clear SF structures were utilized. Other forces, including the

– limited – provision of IP connectivity from non-private ISPs, ensure that less-connected regions are not as thinly connected as an SF structure would indicate.

To conclude, despite the inherent difficulty to approach the Internet from a spatial perspective, the above analysis reveals not only the spatial dimension of the Internet, but also the impact of physical distance on the structure of this complex spatial network. Further research needs also to be done in order to understand the impact of more relational facets of distance and proximity. Such an exercise will enable us to further shed light into the complexity of the Internet by approaching it from a relational point of view.

References

- Adamic, L.A. & Huberman, B.A., 2002. Zipf's law and the internet. *Glottometrics*, 3, 143-150.
- Albert, R. & Barabási, A.-L., 2002. Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74, 47-97.
- Arnaud, B.S., 2009. *Green it/broadband and cyber-infrastructure* [online]. <http://green-broadband.blogspot.com/> [Accessed Access Date 2009].
- Bakis, H., 1981. Elements for a geography of telecommunication. *Geographic Research Forum*, 4, 31-45.
- Balland, P.-A., 2011. Proximity and the evolution of collaboration networks: Evidence from research and development projects within the global navigation satellite system (gnss) industry. *Regional Studies*, forthcoming.
- Baltagi, B.H., 2001. *Econometric analysis of panel data*, 2nd ed. Chichester: John Wiley & Sons.
- Barabási, A.-L., 2002. *Linked: The new science of networks* Cambridge, MA: Perseus.
- Barabási, A.-L. & Albert, R., 1999. Emergence of scaling in random networks. *SCIENCE*, 286 (15 OCTOBER), 509-512.
- Batty, M., 1997. Virtual geography. *Futures*, 29 (4/5), 337-352.
- Boschma, R., 2004. Competitiveness of regions from an evolutionary perspective. *Regional Studies*, 38 (9), 1001-1014.
- Boschma, R., 2005. Proximity and innovation: A critical assessment. *Regional Studies*, 39 (1), 61-74.
- Brun, J.-F., Carrere, C., Guillaumont, P. & Melo, J.D., 2005. Has distance died? Evidence from a panel gravity model. *The World Bank Economic Review*, 19 (1), 99-120.
- Buchanan, M., 2002. *Nexus: Small worlds and the groundbreaking science of networks* New York: W. W. Norton.
- Cairncross, F., 2001. *The death of distance 2.0* London: Texere Publishing Limited.
- Carmi, S., Havlin, S., Kirkpatrick, S., Shavitt, Y. & Shir, E., 2007. A model of internet topology using k-shell decomposition. *PNAS*, 104 (27), 11150-11154.
- Chaturin, B., 1988. *What am i doing here?* London: Penguin.

- Cohen-Blankshtain, G. & Nijkamp, P., 2004. The appreciative system of urban ict policies: An analysis of perceptions of urban policy makers. *Growth and Change*, 35 (2), 166-97.
- Cohen, G., Salomon, I. & Nijkamp, P., 2002. Information-communications technologies (ict) and transport: Does knowledge underpin policy? *Telecommunication Policy*, 26 (1-2), 31-52.
- Cornell University, 2010. *About regional science* [online]. <http://aap.cornell.edu/crp/programs/regsci/> [Accessed Access Date 2010].
- Cutrini, E., 2009. Specialization and concentration from a twofold geographical perspective: Evidence from europe. *Regional Studies*, forthcoming.
- D'ignazio, A. & Giovannetti, E., 2007. Spatial dispersion of interconnection clusters in the european internet *Spatial Economic Analysis*, 2 (3), 219-36.
- Dimes, 2010. *Dimes project* [online]. <http://www.netdimes.org/new/> [Accessed Access Date 2010].
- Dodge, M. & Zook, M.A., 2009. Internet based measurement. In Kitchin, R. & Thrift, N. eds. *The international encyclopedia of human geography*. Elsevier.
- Drucker, P.F., 1998. From capitalism to knowledge society. In Neef, D. ed. *The knowledge economy*. Woburn, MA: Butterworth-Heinemann.
- Ec, 1999. *Esdp - european spatial development perspective*. Luxembourg.
- Etzo, I., 2011. The determinants of the recent interregional migration flows in italy: A panel data analysis. *Journal of Regional Science*, forthcoming.
- Eurostat, 2011. *Regional data* [online]. <http://epp.eurostat.ec.europa.eu/portal/page/portal/eurostat/home/> [Accessed Access Date 2010].
- Evans-Cowley, J., Malecki, E.J. & Mcintee, A., 2002. Planning responses to telecom hotels: What accounts for increased regulation of co-location facilities? *Journal of Urban Technology*, 9 (3), 1-18 Available from: <Go to ISI>://000180572300002
- Faloutsos, M., Faloutsos, P. & Faloutsos, C., 1999. On power-law relationships of the internet topology. *Computer Communication Review*, 29, 251-262.
- Gilder, G. 1995 *Forbes* ASAP, February 27.
- Gorman, S.P. & Kulkarni, R., 2004. Spatial small worlds: New geographic patterns for an information economy. *Environment and Planning B*, 31 (2), 273-296 Available from: <Go to ISI>://000232718600007.
- Graham, S., 1998. The end of geography or the explosion of place? Conceptualizing space, place and information technology. *Progress in Human Geography*, 22 (2), 165-185.
- Hepworth, M., 1989. *Geography of the information economy* London: Belhaven Press.
- Hoekman, J., Frenken, K. & Van Oort, F., 2009. The geography of collaborative knowledge production in europe. *The Annals of Regional Science*, 43 (3), 721-738.
- Hsiao, C., 2003. *Analysis of panel data*, 2nd ed. Cambridge: Cambridge University Press.
- Kellerman, A., 1993. *Telecommunications geography* London: Belhaven Press.
- Kirat, T. & Lung, Y., 1999. Innovation and proximity : Territories as loci of collective learning processes. *European Urban and Regional Studies*, 6 (1), 27-38.
- Kirsch, S., 1995. The incredible shrinking world? Technology and the production of space. *Environment and Planning D: Society and Space*, 13 (5), 529-555.
- Kitchin, R., 1998a. *Cyberspace* New York: Wiley.
- Kitchin, R., 1998b. Towards geographies of cyberspace. *Progress in Human Geography*, 22 (3), 385-406.
- Kolko, J., Year. The death of cities? The death of distance? Evidence from the geography of commercial internet usageed.^eds. *Selected Papers from the Telecommunications Policy Research Conference 1999*, Newcastle.
- Latora, V. & Marchiori, M., 2001. Efficient behavior of small-world networks. *Physical Review Letters*, 87.

- Malecki, E.J., 2002. The economic geography of the internet's infrastructure. *Economic Geography*, 78 (4), 399-424.
- Malecki, E.J., 2004. Fibre tracks: Explaining investment in fibre optic backbones. *Entrepreneurship and Regional Development*, 16 (1), 21-39 Available from: <Go to ISI>://000220253300003.
- Malecki, E.J. & Gorman, S.P., 2001. Maybe the death of distance, but not the end of geography: The internet as a network. In Leinbach, T.R. & Brunn, S.D. eds. *Worlds of e-commerce: Economic, geographical and social dimensions*. West Sussex: Wiley, 87-105.
- Mitchell, W.J., 1995. *City of bits: Space, place and the infobahn* Cambridge, MA: MIT Press.
- Moss, M.L. & Townsend, A., 1997. Tracking the net: Using domain names to measure the growth of the internet in us cities. *Journal of Urban Technology*, 4 (3), 47-60 Available from: <Go to ISI>://000071345200003.
- Newman, M.E.J., 2003. The structure and function of complex networks. *SIAM Review*, 45, 167-256.
- Newman, M.E.J., 2005. Power laws, pareto distributions and zipf's law. *arXiv:cond-mat/0412004v3*.
- Nooteboom, B., 2000. Learning by interaction: Absorptive capacity, cognitive distance and governance. *Journal of Management and Governance*, 4 (1-2), 69-92.
- Nooy, W.D., Mrvar, A. & Batagelj, V., 2005. *Exploratory social network analysis with pajek*: Cambridge University Press.
- North, D.C., 1993. Institutional change: A framework of analysis. In Sjöstrand, S.E. ed. *Institutional change: Theory and empirical findings*. New York: M.E. Sharpe.
- Openshaw, S., 1984. The modifiable areal unit problem. *Concepts and Techniques in Modern Geography*, 38, 41.
- Paci, R. & Usai, S., 2000. Technological enclaves and industrial districts: An analysis of the regional distribution of innovative activity in europe. *Regional Studies*, 34 (2), 97-114.
- Pastor-Satorras, R. & Vespignani, A., 2004. *Evolution and structure of the internet* Cambridge: Cambridge University Press.
- Patuelli, R., Reggiani, A., Gorman, S.P., Nijkamp, P. & Bade, F.-J., 2007. Network analysis of commuting flows: A comparative static approach to german data. *Networks and Spatial Economics*, 7, 315-331.
- Reggiani, A., 2009. From complexity to simplicity. In Reggiani, A. & Nijkamp, P. eds. *Complexity and spatial networks*. Berlin: Springer-Verlag.
- Reggiani, A., Bucci, P. & Russo, G., 2010. Accessibility and network structures in the german commuting. *Networks and Spatial Economics*, 1-21 Available from: <http://dx.doi.org/10.1007/s11067-010-9149-0>.
- Reggiani, A. & Nijkamp, P., 2009. *Complexity and spatial networks* Berlin: Springer-Verlag.
- Reggiani, A. & Vinciguerra, S., 2007. Network connectivity models: An overview and applications to the space-economy. In Friesz, T. ed. *Network science, nonlinear science and infrastructure systems*. New York: Springer-Verlag.
- Rodríguez-Pose, A. & Tselios, V., 2009b. Mapping regional personal income distribution in western europe: Income per capita and inequality. *Czech Journal of Economics and Finance*, 5 (1), 41-70.
- Schafer, A., 1988. The global reward for motorized mobility. *Transportation Research A*, 32 (6), 455-477.
- Schintler, L., Gorman, S.P., Reggiani, A., Patuelli, R., Gillespie, A., Nijkamp, P. & Rutherford, J., 2004. Complex network phenomena in telecommunication systems. *Networks and Spatial Economics*, 5 (4), 351-370.
- Shavitt, Y. & Shir, E., 2005. Dimes: Let the internet measure itself. *ACM SIGCOMM Computer Communication Review*, 35 (5), 71-74.
- Star, S.L., 1999. The ethnography of infrastructure. *American Behavioral Scientists*, 43 (3), 377-91.

- Swyngedouw, E., 1993. Communication, mobility and the struggle for power over space. In Giannopoulos, G. & Gillespie, A. eds. *Transport and communications in the new europe*. London: Belhaven, 305-325.
- Tasse, G., 1992. *Technology infrastructure and competitive position* Boston, MA: Kluwer Academic Press.
- Tasse, G., 2008. Modeling and measuring the economic roles of the technology infrastructure. *Economics of Innovation and New Technology*, 17 (7&8), 615-29.
- Toffler, A., 1980. *Third way* New York: William Morrow.
- Torre, A. & Gilly, J.-P., 2000. On the analytical dimension of proximity dynamics. *Regional Studies*, 34 (2), 169-180.
- Torre, A. & Rallet, A., 2005. Proximity and localization. *Regional Studies*, 39 (3), 47-59.
- Townsend, A., 2003. *Wired / unwired: The urban geography of digital networks*. MIT.
- Tranos, E., 2010. *The geography of the internet infrastructure in europe*. Newcastle University.
- Tranos, E., 2011. The topology and the emerging urban geographies of the internet backbone and aviation networks in europe: A comparative study *Environment and Planning A*, 43 (2), 378-92.
- Tranos, E. & Gillespie, A., 2011. The urban geography of internet backbone networks in europe: Roles and relations. *Journal of Urban Technology*, 18 (1), 35– 49.
- Tranos, E. & Nijkamp, P., forthcoming. *Cyber-place and proximities: A test on quantitative patterns*. Amsterdam: VU, Dept. of Spatial Economics.
- Vinciguerra, S., Frenken, K. & Valente, M., 2010. The geography of internet infrastructure: An evolutionary simulation approach based on preferential attachment. *Urban Studies*, 47 (9), 1969–1984.
- Watts, D.J., 2003. *Six degrees: The science of a connected age* New York: W. W. Norton & Company.
- Watts, D.J., 2004. The 'new' science of networks. *Annual Review of Sociology* 30, 243-270.
- Watts, D.J. & Strogatz, S.H., 1998. Collective dynamics of small-world networks. *Nature*, 393, 440-442.
- Waxman, B.M., 1988. Routing of multipoint connections. *IEEE Journal on Selected Areas in Communications*, 6 (9), 1617 - 1622.
- Wheeler, D.C. & O'Kelly, M.E., 1999. Network topology and city accessibility of the commercial internet. *Professional Geographer*, 51 (3), 327-339.
- Wooldridge, J.M., 2003. *Introductory econometrics: A modern approach*, 2nd ed. Mason, Ohio: South-Western.

Table 1: Network analysis

Network statistics	Degree		distribution fit	Degree		
	2005	2008		2005	2008	
European nodes	1376	1276	Exponential	R ²	0.679	0.632
Intra-European IP links	23352	19521		Coef.	0.0003	0.0002
Av. degree *	1084	1490	Power	R ²	0.733	0.712

Max. degree *	44313	77692		Coef.	-0.481	-0.435
Gini coef.	0.727	0.741	Tanner function	R ²	0.909	0.889
Density *	0.024	0.023		Power Coef.	-0.323	-0.305
Av. dist. *	2.295	2.176		Exp. Coef.	-0.0002	-0.0001
<i>Av. dist. RN</i>	<i>2.831</i>	<i>2.891</i>	* for these metrics, links between Europe and the rest of the world were also included in the analysis.			
CC*	0.71	0.69				
<i>CC RN</i>	<i>0.012</i>	<i>0.012</i>				

Table 2: IP ranks

rank	2005	2008
1	Inner London - East (London)	Inner London - East (London)
2	Stockholms län (Stockholm)	Stockholms län (Stockholm)
3	Groot-Amsterdam (Amsterdam)	Groot-Amsterdam (Amsterdam)
4	Paris (Paris)	Frankfurt am Main, Kreisfreie Stadt (Frankfurt)
5	Roma (Rome)	Offenbach, Landkreis (adjacent to Frankfurt)
6	Cambridgeshire CC (Cambridge)	Paris (Paris)
7	Frankfurt am Main, Kreisfreie Stadt (Frankfurt)	Madrid (Madrid)
8	Zürich (Zürich)	Roma (Rome)
9	Madrid (Madrid)	Milano (Milan)
10	Arr. de Bruxelles-Capitale (Brussels)	Zürich (Zürich)
11	Milano (Milan)	Miasto Warszawa (Warsaw)
12	Offenbach, Landkreis (adjacent to Frankfurt)	Cambridgeshire CC (Cambridge)
13	Agglomeratie 's-Gravenhage (Den Haag)	Bucureşti (Bucharest)
14	Uusimaa (Helsinki)	Outer London - West and North West (London)
15	Stuttgart, Stadtkreis (Stuttgart)	Agglomeratie 's-Gravenhage (Den Haag)
16	Bern (Bern)	Oldenburg (Oldenburg)
17	Østjylland (Aarhus)	Surrey (Surrey)
18	Surrey (Surrey)	Københavns omegn (Copenhagen)
19	Wien (Wien)	Bern (Bern)
20	Berlin (Berlin)	Hauts-de-Seine (adjacent to Paris)
21	Københavns omegn (Copenhagen)	Västra Götalands län (Gothenburg)
22	Oldenburg (Oldenburg)	Hertfordshire (adjacent to London)

23 Miasto Warszawa (Warsaw)

Arr. de Bruxelles-Capitale (Brussels)

24 Hauts-de-Seine (adjacent to Paris)

Utrecht (Utrecht)

25 Hertfordshire (adjacent to London)

Wien (Wien)

Table 3: Panel data regressions on the intensity of IP links (natural logarithm)

	(1)	(2)	(3)	(4)	(5)
dist_ln	-0.939 (0.008)***		-0.361 (0.009)***	-0.391 (0.009)***	-0.353 (0.011)***
inter		3.39 (0.041)***	2.87 (0.043)***	2.64 (0.042)***	2.628 (0.042)***
cntr		2.334 (0.018)***	1.828 (0.022)***	1.716 (0.022)***	4.631 (2.051)**
south				1.02 (0.030)***	0.969 (1.75)
nordic				1.062 (0.045)***	2.647 (2.404)
se				1.031 (0.032)***	4.345 (2.901)
ip_o_ln	0.428 (0.005)***	0.535 (0.005)***	0.543 (0.005)***	0.567 (0.005)***	0.613 (0.005)***
ip_d_ln	0.377 (0.005)***	0.536 (0.005)***	0.542 (0.005)***	0.571 (0.005)***	0.616 (0.005)***
t1					0.356 (0.014)***
t2					0.121 (0.014)***
t3					0.066 (0.014)***
Constant	1.901 (0.059)***	-7.291 (0.067)***	-4.889 (0.090)***	-5.183 (0.089)***	-8.262 (2.397)***
Observations	83700	83700	83700	83700	83700
Number of link	44518	44518	44518	44518	44518

Standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 4: Cross section regressions on the intensity of IP links (natural logarithm) for The Netherlands

	(1)	(2)
dist_ln	-0.356 (0.040)***	-0.07 (0.046)
intramuni		1.743 (0.154)***
ip_o_ln	0.268 (0.015)***	0.298 (0.014)***
ip_d_ln	0.384 (0.016)***	0.392 (0.016)***
Constant	1.48 (0.432)***	-2.036 (0.521)***
Observations	1953	1953
R-squared	0.312	0.355

Standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

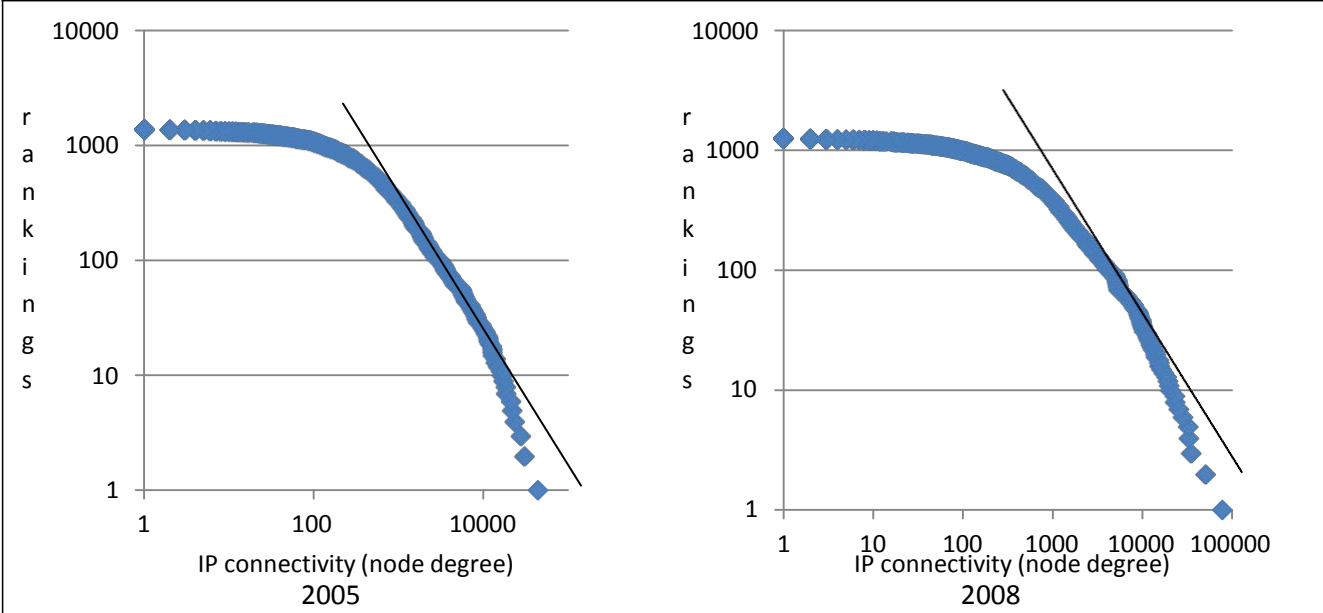


Figure 1: Cumulative degree distribution of NUTS-3 regions based on IP links

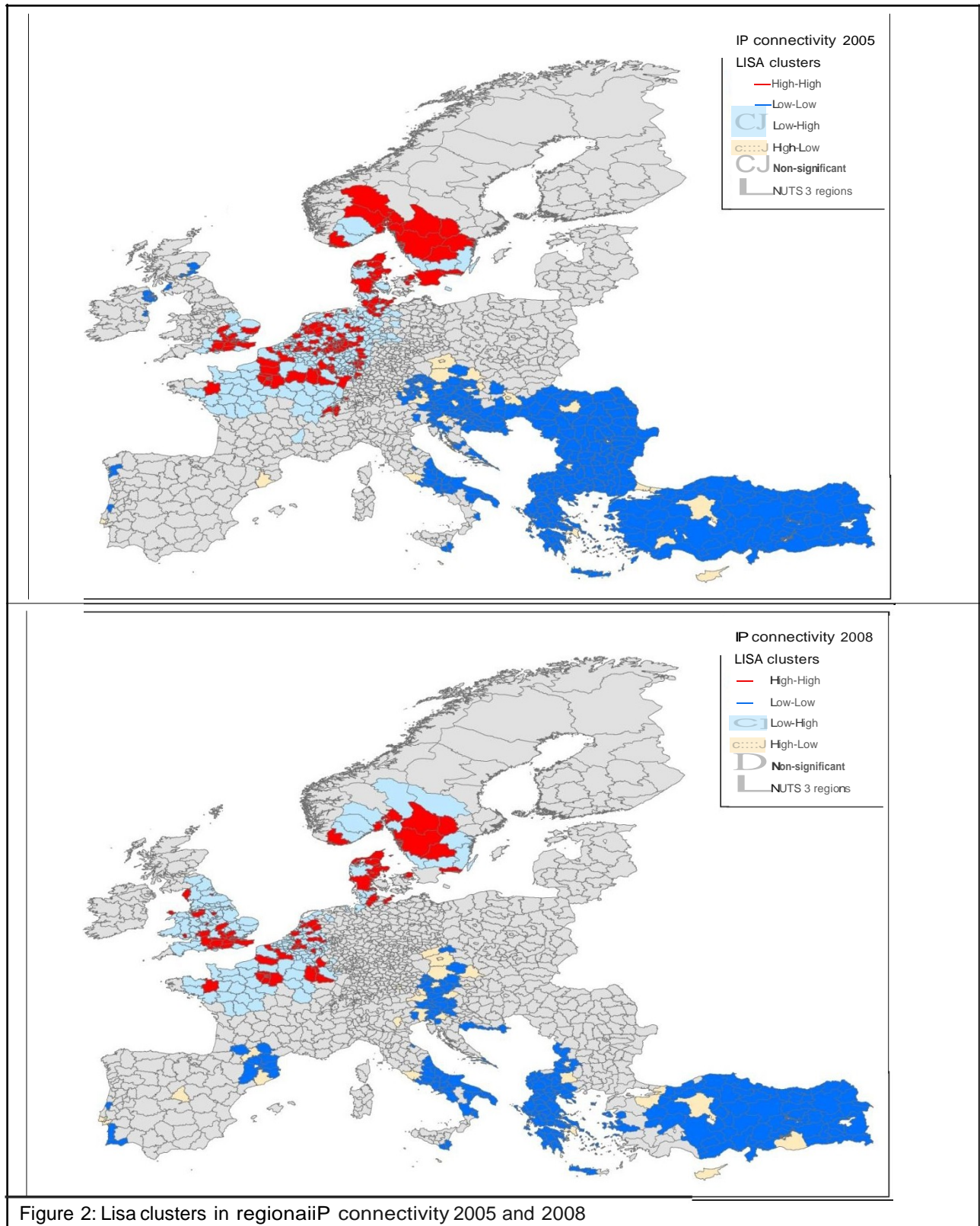


Figure 2: Lisa clusters in regional IP connectivity 2005 and 2008

ⁱ Exceptions include the work of Gorman and Kulkarni (2004), Schintler et al. (2004), Vinciguerra et al. (2010), and Tranos (2011).

ⁱⁱ Traceroutes are specific programs, which map the route that a data packet follows through different nodes in order to reach its final destination (Dodge and Zook 2009).

ⁱⁱⁱ These links function at level 3 of the OSI model. As noted elsewhere (Tranos 2010), the first three layers of the OSI model represent physical infrastructural capital, while the four highest layers reflect 'infrotechnologies' (Tassey 1992, Tassej 2008).

^{iv} NUTS is the French acronym for the Nomenclature for Territorial Units of Statistics, and NUTS-3 is the most detailed level usually representing a province.

^v Because there are usually numerous different ways to connect any two given nodes (known as *walks*), research commonly focuses on the shortest path, known as distance (Nooy *et al.* 2005).

^{vi} The clustering coefficient of node i is the ratio between the number of edges E_i that exist among its nearest neighbours (nodes which are directly connected with node i) and the maximum number of these edges, where k_i is the number of nodes in clique i : $C_i = \frac{2E_i}{k_i(k_i - 1)}$ (Latora and Marchiori 2001).

^{vii} For a review of the new science of networks from a spatial economics perspective, the reader is referred to Reggiani and Vinciguerra (2007), and, for an application of CNA on the Internet infrastructure, to Tranos (2011).

^{viii} The maximum number of edges in a network is $= n(n-1)/2$.