

Making the Web-Science Operational for Interregional Commuting Analysis: Evidence from Greece

Dimitrios Tsiotas¹  · Serafeim Polyzos¹

Received: 8 March 2014 / Accepted: 23 June 2015 / Published online: 9 July 2015
© Springer Science+Business Media New York 2015

Abstract This article proposes a Network Theoretical model for the interregional commuting analysis, estimating the volume of commuters that operate in a commuting network, based on the conceptual framework of the term “Web,” as provided by the Web-Science. Initially, the Web-Science’s conceptual framework is being generalized to operate for any kind of network and, afterwards, the proposed model is constructed, under the rationale that each conceptual component in the generalized Network Theoretical framework should be represented by different variables. The applicability of the proposed model is evaluated on real data of the Greek interregional commuting system, formulated by the capital cities of the Greek non-insular prefectures ($|V(G)|=n=39$) and by their respective interconnection axes ($|E(G)|=m=72$). The results of the analysis show respect to the categories of the proposed conceptual framework, where the overall empirical analysis of the Greek interregional commuting network elected the population variable as the most determinative factor for the commuting phenomenon.

Keywords Network analysis · Network theory · Interregional commuting · Web-science

Introduction

Commuting is generally defined as the act of daily intercity traveling for employment purposes, and it suggests a multi-parametric phenomenon determined by a set of socio-economic and geopolitical factors (Polyzos 2011; Polyzos et al. 2013). The socioeconomic and geopolitical framework describing commuting renders the analysis and thus the comprehension of this phenomenon a very important procedure for the effective and sustainable transportation planning and policy (Evans et al. 2002; Van Ommeren and Rietveld 2005).

✉ Dimitrios Tsiotas
dtsiot@yahoo.gr
Serafeim Polyzos
spolyzos@uth.gr

¹ Department of Engineering in Regional and Urban Planning, University of Thessaly, Pedion Areos, Volos, P.O. 38.334, Karditsa, Greece

The multivariate nature of commuting obviously sets the holistic microeconomic analysis of this phenomenon to be a particularly complex procedure, a fact that has favored so far mostly the decomposition approach in the research field. Many aspects of the commuting phenomenon have been studied so far by regional researchers, such as suggestively are *transportation (distance and time) cost* (Van Ommeren and Fosgerau 2009; Tsiotas and Polyzos 2013a), *travel stress* (Koslowsky et al. 1995), *travel accident probability* (Ozbay et al. 2007), *transportation and route alternatives* (Murphy 2009; Liu and Nie 2011), and *productivity changes* (Van Ommeren and Rietveld 2005).

One of the modern scientific sectors capable in providing a modeling toolkit for the macroscopic geographic and socioeconomic analysis of spatial communication systems is the so called (complex) *Network Analysis* (Brandes and Erlebach 2005; Barthelemy 2011; Tsiotas and Polyzos 2014, 2015). This recently established analytical framework, and in particular *Social Network Analysis* (Easley and Kleinberg 2010), became extremely popular during the rapid sprawl of the *Social Network Sites (SNS)* (Kalantzi and Tsiotas 2012) through the *World Wide Web (WWW)* (Berners-Lee et al. 2007; Tsiotas and Vafopoulos 2010), which managed today to overcome the initial borders of its maternal science, *Sociology*, and to introduce an individual scientific sector cited by many researchers as *Network Theory* (Zuckerman 1999; Easley and Kleinberg 2010; Borgatti and Halgin 2011).

Network Theory interprets communication systems as *networks* $G(V,E)$ (Easley and Kleinberg 2010; Borgatti and Halgin 2011; Tsiotas and Polyzos 2013a), which are being represented as a set of interconnected entities, called *vertices* or *nodes* $V(G)$, with their connections, called *edges* or *links* $E(G)$. Whether considering that the existence of communication systems is diachronic in human history (Apicella et al. 2012), then a network does not suggest a modern concept. Nevertheless, the modern *post-Social Network Theoretic* aspect of networks, expressing the everyday life's experience of "being connected" (Christakis and Fowler 2009), has transformed these communication structures into a modern way of thought. According to the above conceptual consideration of networks, the interregional commuting system can also suggest a network, since at this level of scale it refers to a set of interconnected cities that daily communicate by exchanging commute labor potential.

This article proposes a Network Theoretical model for the interregional commuting analysis, which estimates the number of commuters that operate in a commuting network, based on the conceptual framework provided by the *Web-Science* for the structural decomposition of the Web, as it was presented by Tsiotas and Polyzos (2013b). The proposed model is being constructed on the interregional communication system consisting of the 39 non-insular prefectures of Greece, and its applicability and effectiveness are tested on real commuting and socioeconomic data. The conceptual width of the model's components renders generality to the utility of the proposed model that is believed to be capable to constitute an effective tool for the principal analytic manipulation of a network under an insignificant loss of information.

This paper is organized as follows: "**Methodology**" presents the methodological part: the conceptual framework for network analysis is described, the interregional commuting model is constructed, and the necessary quantitative framework is presented. "**Results and Discussion**" presents and discusses the results of the analysis, under a regional economic perspective. Finally, in "**Conclusions**," conclusions are given.

Methodology

The proposed interregional commuting model of this paper is based on the generalized conceptual framework of the term “network,” originated from the theoretical matter of the Web-Science (Berners-Lee et al. 2007), as it was introduced by Tsiotas and Polyzos (2013b). This conceptual framework provides decomposition rationale that leads to the construction of a linear regression model, capable to describe the Greek interregional commuting network under an insignificant loss of information.

A New Conceptual Framework for Network Theory

Recently, Tsiotas and Polyzos (2013b) proposed a generalized *Web-Science-based* conceptual framework for networks. The rationale of this framework, illustrated at Fig. 1, is based on the consideration that the theoretical matter of the Web is wide enough to provide a generalized documentation for the description of all kinds of networks (physical or immaterial), since it represents perhaps the most advanced, developed, complex, and complete socio-technological communication system of global scale that has been diachronically made by humans.

According to the perspective of the Web-Science, as it has been established and it is expressed by Berners-Lee et al. (2007), the Web constitutes a multilayer network composed by three ordered sub-network components (layers), the *Internet*, the World Wide Web, and the *Society* (Fig. 1a). The undermost component, called the Internet, consists of the *interconnected hardware devices* (PC's, servers, routers, cables, links, etc.) that materialize the physical part of the Web. The middle component of the Web consists of the *software interconnections*, referring to the total of the *operational systems, browsers, applications, interfaces*, etc. that are responsible for the communication conduct in the Web. This component is called the World Wide Web (WWW). Finally, the upper component refers to the Society, which is the living matter of the Web that meets the utility of the previous underling structures.

The corresponding conceptual framework provided by Network Theory (Easley and Kleinberg 2010) distinguishes in every network two components: a *structural* and a *behavioral* (or *functional*) (Fig. 1a). The structural component consists of the set of the infrastructures and their supporting facilities and suggests the constructed background that materializes the communication among the interconnected entities of the network. All the other factors that are related with the conduct of communication, expressing flows, signal processing, and generally the “moving matter” of the network suggest the *behavioral component*, referring to the *mobile* and *dynamic* part of the network.

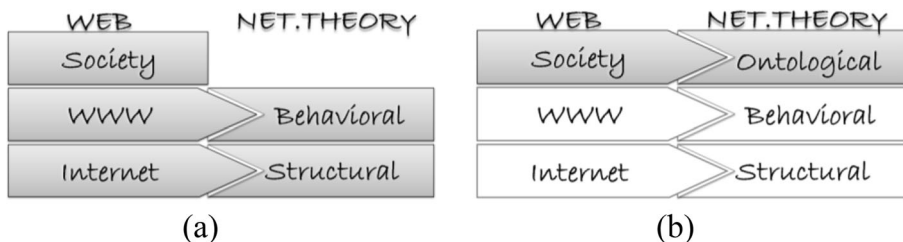


Fig. 1 a Conceptual components for the terms “Web” and “network.” b The proposed expansion of the conceptual framework “network”

The above classification, provided by Easley and Kleinberg (2010), enjoys the generality to formulate a primary theoretical framework for introducing the notion of network, but it seems insufficient to distinctly refer to the *social origin* that shaped the modern analytical aspect and boosted the development of Network Theory. Perhaps some indications of this social gender can be traced within the conceptual content of the behavioral component, but such indications are latent and cannot be considered as attributive, since similar indications can be also traced within the concept of the structural component.

Whether considering the previous two conceptual frameworks (Berners-Lee et al. 2007; Easley and Kleinberg 2010) under a jointed perspective, then the semantic correspondences *structural-internet* and *functional-Web* can be shaped, where Tsiotas and Polyzos (2013b) introduced a conceptual component, named “ontological,” shaping a third pair of correspondence *society-ontological*. According to this rationale, the ontological component refers to the *society of the interconnected entities* in a network and includes the set of its surrounding *socioeconomic, cultural, ethic, political, and cognitive* attributes.

Graph Theoretical Model and Data

The Greek interregional commuting network is modeled by a non-directed bond graph (Fig. 2) $G(V,E)$ (Tsiotas and Polyzos 2013a, b), where the set of vertices represent the capital cities of Greek regions and the set of edges represent distances. In particular, every terrestrial prefecture P_i ($i=1,\dots,39$) is projected to a vertex ($P_i \rightarrow v_i$) $v_i \in V$ that is located at the geographic center of each capital city and their direct transportation road connections are drawn as edges $e_{ij} \in E$. Each edge is described by a weight value w_{ij} expressing the kilometric or spacetime (necessary time, in minutes, for covering the certain edge) distances. The available edge data is organized in diagonal adjacency matrices $D^{39 \times 39}$ (Diestel 2005) and concern records of the year 2010 (Tsiotas et al. 2012).

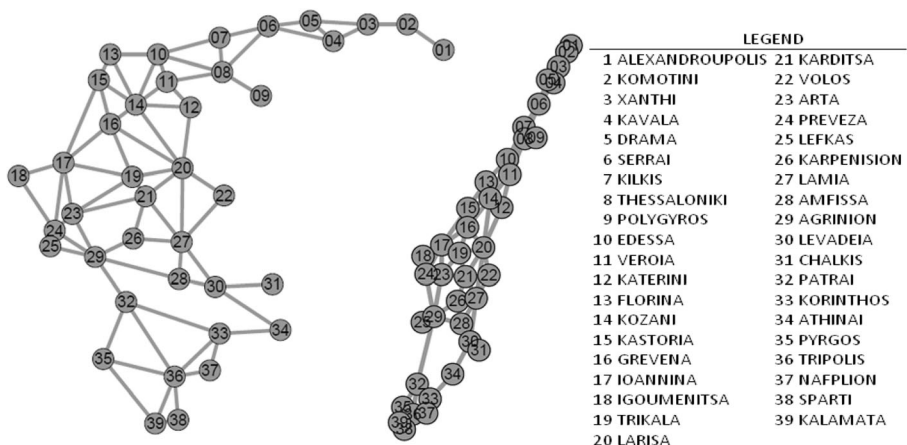


Fig. 2 The Graph Theoretical model of the Greek interregional commuting system (left) and its Force Atlas (Bastian et al. 2009) transformation (right)

The Greek interregional commuting network is considered additionally to edge—also as node-weighted (Tsiotas and Polyzos 2013a), meaning that each node $v_i \in V(G)$ is described by a set of weight values $w_i = \{w_{ij}\}$ that correspond to node socioeconomic attributes, which operate accordingly to the way that the degree value (Diestel 2005) describes a node. This consideration renders the set of nodes in the interregional commuting network into a vector space of dimension $|V(G)|=39$, where the sets with the node weights for each attribute are treated as separate statistical variables.

Under this treatment, a group of 30 variables (Y, X_1, \dots, X_{29}) was formulated for the analysis of the Greek interregional commuting network, which were collected to cover (as possible) the very range of the determining commuting factors, as it derives from the literature review (Glaeser and Kohlhase 2003; Clark et al. 2003; Ozbay et al. 2007; Van Ommeren and Fosgerau 2009; Murphy 2009; Liu and Nie 2011; Polyzos 2011). The available variables to the analysis were further grouped into structural, behavioral, and ontological, in accordance to the conceptual framework presented previously.

Table 1 shows the variables participated to the Greek interregional commuting analysis with their descriptions, measures, and references per case.

Proposing an Interregional Commuting Model

This paper proposes a Network Theoretical model for the interregional commuting analysis, estimating the volume of commuters that operate in a commuting network, based on the aforementioned Network Theoretical conceptual framework (Tsiotas and Polyzos 2013b) and on the non-directed interregional commuting network model. The proposed model utilizes Pearson's bivariate correlation (Norusis 2004; Devore and Berk 2012) and linear regression analysis (Norusis 2004, 2005).

The algorithm of the proposed model's construction consists of three steps. At the first step, the available node variables of the interregional commuting network are being shorted into three distinct thematic groups or classes, according to their relevance to the structural, behavioral, or *operational* components of the commuting network. Indicatively, the degree, the distance-based, and related variables are grouped in the structural component's class; the variables related to causes producing flows are grouped in the behavioral component's class, and the variables referring to qualitative attributes of the society constituting the network are grouped in the ontological component's class. The three component classes constitute distinct and independent group sets (X_S, X_B and X_O), according to relation (1).

$$\begin{aligned} \mathbf{X} &\equiv \{X_i, i = 1, \dots, |\mathbf{X}|\} \\ \mathbf{X} &\equiv \mathbf{X}_S \cup \mathbf{X}_B \cup \mathbf{X}_O, \quad \mathbf{X}_i \cap \mathbf{X}_j = \emptyset, \quad j \neq i, j = \{S, B, O\} \end{aligned} \quad (1)$$

At the second step, the algorithm distinguishes the most representative node variables per component class. This is conducted with the use of Pearson's bivariate coefficients of correlation (Norusis 2004; Devore and Berk 2012; Tsiotas and Polyzos

Table 1 Examined variables in the commuting network analysis

| Variable | Description | Source |
|--|---|--|
| (Y) Number of commuters | Number of commuters that originate from a capital city | (GNSS—Greek National Statistics Service ESYE 2007) |
| (X ₁) Degree | Degree of each city for the Greek interregional road network of the year 2012 | (SE; Google Maps 2012) |
| (X ₂) Commuting degree | Degree of each city in the commuting network (number of commuting destinations) | (SE; GNSS—Greek National Statistics Service ESYE 2007) |
| (X ₃) Degree difference | $X_{S3} = X_{S1} - X_{S2}$ | SE |
| (X ₄) Closeness centrality | Accessibility index that lies in inverse analogy with the average distance of a city to the rest of the destination cities of the network | (Tsiotas et al. 2012) |
| (X ₅) Mobility centrality | Centrality index capable in capturing the ability of a node to produce flows for a given gravity cause (node weight) | (Tsiotas and Polyzos 2013a) |
| (X ₆) Population | Population of each prefecture for the year 2006 | (Polyzos 2011) |
| (X ₇) Labor population percentage | Population percentage of each prefecture for ages $20 \leq A \leq 65$ | (Polyzos 2011) |
| (X ₈) Gravity sign | Statistically tested sign of a city’s commuting role (expelling, neutral, attractive) | SE (based on X ₁) |
| (X ₉) Educational index | Composite index of each prefecture’s educational level | (Polyzos 2011) |
| (X ₁₀) GDP | Gross domestic product of each prefecture for the year 2006 | (Epilogi 2006) |
| (X ₁₁) Welfare index | Composite index showing each prefecture’s welfare lever for the year 2005 | (Epilogi 2006) |
| (X ₁₂) Public servants | number of public servants of each prefecture | (Epilogi 2006) |
| (X ₁₃) Direct commuters | $= \max \{ \text{incoming, outgoing commuters} \} \times S_7$ | SE |
| (X ₁₄) Minimum commuting distance | Distance of each city’s closest commuting destination | (SE; Google Maps 2013) |
| (X ₁₅) Average neighbor distance | Average distance of variable X ₂ destinations | (SE; Google Maps 2013) |
| (X ₁₆) Average bus route frequency | Each city’s average number of weekly bus routes for all available destinations | SE |
| (X ₁₇) Bus route destinations | Number of available bus trip destinations per capital city | SE |
| (X ₁₈) Bus flow index | $B_3 = B_2 \times S_{10}$ | SE |
| (X ₁₉) Average train route frequency | | SE |

Table 1 (continued)

| Variable | Description | Source |
|--|--|----------------------------|
| (X ₂₀) Train route destinations | Each city’s average number of weekly train routes for all available destinations Number of available train trip destinations per capital city | SE |
| (X ₂₁) Train flow index | $B_5 = B_4 \times S_{11}$ | SE |
| (X ₂₂) Transportations GDP | Each prefectures’s GDP in transportations | (Epilogi 2006) |
| (X ₂₃) Car number | Number of private cars at each prefecture | (Epilogi 2006) |
| (X ₂₄) Bus number | Number of buses at each prefecture | (Epilogi 2006) |
| (X ₂₅) Taxi number | Number of taxis at each prefecture | (Epilogi 2006) |
| (X ₂₆) Accidents | Number of accidents per prefecture | (Epilogi 2006) |
| (X ₂₇) Accidents percentage | Per capita number of accidents for each prefecture | (Epilogi 2006) |
| (X ₂₈) Productivity dynamism | Complex index depending from GDP change, unemployment percentage, labor productivity, and labor percentage | (Polyzos 2011) |
| (X ₂₉) Unemployment inequalities | Inequalities index for unemployment per prefecture for the year 2007 | (Polyzos and Tsiotas 2012) |

SE self edited

2014, 2015), as they are defined at relation (2), where $cov(x,y) \equiv s_{xy}$ stands for the covariance of variables x,y and $\sqrt{var(x)} \equiv s_x, \sqrt{var(y)} \equiv s_y$ are their respective standard deviations.

$$r(x,y) \equiv r_{xy} = \frac{cov(x,y)}{\sqrt{var(x) \cdot var(y)}} \equiv \frac{s_{xy}}{s_x \cdot s_y} \tag{2}$$

The representative node variables of each component class are chosen under the criterion to have the largest sum of correlation coefficients squares among the significant (chosen level of significance $\leq 10\%$) correlation pairs calculated per variable, as it is described in relation (3).

$$\begin{aligned} X_S \in \mathbf{X}_S : \forall X_i, X_j \in \mathbf{X}_S \mid \sum_i r^2(X_S, X_i) &= \max \left\{ \sum_i r^2(X_i, X_j) : P[r(X_i, X_j) = 0] \leq 0, 10 \right\} \\ X_B \in \mathbf{X}_B : \forall X_i, X_j \in \mathbf{X}_B \mid \sum_i r^2(X_B, X_i) &= \max \left\{ \sum_i r^2(X_i, X_j) : P[r(X_i, X_j) = 0] \leq 0, 10 \right\} \\ X_O \in \mathbf{X}_O : \forall X_i, X_j \in \mathbf{X}_O \mid \sum_i r^2(X_O, X_i) &= \max \left\{ \sum_i r^2(X_i, X_j) : P[r(X_i, X_j) = 0] \leq 0, 10 \right\} \end{aligned} \tag{3}$$

After electing the representatives of the component classes, the chosen variables are set as predictor variables (X_i) to a *Linear Regression Model* (Norusis 2004, 2005)

having response variable (Y) the number of commuters of the interregional commuting network (Table 1). The linear regression model produces an estimation of a linear equation that best describes the relation between the dependent variable and the set of independent, under the *Least Squares* optimization method and the constraint that the standard errors follow a normal distribution. The versions of the linear regression algorithm used here is the *Enter* method, where all inserted variables are calculated in the model at once.

According to the aforementioned algorithm, the proposed model of the interregional commuting in Greece is totally described in relation (4).

$$\begin{aligned}
 & \mathbf{X} \equiv \{X_i, i = 1, \dots, |\mathbf{X}|\} \\
 & \mathbf{X} \equiv \mathbf{X}_S \cup \mathbf{X}_B \cup \mathbf{X}_O, \quad \mathbf{X}_i \cap \mathbf{X}_j = \emptyset, \quad j \neq i, j = \{S, B, O\} \\
 & \text{Commuting_Network_Gy}(V, E) : \\
 & Y = f(X_{\text{structural}}, X_{\text{functional}}, X_{\text{ontological}}) = f(X_S, X_B, X_O) = b_S \cdot X_S + b_B \cdot X_B + b_O \cdot X_O + c \\
 & X_S \in \mathbf{X}_S : \forall X_i, X_j \in \mathbf{X}_S \left| \sum_i r^2(X_S, X_i) = \max \left\{ \sum_i r^2(X_i, X_j) : P[r(X_i, X_j) = 0] \leq 0, 10 \right\} \right. \\
 & X_B \in \mathbf{X}_B : \forall X_i, X_j \in \mathbf{X}_B \left| \sum_i r^2(X_B, X_i) = \max \left\{ \sum_i r^2(X_i, X_j) : P[r(X_i, X_j) = 0] \leq 0, 10 \right\} \right. \\
 & X_O \in \mathbf{X}_O : \forall X_i, X_j \in \mathbf{X}_O \left| \sum_i r^2(X_O, X_i) = \max \left\{ \sum_i r^2(X_i, X_j) : P[r(X_i, X_j) = 0] \leq 0, 10 \right\} \right.
 \end{aligned} \tag{4}$$

Limitations of the Proposed Model

The proposed model is subjected to a set of limitations evident from the quantification procedure of the Greek commuting phenomenon. The first limitation regards the fact that the commuting phenomenon is modeled as a node attribute in the interregional commuting network and not as an attribute describing the network’s edges. This limitation ignores the distribution of the commuting quantities within the network channels and allows just considering the incoming and outgoing sums of commuters that arrive or originate to each node. However, such limitation does not suggest a significant concern, since the total directed quantities of commuting flows are being captured in each node and thus the total amount of information is preserved. Besides, the node attribute of commuting is essential to the model for setting a common reference base for the correlation and linear regression analyses.

Another limitation is that the proposed model uses one representative variable per component class, a fact that causes the reduction of the model’s resolution, leading to an inevitable loss of information that probably depends on the level of colinearity occurring among variables of the same class. This limitation originates from a economizing rationale describing the construction of the model and it can accordingly be surpassed whether constructing a more complex model, where its representatives would be sub-models $f_S(X_S)$, $f_B(X_B)$, and $f_O(X_O)$ of the representative vector variables X_S, X_B, X_O , which they can be either linear or not and they can consist of independent or uncorrelated components within each component class.

Such a treatment would render an alternative expression to the interregional commuting model of relation (4) that considers the simple linear case $f_S(X_S) = b_S X_S, f_B(X_B) =$

$b_S X_B$, and $f_O(X_O)=b_O X_O$ and perhaps to an alternative optimization criterion of electing the representative variables. Nevertheless, the proposed model in this paper seems to operate almost perfectly on the Greek interregional commuting data and thus it sets no further requirement for constructing a more complex model.

The final limitation of the proposed model regards the grouping procedure of the available node variables describing the interregional commuting network. At the proposed algorithm, the classification of the model's variables to the structural, behavioral, and ontological classes depends on the researcher's volition, on how he interprets that a variable belongs to a group, although most of the cases are quite distinct. This may render the model to become sensitive to errors of grouping by eliminating the significance of potentially representative variables if they are placed to incorrect component classes. Such limitation can be surpassed through the try and error process, whether moving the uncorrelated variables from one component class to another and including them to groups where they are mostly correlated.

Results and Discussion

This paragraph presents and discusses the results of the analysis. At the first step of the algorithm of the proposed framework, the available variables were shorted into three classes, according to their relevance to the structural, behavioral, or operational components of the commuting network. Table 2 shows the results of this classification. For standardization purposes, the independent variable Y (number of commuters) was included at the behavioral component's class. Next, the representative values per class were calculated, according to the presented methodology. For evaluation purposes this time the sums of square coefficients of correlation were calculated twice, firstly including only for the variables inside each class (in-class case) and secondly including all the

Table 2 The classification of the available node variables

| Structural class | Behavioral (Functional) class | Ontological class |
|---------------------------------------|---|--|
| (S_1) Degree | (Y) Number of commuters | (O_1) Labor population percentage |
| (S_2) Commuting degree | (B_1) Direct commuters | (O_2) Educational index |
| (S_3) Degree difference | (B_2) Average bus route frequency | (O_3) GDP |
| (S_4) Closeness centrality | (B_3) Bus flow index | (O_4) Welfare index |
| (S_5) Mobility centrality | (B_4) Average train route frequency | (O_5) Public servants |
| (S_6) Population | (B_5) Train flow index | (O_6) Transportations GDP |
| (S_7) Gravity sign | (B_6) Car number | (O_7) Accidents |
| (S_8) Minimum commuting distance | (B_7) Bus number | (O_8) Accidents percentage |
| (S_9) Average neighbor distance | (B_8) Taxi number | (O_9) Productivity dynamism |
| (S_{10}) Bus route destinations | | (O_{10}) Unemployment inequalities |
| (S_{11}) Train route destinations | | |

Table 3 Selected node variables for the model construction

| In-class Rank | Structural class | | Total | | Behavioral class | | Total | | Ontological class | | Total | |
|---------------|------------------|------------------|----------|----------------|------------------|------------------|----------|----------------|-------------------|------------------|----------|----------------|
| | Variable | $\Sigma_S (r^2)$ | Rank | $\Sigma (r^2)$ | Variable | $\Sigma_B (r^2)$ | Rank | $\Sigma (r^2)$ | Variable | $\Sigma_O (r^2)$ | Rank | $\Sigma (r^2)$ |
| 1 | S6 | 2.746 | 2 | 12.116 | Y | 5.285 | 1 | 12.181 | O2 | 4.091 | 9 | 10.451 |
| 2 | S5 | 2.738 | 3 | 12.058 | B6 | 5.190 | 4 | 11.905 | O6 | 3.990 | 7 | 11.703 |
| 3 | S10 | 2.493 | 11 | 7.877 | B7 | 5.152 | 6 | 11.730 | O7 | 3.983 | 5 | 11.805 |
| 4 | S1 | 2.448 | 23 | 2.596 | B8 | 5.048 | 8 | 11.428 | O4 | 3.256 | 13 | 6.443 |
| 5 | S2 | 2.188 | 22 | 2.945 | B3 | 3.910 | 10 | 9.400 | O9 | 2.612 | 19 | 3.859 |
| 6 | S9 | 1.777 | 25 | 2.137 | B5 | 3.362 | 12 | 7.177 | O1 | 2.507 | 14 | 5.169 |
| 7 | S4 | 1.753 | 27 | 1.838 | B4 | 2.622 | 16 | 4.883 | O3 | 2.236 | 21 | 3.307 |
| 8 | S11 | 1.576 | 17 | 4.404 | B1 | 1.628 | 20 | 3.467 | O5 | 1.935 | 18 | 3.898 |
| 9 | S8 | 1.570 | 28 | 1.790 | B2 | 1.000 | 30 | 1.206 | O8 | 1.849 | 15 | 5.080 |
| 10 | S3 | 1.569 | 24 | 2.169 | | | | | O10 | 1.368 | 29 | 1.458 |
| 11 | S7 | 1.158 | 26 | 1.931 | | | | | | | | |

Higher values per class are shown in bold

available 30 node variables (total case), but without changing the classification of Table 2. The results of this procedure are shown in Table 3.

According to Table 3, the representative variables resulted from the *in-class* analysis are variable S_6 (population) for the structural component's class, variable B_6 (car number) for the behavioral component's class, and variable O_2 (educational index) for the *ontological* component's class. The respective results for the total case slightly differ from the in-class case, resulting to the variables S_6 , B_6 , and O_7 (number of accidents), in correspondence. The ranking of variable O_2 in the total case moves to place 3, following the variables O_7 and O_6 . For this reason, we apply three corresponding linear regression models.

An interesting also observation from Table 3 is that the dependent variable Y (number of commuters) is placed at the first place in the ranking, showing the highest value of correlation sums. This seems reasonable, since the available independent node variables are theoretically related to commuting, constituting (in different scale) determining factors of this phenomenon. Another interesting observation is that the population variable (S_6) is the highest correlated variable among the predictor variables. This renders to the commuting phenomenon, a gravity characteristic that is verified at the part of the linear regression analysis.

Continuing, Table 4 shows the results of the linear regression analysis applied for three different predictor sets (S_6, B_6, O_2), (S_6, B_6, O_7), and (S_6, B_6, O_6). As it can be observed from Table 4, all these three models have almost

Table 4 Linear regression results

| Determination | Model predictors ^a | Unstandardized coefficients B | Std. error | Standardized coefficients Beta | t | Sig. |
|--|-------------------------------|-------------------------------|------------|--------------------------------|--------|--------------------|
| (1). Predictors: (constant), S_6 , B_6 , O_2 | | | | | | |
| $R=0.999$ | (Constant) | -881.91 | 174.10 | | -5.066 | 0.000 |
| R square=0.998 | S_6 | 0.011 | 0.002 | 0.600 | 4.840 | 0.000 |
| Std. error of the estimate=600.75 | B_6 | 0.011 | 0.004 | 0.345 | 3.003 | 0.005 |
| | O_2 | 41.422 | 11.729 | 0.065 | 3.532 | 0.001 |
| (2). Predictors: (constant), S_6 , B_6 , O_7 | | | | | | |
| $R=0.998$ | (Constant) | -674.49 | 212.47 | | -3.174 | 0.003 |
| R square=0.997 | S_6 | 0.015 | 0.002 | 0.816 | 6.151 | 0.000 |
| Std. error of the estimate=693.20 | B_6 | 0.02 | 0.005 | 0.049 | 0.335 | 0.740 ^b |
| | O_7 | 1.198 | 1.482 | 0.134 | 0.809 | 0.424 ^b |
| (3). Predictors: (constant), S_6 , B_6 , O_6 | | | | | | |
| $R=0.999$ | (Constant) | -532.81 | 199.46 | | -2.671 | 0.011 |
| R square=0.997 | B_6 | 0.016 | 0.002 | 0.834 | 7.991 | 0.000 |
| Std. error of the estimate=643.41 | B_6 | 0.021 | 0.007 | 0.627 | 2.796 | 0.008 |
| | O_6 | -6.337 | 2.508 | -0.462 | -2.527 | 0.016 |

^a Enter method

^b Coefficients with significances greater than $p=0.1$ are considered of zero contribution to the model and thus they are rejected

absolute (~ 1) coefficients of determination, implying their almost perfect ability to describe the variation of the data (Norusis 2004). However, the determination ability of model 2 (S_6, B_6, O_7) is rejected, because the coefficients beta of the predictors B_6, O_7 are insignificant to the model.

Nevertheless, the high r^2 results of all these models in Table 4 highlight the utility of the proposed methodology in capturing in an almost absolute level of determination ($\geq 99.8\%$) the variation of the commuting phenomenon, by considering only in the models the $\sim 10\%$ of the available information (3 out of 29 variables). This renders a contractive attribute to the proposed methodology, operating as an alternative principal component analysis based on the components that the network theoretical consideration (Tsiotas and Polyzos 2013b) elected.

Finally, the results of Table 4 verify that the population (variable S_6) is the most significant component in the interregional commuting model, suggesting obviously the most important driving factor for the commuting phenomenon in Greece. The level of such contribution is about 60, 81, and 83% for the models 1, 2, and 3, in correspondence, where it should be taken under consideration that population also affects many behavioral and ontological variables of commuting (Polyzos 2011; Tsiotas and Polyzos 2013b).

Conclusions

This paper proposed a Network Theoretical model for interregional commuting analysis, measuring the volume of commuters that operate in a commuting network, based on Web-Science's conceptual framework that has been generalized to operate for any kind of network. This framework decomposed the notion of "network" into a structural, a behavioral, and an ontological component, providing classes for shorting the available commuting network variables into such attribute classes.

The proposed methodology is ruled by the rationale that each conceptual component may be sufficiently represented by one characteristic variable per class, which is chosen under the criterion of having the highest sum of square values of the significant coefficients of correlation. The applicability of the proposed model is evaluated on real data of the Greek interregional commuting system, and the results of the analysis show respect to the categories of the proposed conceptual framework.

The foregoing analysis elected the utility of the proposed methodology in capturing almost absolutely the variation of the commuting phenomenon, by considering only a contractive set of the available information. The proposed methodology is capable in operating as an alternative principal component analysis based on the network theoretical consideration.

Finally, the analysis indicated the variable of population as the most important driving factor for the commuting network in Greece, rendering to the commuting phenomenon a gravity characteristic. The impact of this importance is detected both in the functional and ontological component variables composing the interregional commuting model, a fact which draws the assumption that the diachronic architecture of the transportation infrastructure in Greece is developed to serve population needs.

References

- Apicella, C., Marlowe, F., Fowler, J., & Christakis, N. (2012). Social networks and cooperation in hunter-gatherers. *Nature*, *481*, 497–502.
- Barthelemy, M. (2011). Spatial networks. *Physics Reports*, *499*, 1–101.
- Bastian, M., Heymann, S., Jacomy, M. (2009). “Gephi: an open source software for exploring and manipulating networks”, *Proceedings of the Third International ICWSM Conference*, pp.361-362.
- Berners-Lee, T., Hall, W., Hendler, J., O’ Hara, K., Shadbolt, N., & Weitzner, D. (2007). *A framework for Web science*. Netherlands: Publishers Inc.
- Borgatti, S., Halgin, D. (2011). “On network theory”. *Organization Science, Articles in Advance*, pp.1-14, doi 10.1287/orsc.1110.0641.
- Brandes, U., & Erlebach, T. (2005). *Network analysis*. Berlin: Springer-Verlag Publications.
- Christakis, N., & Fowler, J. (2009). *Connected: the surprising power of our social networks and how they shape our lives* (1st ed.). New York: Little, Brown and Company Hachette Book Group Publications.
- Clark, W., Huang, Y., & Withers, S. (2003). Does commuting distance matter? Commuting tolerance and residential change. *Regional Science and Urban Economics*, *33*(2), 199–221.
- Devore, J., & Berk, K. (2012). *Modern mathematical statistics with applications*. London: Springer-Verlag Publications.
- Diestel, R. (2005). *Graph theory* (3rd ed.). Heidelberg: Springer-Verlag Publications.
- Easley, D., & Kleinberg, J. (2010). *Networks, crowds, and markets: reasoning about a highly connected world*. Oxford: Cambridge University Press.
- Epiligi. (2006). *The prefectures of Greece*. Athens: All Media Publications.
- Evans, G., Wener, R., & Phillips, D. (2002). The morning rush hour: predictability and commuter stress. *Environment and Behavior*, *34*(4), 521–530.
- Glaeser, E., & Kohlhase, J. (2003). Cities, regions and the decline of transport costs. *Papers in Regional Science*, *83*(1), 197–228.
- GNSS—Greek National Statistics Service (ESYE). (2007). *Statistical Yearbook*, Data from National Census 2001.
- Google Maps. (2012). <http://maps.google.gr/> (getting directions application) [accessed 15-12-2012].
- Google Maps. (2013). *Google Mapping Services*, Available at the URL: www.google.gr/maps?hl=el. Accessed 11 Oct 2013.
- Kalantzi, O., & Tsiotas, D. (2012). Social network sites: the case of FacebookTM in education. In D. Tseles, K. Malafantis, & A. Pamouktsoglou (Eds.), *Education and society: research and innovation in new technologies* (pp. 85–93). Athens: Sigxroni Ekdotiki (Modern Publishing) Publications.
- Koslowsky, M., Kluger, A., & Reich, M. (1995). *Commuting stress: causes, effects, and methods of coping*. London: Plenum Press Publications.
- Liu, Y., & Nie, Y. (2011). Morning commute problem considering route choice, user heterogeneity and alternative system optima. *Transportation Research Part B: Methodological*, *45*(4), 619–642.
- Murphy, E. (2009). Excess commuting and modal choice. *Transportation Research Part A: Policy and Practice*, *43*(8), 735–743.
- Norusis, M. (2004). *SPSS 13 Advanced statistical procedures companion*. New Jersey: Prentice Hall Publications.
- Norusis, M. (2005). *SPSS 14.0 statistical procedures companion*. New Jersey: Prentice Hall Publications.
- Ozbay, K., Bartin, B., Yanmaz-Tuzel, O., & Berechman, J. (2007). Alternative methods for estimating full marginal costs of highway transportation. *Transportation Research Part A*, *41*, 768–786.
- Polyzos, S. (2011). *Regional development*. Athens: Kritiki Publications.
- Polyzos, S., & Tsiotas, D. (2012). The evolution and spatial dynamics of coastal cities in Greece. In S. Polyzos (Ed.), *Urban development* (pp. 275–296). Croatia: Intech Publications.
- Polyzos, S., Tsiotas, D., & Minetos, D. (2013) “Determining the Driving Factors of Commuting: An Empirical Analysis from Greece”. *Journal of Engineering Science and Technology Review*, *6*(3), pp. 46–55.
- Tsiotas, D., & Polyzos, S. (2013a). Introducing a new centralities measure from the interregional road transportation network analysis in Greece. *Annals of Operations Research*. doi:10.1007/s10479-013-1434-0.
- Tsiotas, D., Polyzos, S. (2013b). “Interregional commuting under the Network Theoretical perspective: an empirical analysis from Greece”. *Presented at the Multidisciplinary Academic Conference on Transport, Logistics and, Information Technology*, Prague, Czech Republic, May 09–11, 2013, ISBN 978- 80-905442-0-8.
- Tsiotas, D., & Polyzos, S. (2014) “Analyzing the maritime transportation system in Greece: a complex network approach”. *Networks and Spatial Economics*. doi:10.1007/s11067-014-9278-y.

- Tsiotas, D., & Polyzos, S. (2015) “Decomposing multilayer transportation networks using complex network analysis: a case study for the Greek aviation network”. *Journal of Complex Networks*. doi:10.1093/comnet/cnv003.
- Tsiotas, D., Vafopoulos, M. (2010). “Mathematic significances as didactic transformations of the Web science”. *Proceedings of the Scientific Conference eRA-5, the SynErgy Forum* (15–17 Sep 2010, Piraeus, Greece), pp.237-250.
- Tsiotas, D., Polyzos, S., Alexiou, A. (2012). “Analyzing the interregional transportation network in Greece by using Graph Theory”, *Proceedings of the 3rd Pan-Hellenic Conference in Urban and Regional Planning and Development* (Volos, Greece, 27–30 Sep 2012), pp. 1100–1106 (ISBN 978-960-9439-13-8) [in Greek].
- Van Ommeren, J., & Fosgerau, M. (2009). Workers’ marginal costs of commuting. *Journal of Urban Economics*, 65(1), 38–47.
- Van Ommeren, J., & Rietveld, P. (2005). The commuting time paradox. *Journal of Urban Economics*, 58(3), 437–454.
- Zuckerman, E. (1999). The categorical imperative: securities analysts and the illegitimacy discount. *American Journal of Sociology*. 104(5), 1398–1397.