Characterizing the Composition of Economic Activities in Central Locations: A Graph-Theoretic Approach to Urban Network Analysis

Srinath K. Ravulaparthy (corresponding author)
Graduate Student, Department of Geography and GeoTrans Lab
University of California Santa Barbara
Santa Barbara, CA 93016
Email: srinath@geog.ucsb.edu
Phone: (805) 893-3663
Fax: (805) 893-3146

Konstadinos G. Goulias
Professor, Department of Geography and GeoTrans Lab
University of California Santa Barbara
Santa Barbara, CA 93106
Email: goulias@geog.ucsb.edu
Phone: (805) 893-4190
Fax: (805) 893-3146

Paper submitted for presentation and publication to the 93rd Annual Meeting of the Transportation Research Board, January 12 – 16, 2014, Washington, D.C.

Paper Number: 14-0958
Submitted in response to call for papers from:
ADD30 – Transportation and Land Development Committee

Word Count Abstract: 250
Word Count Manuscript: 5,890 + 3 Tables (750) + 4 Figures (1000) = 7,640

Manuscript revised and submitted on August 1, 2013.

ABSTRACT

This paper draws on advances in spatial networks by representing a city as a weighted primal graph of a street network, which takes into account the context of location and its importance. We introduce the link-based multiple centrality indices (L-MCI) to represent location properties in terms of closeness, intermediacy, straightness and accessibility to all other locations. The proposed methodology is built on concepts of multiple centrality assessment (MCA) model. Results from L-MCI clearly identify the major city centers in Santa Barbara County based on the geometric configuration of the network. Moreover, these centrality indices also exhibit some unique properties that can be observed across other network structures. We also employ a clustering technique that accounts for spatial dependence in centrality values across multiple spatial scales, which aids in classifying the region into locations of high centrality and low centrality.

We further demonstrate the novelty of this approach in examining the relationship between structural properties of the street network and spatial organization of economic activities in Santa Barbara County, California. Results from this study confirm that link-based network centrality indices play a significant role in spatial distribution of economic activities. Professional services and retail trade form a major proportion of economic activities in locations with very high centrality values, e.g. downtown areas. Locations with high betweenness centrality values are especially attractive to retail trade activities, as they generate a greater potential for business opportunities. The results clearly reveal the presence of core-periphery type of a city model in Santa Barbara County.

Keywords: spatial networks, graph theory, multiple centrality indices, spatial clusters, firm locations

INTRODUCTION

Integration of space into urban economic models was a major development (*1*) in answering questions of location, agglomeration, traffic congestion and the formation of cities. Spatial measures in land-use models include *accessibility measures*, such as cumulative opportunity-type indices, gravity-type indices and utility type indices are typically used to estimate the qualities of location's accessibility that are attributable to surrounding land-use attractions, and *graph theory measures*, that are typically used to estimate the qualities of location's accessibility that are attributable to the geometric pattern of the urban infrastructure.

Graph theory based approach to spatial systems in urban planning and design was formally operationalized by Hillier and Hanson (*2*) on cities by developing the *space syntax* methodology, which follows a dual graph representation of street networks, where streets are turned into nodes and the intersections as edges (or links). The most basic tenet of space syntax is the index of integration, which represents how integrated or central a given link is in the network and is shown to have a significant correlation with traffic and distribution of non-residential land-use activities (*3-4*). As noted by Porta and Crucitti et al. (*5-6*) the main shortcoming of the space syntax approach is that it does not account for metric distances, but rather focuses on topological distance measures (i.e., the number of connections, rather than the length of connections), which significantly underestimate the network system performance measures as pointed by Batty and Ratty (*7-8*). In addition, the dual graph representation in space syntax is also fundamentally different from the traditional network representation of spatial systems.

To overcome these issues, Porta and Crucitti et al. (*5-6*) propose the primal graph representation of spatial systems for network analysis defined within a geographic framework based on metric distance. In this representation urban street patterns are turned into undirected, valued primal graphs where intersections are nodes and streets are edges. The key focus of this approach is a set of centrality measures to spatial systems which is a fundamental concept in social network analysis (*9-10*). This family of primal graph representation of urban street networks and their associated centrality measures form the multiple centrality assessment (MCA) model that defines centrality of a place based on its closeness (proximity) to other places and its connectivity, intermediacy, and directedness (on a straight line between two points), as well as its criticality/importance to other places. The final outcome of the MCA model is to assign a set of centrality values to each street segment with the results being several maps of a street network each of which shows one set of centrality values for links on the network (*11-13*).

In this paper we complement the MCA model in two new ways: (a) accommodate the context of location and its importance through weighted link attributes like roadway capacity, population and opportunities at a place; and (b) accounting for the relative importance of a link in the network across multiple spatial scales and centrality values. Our paper also contributes to the on-going research efforts (*14-17*) in representing the structure and properties of the street network in integrated models of land-use and transportation. We also develop four clusters (or classes) of links (or locations) based on multiple centrality indicators using the Latent Class Cluster Analysis (LCCA) and then examine the composition and characteristics of economic activities by industrial sectors associated with each cluster.

For this purpose, we report the findings for the Santa Barbara County area as a case study, with a total population of 399,347 and an employment of 215,440 as reported for year 2000. In the next section we describe the data used followed by the methodology. After this, the spatial organization along with composition and characteristics of economic activities is presented for each class of links. The final section is a summary and conclusions.

DATA USED

This analysis is conducted for Santa Barbara County and we use the U.S. Census Topologically Integrated Geographic Encoding and Referencing (TIGER or TIGER/Line) network for year 2000. To compute the centrality indices we represent each link in the network as a node (or mid-point of the link) and thereby increase the number of edges by two fold. Thus, the U.S. Census 2000 Tiger/Line network has 32,395 nodes with 64,790 edges with an average edge length of 303.41 meters. We compute the network centrality indices at different spatial scales (e.g. local and regional), because the centrality measures at regional scale may not reveal network properties on a more local environment such as neighborhood. Also, as shown by Porta et al. (*5*) small-scale measures are useful to overcome the edge-effects as they distort the centrality values near the edge of a network. To study the spatial distribution of economic activities in Santa Barbara County, we use the year 2000 geo-referenced National Establishment Time-Series (NETS) database of 20,628 business establishments in the region. The NETS database is a longitudinal dataset with unit of observation being a business establishment that produces goods or services at a single physical location – for example a single store or an establishment (*18*) all classified by North American Industrial Classification System (NAICS). We use sixteen types of economic activities: (*a*) agriculture, forestry, fishing and hunting; (*b*) mining; (*c*) utilities; (*d*) construction; (*e*) manufacturing; (*f*) wholesale trade; (*g*) retail trade; (*h*) transportation and warehousing; (*i*) information; (*j*) professional, scientific, management, administrative and waste-management services; (*k*) health care; (*l*) arts, entertainment, recreation, accommodation and food services; (*m*) other services (except public administration); (*n*) finance, insurance, real estate and rental and leasing [FIRE]; (*o*) public administration and armed force; (*p*) educational services. Table 1 illustrates the descriptive statistics of employment size, sales and number of establishments for the different economic activities by their economic category.

TABLE 1 Descriptive statistics of economic activities by economic sector for year 2000

Economic Sector	Economic Activity Type	Total number of establishments	Number of employees		Sales reported in million USD	
			Mean	Std. Deviation	Mean	Std. Deviation
Primary Sector	Agriculture (11)	535	17.16	43.52	1.625	5.513
	Mining (21)	80	30.56	79.71	4.298	10.389
Secondary Sector	Utilities (22)	32	11.34	11.05	1.923	3.169
	Construction (23)	1,619	5.59	16.20	0.821	3.122
	Manufacturing (31 – 33)	1,210	19.05	84.54	2.546	14.258
Tertiary Sector	Wholesale Trade (42)	943	7.44	16.07	1.903	5.540
	Retail Trade (44-45)	3,126	7.26	19.53	1.054	6.298
	Transport & warehousing (48-49)	341	9.25	17.79	0.805	2.248
	Information (51)	544	8.85	22.66	1.382	9.817
	Professional Services (54 – 56)	4,029	5.73	19.11	0.534	1.848
	Healthcare Services (62)	1,881	14.46	87.40	0.894	5.800
	Arts & entertainment (71 – 72)	1,612	15.53	35.29	0.692	2.355
	Other Services (81)	2,249	5.99	43.73	0.459	3.517
	FIRE (52 – 53)	1,896	6.82	22.34	0.973	4.595
Quaternary Sector	Public Administration (92)	153	115.31	776.73	1.127	11.847
	Educational Services (61)	378	36.20	96.79	2.151	5.980

LINK-BASED MULTIPLE CENTRALITY INDICES (L-MCI)

We develop here the link-based multiple centrality indices (L-MCI) that unify the concepts of weighted spatial graph representation and multiple centrality assessment (MCA) model. This modified version of the MCA model (or L-MCI) is based on: (a) primal graph representation of street networks; (b) location instances of activities or opportunities that are distributed along the street network; (c) weighted representation of a primal graph based on the attributes of activities or opportunities; and (d) transferability of the link-based centrality values to these location instances. Furthermore, a subtle distinction of this method over the original MCA model is that we compute the network centrality indices directly for a link, rather than averaging the centrality values from its connected nodes.

The steps of our method are: ***Step1:*** snap (or attach) each location of each business establishment to the nearest street network link within 100 meters; ***Step2:*** construct the database representing the mapping of every activity or opportunity along the street network within 100 meters; ***Step3:*** enumerate the attributes of the business establishments per link using the unique identifiers of activities or opportunities and street network links; ***Step4:*** compute the weighted network centrality indices at the point level representing each street network link using metric distances and weights as attributes obtained in *Step3*. The point location in this case is the mid-point of each street network link in the region (in essence a new set of nodes are part of a primal graph with twice as many edges); ***Step5:*** from the database in *Step2* transfer the centrality indices of the links to their associated activities or opportunities that are distributed along the street network.

For this purpose, the urban street network is represented as an undirected weighted primal graph $\boldsymbol{G}$ with $\boldsymbol{N}$ nodes and $\boldsymbol{K}$ edges. The weighted primal graph representation is defined by nodal attributes (their weights) and edges defined by metric distance. The spatial graph is described by the adjacency $\boldsymbol{N}\times\boldsymbol{N}$ matrix $\mathbf{A}$, whose elements $\boldsymbol{a_{ij}}$ are equal to 1 when there is an edge between links i and j, and 0 otherwise. It should be noted that in our case, the number of nodes (which also represent links) $\boldsymbol{N}$ is half as the number of edges $\boldsymbol{K}$ in the spatial graph, this is because we are representing each link as a node. This representation should not be confused with the space syntax methodology that adopts a dual graph representation. Based on this notation following are the link-based multiple centrality indices (L-MCI) that are used here.

Degree Centrality **($\boldsymbol{C^D}$)** is the count of the number of edges incident upon a given node, i.e. the number of first neighbors. Thus, degree centrality of link i is defined as:

$$C_i^D = \frac{k_i}{N-1} = \frac{\sum_{j \in N} a_{ij} W[j]}{N-1}$$

Where N is the total number of links and k_i is the degree of the link i, i.e., the number of links adjacent to i, and a_{ij} is the element of the adjacency matrix $\mathbf{A}$ and $W[j]$ is the weight associated with link j – the weight is the attribute information derived from *Step3*, by enumerating the attributes of the activities or opportunities per link along the street network.

Closeness Centrality **($\boldsymbol{C^C}$)** for a link i on a network measures to what extent a link is close to all the other links along the shortest paths from one link to another on the network (*19*) defined by:

$$C_i^C = \frac{1}{\sum_{j \in N} d_{ij} W[j]}$$

Where N is the total number of links in the network, and d_{ij} is the network shortest path distance between links i and j and $W[j]$ is the weight associated with the link j. Closeness centrality is a measure of clustering or dispersion of activities or opportunities in a region. Closeness centrality also captures the accessibility of a place, which reflects the cost of overcoming spatial separation between places with population or activities. To have a more meaningful and numeric interpretation we use *distance remoteness centrality* (C^{DR}), the inverse of the closeness centrality measure, defined as:

$$C_i^{DR} = \sum_{j \in N} d_{ij} W[j]$$

Betweenness Centrality **(C^B)** for a link i on a network is based on the idea that a link is more central when it is traversed by a large number of shortest paths connecting any other two links in the network, and this is defined as:

$$C_i^B = \sum_{j,k \in N; j \neq k; k \neq i} \frac{n_{jk}[i]}{n_{jk}} W[j]$$

Where n_{jk} is the number of shortest paths that exist between links j and k, and $n_{jk}[i]$ is the subset of these paths that pass through link i, weighted by $W[j]$. C^B captures the prominence of a link acting as intermediary among many links.

Straightness Centrality **(C^S)** for a link i represents "efficiency of communication" between two links increases when there is a least deviation of their shortest path from the virtual straight line connecting them – that is, a greater straightness of the shortest-path distance. Thus straightness centrality is defined as:

$$C_i^S = \sum_{j \in N; j \neq i} \frac{\delta_{ij}}{d_{ij}} W[j]$$

Where δ_{ij} is the straight line Euclidean distance between links i and j. The rest of the symbols are as defined in the other indicators. This measure of centrality is a corridor property in a region. For example, in downtown regions, where the street networks are in grid format tend have a higher straightness centrality along with high density of economic activities along these street networks. This result is later illustrated in detail with computed centrality measures.

Reach Centrality **(C^R)** for a link i measures the number of other links that can be reached along the shortest path on a network. It is defined as follows:

$$C_i^R = \sum_{j \in N} W[j]$$

Where $W[j]$ is the weight of the destination link j, and the reach metric is equivalent to the cumulative opportunities type accessibility measure as discussed in Bhat et al. (*20*), but applied on a network rather than Euclidean space.

Thus, the link-based multiple centrality indices capture a location's advantage of various places in a city and the importance of individual places that contribute to the spatial interaction between activities or opportunities in a city. Furthermore, these indices also account for the connectivity and configuration of the urban street network. To compute these centrality indices we use Urban Network Analysis toolbox for ArcGIS developed by Svetsuk and Mekonnen (*21*).

RESULTS FROM LINK-BASED MULTIPLE CENTRALITY INDICES

To determine the large-scale and small-scale measures for the centrality indices, we compute centrality indices for different network radii or network buffers surrounding each link of – 2.5km, 6km and 12km along with measures for the entire county, which are the 25th, 50th and 75th percentiles of the pairwise distance distribution. Furthermore, to make the analysis tractable we only report the results based on four centrality indices-C^{DR}, C^{B}, C^{S}, C^{R} computed for the different network buffers as indicated. To further study the correlation between network structure and its properties with location patterns of economic activities we only report the results from un-weighted link-based multiple centrality indices with $W[j] = 1$.

We further simplify the analysis, by normalizing the network centrality measures given as: $C_{di}^{N} = C_{di}^{M}/\max[C_{di}^{M}]$ where $0 \leq C_{di}^{N} \leq 1$ is the normalized centrality measure for link i within a given network radius d being – 2.5km, 6km, 12km and all county buffers for N being the different types of centrality measures – closeness (or remoteness), betweenness, straightness and reach. Similarly, C_{di}^{M} is the non-normalized centrality measure for link i within a given network radius d as computed from the aforementioned equations for M being different types of non-normalized centrality measures. Table 2 provides a summary of these measures.

TABLE 2 Mean and (standard deviation) of normalized network street centrality measures

Centrality Index	Network Buffer			
	All County	12km buffer	6km buffer	2.5km buffer
Remoteness (C^{DR})	0.455 (0.126)	0.408 (0.270)	0.313 (0.276)	0.220 (0.230)
Betweenness (C^{B})	0.017 (0.072)	0.025 (0.068)	0.035 (0.082)	0.033 (0.070)
Straightness (C^{S})	0.884 (0.098)	0.391 (0.286)	0.303 (0.277)	0.213 (0.226)
Reach (C^{R})	0.999 (0.029)	0.430 (0.298)	0.325 (0.285)	0.228 (0.233)

In the interest of brevity, we only present the spatial distribution of network centrality measures within 2.5km network buffer as indicated in Figure 1. It should be noted in Figure 1 that lower values are represented in gray color and higher values in dark red. Recalling from the definition of distance remoteness (C^{DR}), higher values of this measure for all county indicate the farthest links in a region. However, within a defined distance threshold, the remoteness centrality index increases indicating the presence of high network connectivity and concentration. For example, as from Figure 1(a) in downtown areas of Santa Barbara and Santa Maria this is evident, where a center emerges within a network buffer of 2.5km. This is attributed to the fact that with presence of high network links, the cumulative distance also increases within this buffer. Therefore, one of the important properties of remoteness centrality index is that when there is a high concentration of network links, without delimiting the study area the index value is lower at these locations.

While, limiting the study area (e.g. using a network buffer of 2.5km), these locations with high network concentration have an increased value in remoteness centrality.

Betweenness centrality (C^B) at the link level captures the role of a link as a pass-through or a traversing point from an origin to a destination in the region. Therefore, the links with high betweenness values have a special property in the fabric of a region. In this analysis, the un-weighted betweenness centrality can be interpreted as a potential (or gateway) to other network links that a link provides to a passerby from an origin to a destination in the network. Thus, betweenness centrality for the entire county with no network buffer is very high only along the major freeway CA-154 that connects the north and southern regions of the county with certain sections of HWY-246 and US-101. This result is as expected, because for the entire county the links along CA-154 are central and serve only as facilitator or as a hub for any travel or interaction that needs to occur among other links in the network. This result is also a manifestation of the property of the network structure like a star network (or hub-spoke network), in which the central node has the highest betweenness value (*22*).

However, the betweenness centrality values vary significantly at different network buffers in representing the critical links for a passerby between any two locations in the network. For example, betweenness centrality index for 2.5km network buffer is very high along the collector and local roads that serve as connectors to major arterials and freeways. This result is illustrated in Figure 1(b) only for the downtown Santa Barbara and Goleta areas. As noticed, this result highlights the significance of local and collector roads in the downtown area relative to the regional network. Moreover, betweenness centrality index for the entire study region has more traversing links along state highways and freeways like CA-154 and US-101, which are not traversed at the 2.5km network buffer. Furthermore, this result also supports the argument by Porta and Wang et al. (*11, 13*) that these links may not attract people or freight as a major trip destination, but they are pass-through nexus to generate business opportunities.

As seen from Figure 1(c), straightness centrality (C^S) represents the extent to which the shortest-paths from a link of interest to all the other links resemble the straight line Euclidean distance. Higher values of straightness centrality indicate larger deviations of network shortest-path distance from straight line distance. This suggests that the longer the commute, the less likely the path to resemble a straight line. As expected, higher values of straightness index are in the areas with high concentration of roadway network links, especially in the downtown areas of Santa Barbara, Santa Maria and Lompoc. Moreover, straightness centrality within 2.5km network buffer clearly distinguishes the urban centers with higher values, thereby capturing the effects of local network structure. Also, the network links in these urban centers (or downtown areas) also happen to have a design of straightest thoroughfares like a grid network, which represent the efficiency of links in the downtown areas.

Finally, reach centrality (C^R) is the same as cumulative opportunity type accessibility measure, which captures the total number of opportunities that can be reached from a link to all other links in the network. In our analysis, for the un-weighted case the reach centrality is the total number of network links that can be reached within a given network buffer. For example, within the 2.5km network buffer as shown in Figure 1(d) high values of reach centrality are concentrated in downtown areas of Santa Barbara, Santa Maria and Lompoc regions. These are also the areas with high concentration of network links. Furthermore, reach centrality can also be interpreted as a special case of straightness centrality index, where the network shortest-path and the straight line Euclidean distance are identical or $\frac{\delta_{ij}}{d_{ij}} = 1$.

Based on the spatial distribution of the network centrality measures; for example, areas with high remoteness values within the 2.5km network buffer also happen to be locations with high intensity of economic activity and opportunities. Furthermore, these locations with high intensity of economic activities also in-turn are along the corridors in the network with high straightness and reach centrality indices for all network buffers. Similarly, corridors with high betweenness centrality index, mainly being in the downtown areas also contain a relatively high density of retail activity, suggesting that betweenness may be an important factor in spatial distribution of retail activities. Therefore, to further understand the linkages between network centrality and distribution of economic activities in a region, we further explore these spatial patterns in a comprehensive manner as described in later sections.

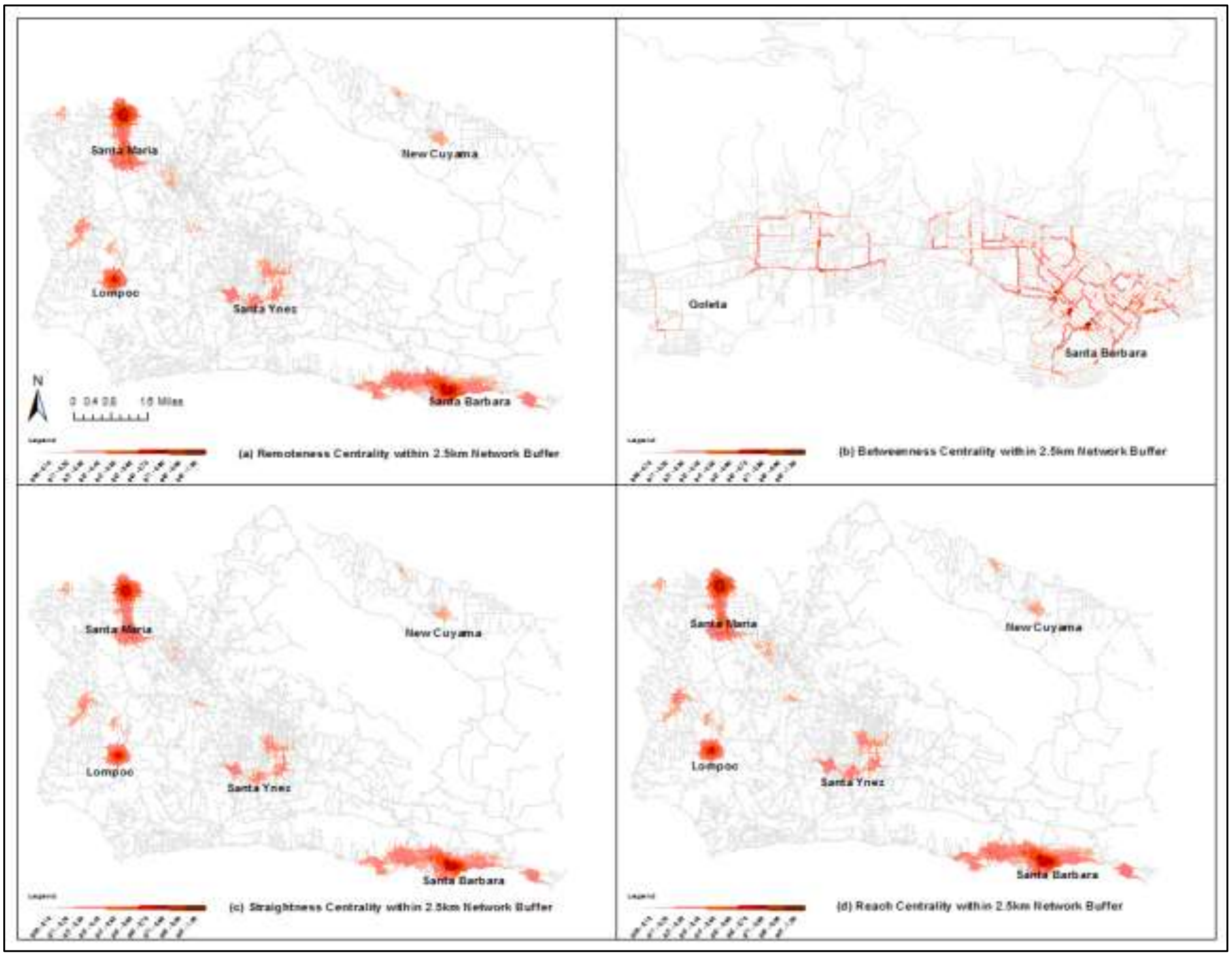

FIGURE 1 Spatial distribution of normalized network centrality measures within a network buffer of 2.5km – (a) remoteness centrality (b) betweenness centrality (c) straightness centrality (d) reach centrality

LATENT CLASS CLUSTER ANALYSIS

Our next objective in the analysis is jointly classifying the region into places (or corridors) that are distinctly different in terms of their network centrality values. It is also desirable to develop a small number of groups that depict the spatial patterns of the centrality measures in the study region. We use Latent Class Cluster Analysis (LCCA) that uses patterns of variation among many dependent variables (in this case the normalized network centrality measures) and identifies groups of links with relatively homogenous scores (*23*). Furthermore, the LCCA technique also helps in addressing the issue of case-relativity as raised by Porta et al. (*12*) due to different network distance thresholds.

For this analysis, sixteen continuous normalized centrality measures(C_{di}^{N}) for network distance thresholds of – 2.5km, 6km, 12km and all county are analyzed jointly using the Latent GOLD® 4.5 software. Starting with a single cluster LC baseline model through an iterative process of testing multiple numbers of latent class clusters we reached the four cluster latent class model as the ideal balance between parsimony and fit with the lowest Bayesian Information Criterion (BIC) value of *-183772.26* and least classification error of *0.0127*. Table 3 illustrates the mean of normalized network centrality scores along with a share of network links that belong to each of the four clusters.

From the results, network links that belong to Cluster#1 (24.9% of roadway links) have significantly high centrality scores when compared with links that belong to Cluster#4 (27.4% of roadway links), which contains the lowest centrality scores. However, centrality scores for the links in Cluster#2 (32.8% of roadway links) and Cluster#3 (14.9% of roadway links) show a distinct pattern for different network distance thresholds. For example, centrality scores within the 12km network buffer for roadway links in Cluster#2 are consistently higher when compared with roadway links in Cluster#3. In contrast, the same roadway links in Cluster#2 within the 2.5km buffer have relatively lower centrality scores when compared with roadway links in Cluster#3. Based on this result, the roadway links within Cluster#2 and Cluster#3 are quite heterogeneous in their centrality measures when compared across different network buffers.

Therefore, Cluster#1 represents roadway links that are highly central. Cluster#2, contains roadway links that are uniformly central. Cluster#3, contains roadway links that are moderately central and the roadway links in Cluster#4 are the least central. It should also be noted that Cluster#3 (moderately central) consistently has higher betweenness centrality values when compared with links in Cluster#2. This result indicates that network links in both Cluster#1 and Cluster#3 are the most critical links in the network structure that serve as pass through points for different origins and destinations in the region.

The spatial distribution of these clusters mapped in Figure 2 show clearly defined regions of highly central places from Cluster#1 (color coded in red) and they are located in northern and southern regions of the county. The uniformly central places (Cluster#2) in blue are surrounding the highly central places in these parts of the county. Furthermore, the uniformly central places in cities of Lompoc and Solvang are also clearly identifiable. Moderately central places from Cluster#3 (color coded in green) are sparsely spread across the county. For example, roadway links in these central places like CA-154 serve only as major connectors between the north and southern regions of the county. However, all the roadway links in New Cuyama are moderately central within the region. This indicates that these network links are predominantly localized and do not serve as critical links in the overall structure of the network in the county. The least central places as in Cluster#4 (color coded in grey) are widely spread across the county indicating their least prominent role or functionality in the network structure of the region. Most important, the locations of highly central links are in the downtown areas of Santa Barbara and Santa Maria. This signifies the criticality of these links which includes major freeways and local roads in the fabric of the regional network structure and its functionality.

TABLE 3 Profile of the four latent class clusters

Network Buffer (d)	Cluster Size (n = 32,395)	**Cluster-1 (Highly Central)**	**Cluster-2 (Uniformly Central)**	**Cluster-3 (Moderately Central)**	**Cluster-4 (Least Central)**
		24.9%	32.8%	14.9%	27.4%
	Mean Normalized Network Centrality Scores				
All County	Reach	1.000	1.000	0.996	1.000
	Betweenness	0.016	0.001	0.068	0.012
	Straightness	0.951	0.909	0.867	0.805
	Remoteness	0.445	0.421	0.458	0.504
Within 12km	Reach	0.812	0.464	0.337	0.099
	Betweenness	0.071	0.003	0.041	0.001
	Straightness	0.771	0.406	0.302	0.079
	Remoteness	0.708	0.466	0.328	0.116
Within 6km	Reach	0.730	0.303	0.229	0.037
	Betweenness	0.110	0.005	0.043	0.001
	Straightness	0.702	0.268	0.215	0.032
	Remoteness	0.697	0.304	0.208	0.038
Within 2.5km	Reach	0.522	0.189	0.204	0.023
	Betweenness	0.096	0.008	0.049	0.001
	Straightness	0.495	0.168	0.197	0.021
	Remoteness	0.516	0.179	0.186	0.020

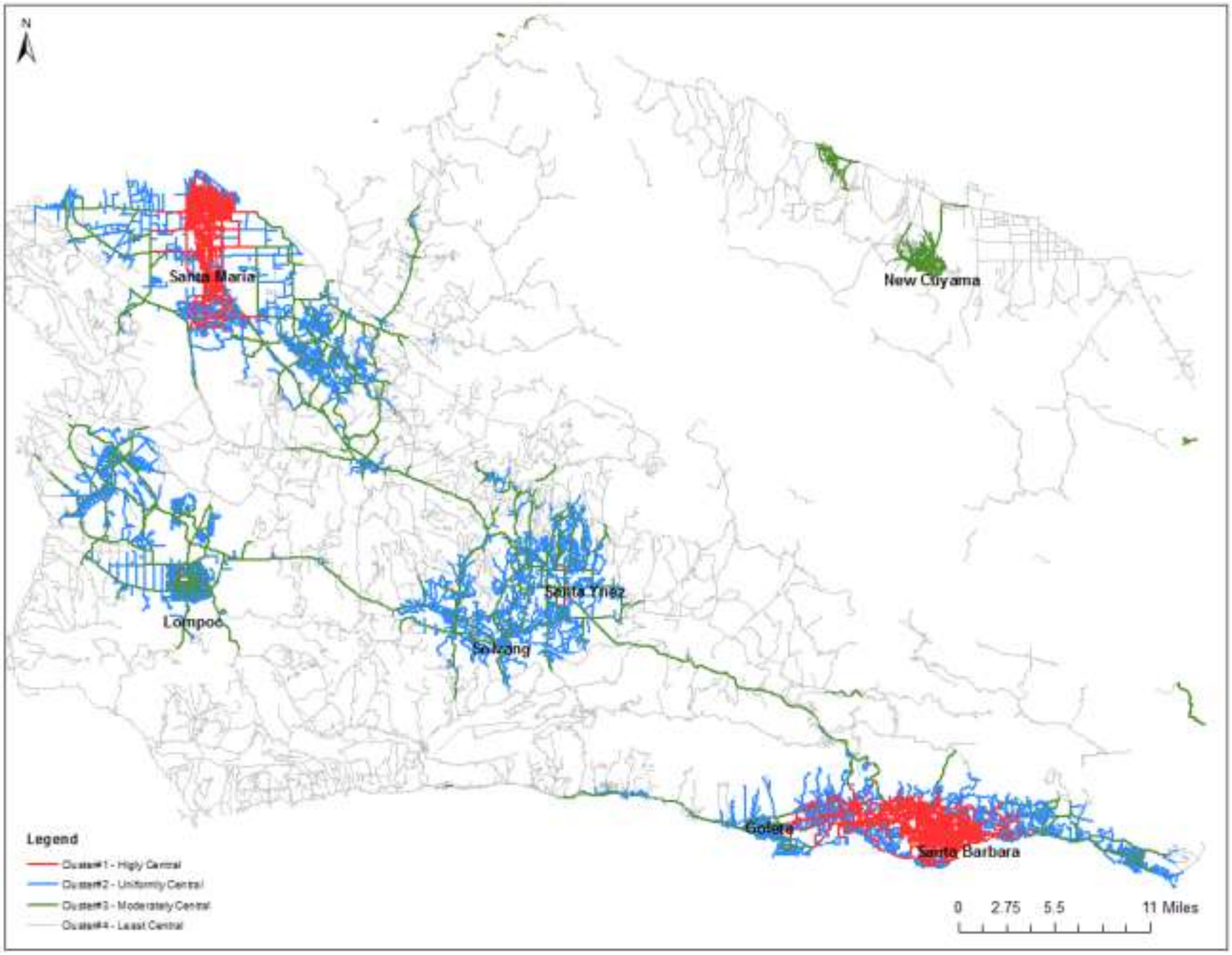

FIGURE 2 Spatial distribution of highly central and least central places in Santa Barbara County (we display the clusters with red indicating highly central, blue indicating uniformly central, green indicating moderately central and grey indicating least central)

LINKING NETWORK CENTRALITIES WITH ECONOMIC ACTIVITIES

Recent studies (*11*-*13*) indicate a significant correlation between spatial distribution of economic activities and street centrality indices. However, there is very little empirical evidence correlating network centrality indices with different categories of economic activities and their characteristics. Most of these studies have used the multiple centrality assessment (MCA) model and kernel density estimation methods in continuous space to establish this correlation. However, the composition of these economic activities along with their characteristics relative to the central places based on the network structure in the region is also important to consider in firm location choice models. Using the four network centrality clusters as derived from the LCCA model we study the composition and characteristics of these economic activities that are located across the four clusters.

Figure 3 illustrates the distribution of economic activities in Santa Barbara County across the four clusters. The tertiary sector of the economy dominates Santa Barbara County, with major economic activities like professional services, retail trade and arts and entertainment (including food and accommodation services) located primarily in Cluster#1. Furthermore, 78.2% of healthcare services in the county are located in Cluster#1 (highly central). The primary sector of the economy like agriculture and mining are uniformly distributed across the four clusters.

Moreover, the results also support the presence of a core-periphery or a center-sub-center type of a city formation in Santa Barbara County. For example, in southern region, downtown Santa Barbara is a major center (Custer#1 - highly central) with city of Goleta and portions of Santa Barbara acting as sub-centers (Cluster#2 - uniformly central). A similar spatial pattern is also observed in the northern region in City of Santa Maria supporting the argument of emerging polycentric cities within a region (*24-25*).

We further report the composition of these economic activities within each of the four clusters as depicted in Figure 4. Based on the composition, it is evident that there are significantly different locational patterns for different economic activities across each of the network centrality clusters. As seen from the distribution of economic activities in Cluster#1 (highly central), professional services (20.3%), retail trade (15.5%) and healthcare services (11.39%) dominate a major share of the composition of economic activities in highly central places. We also find here other service based economic activities like finance, insurance and real estate [FIRE], arts and entertainment and other services. Therefore, based on this composition it can be inferred that in Santa Barbara County, retail and service based economic activities tend to concentrate in the areas that have very high centrality values. This could be attributed to the property of the network structure. For example, the roadway links in Cluster#1 have the highest straightness centrality value indicating smaller deviations of the network shortest-path distance from the straight line distance. The smaller deviation could be especially appealing to retail sector, as these areas can be reached directly in-terms of potential business opportunities.

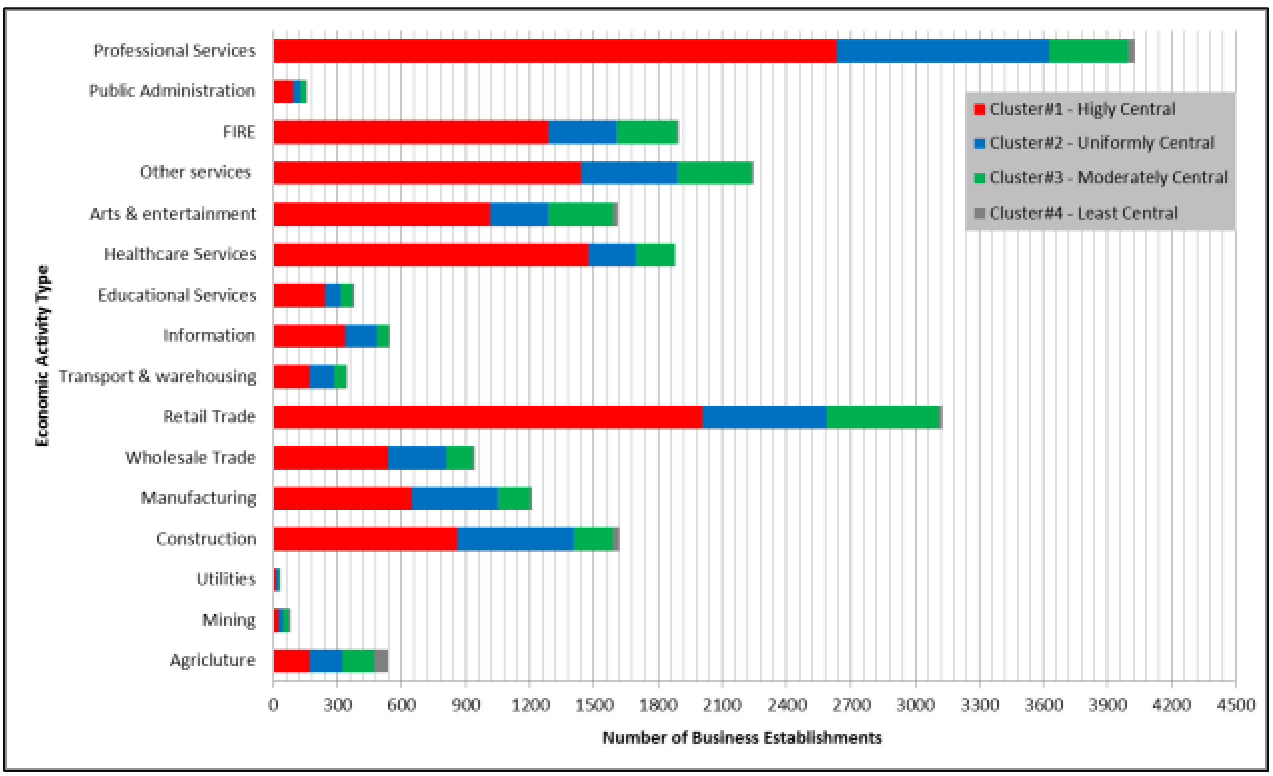

FIGURE 3 Distribution of economic activities across the four clusters

Although, Cluster#1 is dominated by service based economic activities, there are also establishments in agriculture, construction and manufacturing. Overall, the results and composition of economic activities in Cluster#1 that central locations in an urban area have the potential to sustain higher densities of retail and service activities (*26*).

In Cluster#2 the major proportion of economic activities are also professional services (21.5%) and retail trade (12.5%). Although, there is relatively high concentration of secondary economic activities like construction (11.75%) and manufacturing (8.7%) when compared with Cluster#1. This difference in distribution of economic activities related to tertiary and secondary sector of the economy can be attributed to the fact that central places command higher real estate values with limited land availability. Secondary sectors of the economy may need higher land area with higher employment for their occupancy when compared with tertiary sectors of the economy. It should also be noted that in Cluster#2 the network centrality values are also consistently high at local and neighborhood levels, which might offer a locational advantage for the service businesses located in these areas. This is especially important as there is a significant positive correlation for local street network design and distribution of non-residential economic and service activities (*2,3,27*). Also, Cluster#2 supports other types of economic activities related to tertiary and quaternary sectors of the economy namely, healthcare services, public administration and educational services. Moreover, Cluster#2 in city of Goleta also consists of University of California, Santa Barbara, which is the largest educational institution in this region.

The core-periphery organization in South Santa Barbara County is evident with downtown Santa Barbara serving as the major center or CBD, while, Goleta serving as a minor center. This result supports the argument that cities are increasingly polycentric and firms in the region value access to the major center (CBD) and the nearest minor or sub-centers (*29-29*). Also, Lang (*30*) argues that when cities become edgeless and polycentric, their composition of activities are heavily oriented towards service based economic activities. However, there are also other spatial economic factors like agglomeration and dispersion that have an impact on the urban spatial structure (*31-32*).

In contrast, the composition of economic activities in Cluster#3 (moderately central) is dominated by retail trade (18.6%), professional services (12.9%) and other services (12.3%) that include automotive repair, beauty salons and professional organizations. It should be also noted that Cluster#3 consistently scores higher betweenness centrality values across all geographic scales when compared with Cluster#2. This composition of economic activities can be attributed to high betweenness centrality values, in which a place itself may not serve as a final trip destination, but it may take advantage of its unique location in the system as merely a pass-through nexus to generate great economic opportunities. Therefore, this result also supports the finding by Porta et al. (*11*) that high value of betweenness centrality often implies a higher concentration of commercial or service activities as observed in Cluster#3. Furthermore, Cluster#3 is also home to the Vandenberg Air Force Base in city of Lompoc, which serves as a major attractor for other manufacturing and service related economic activities as seen from Figure 5, in which betweenness centrality plays a significant role in connecting these locations with the rest of the county. Finally, the composition of economic activities in Cluster#4 (least central locations) are heterogeneous and include a major proportion of agricultural activities (26.2%) in the region with construction and professional services also reported to be high (13.5%). This type of economic composition can be as expected in the least central locations, because these are rural areas dominated by farm lands, vineyards and ranches.

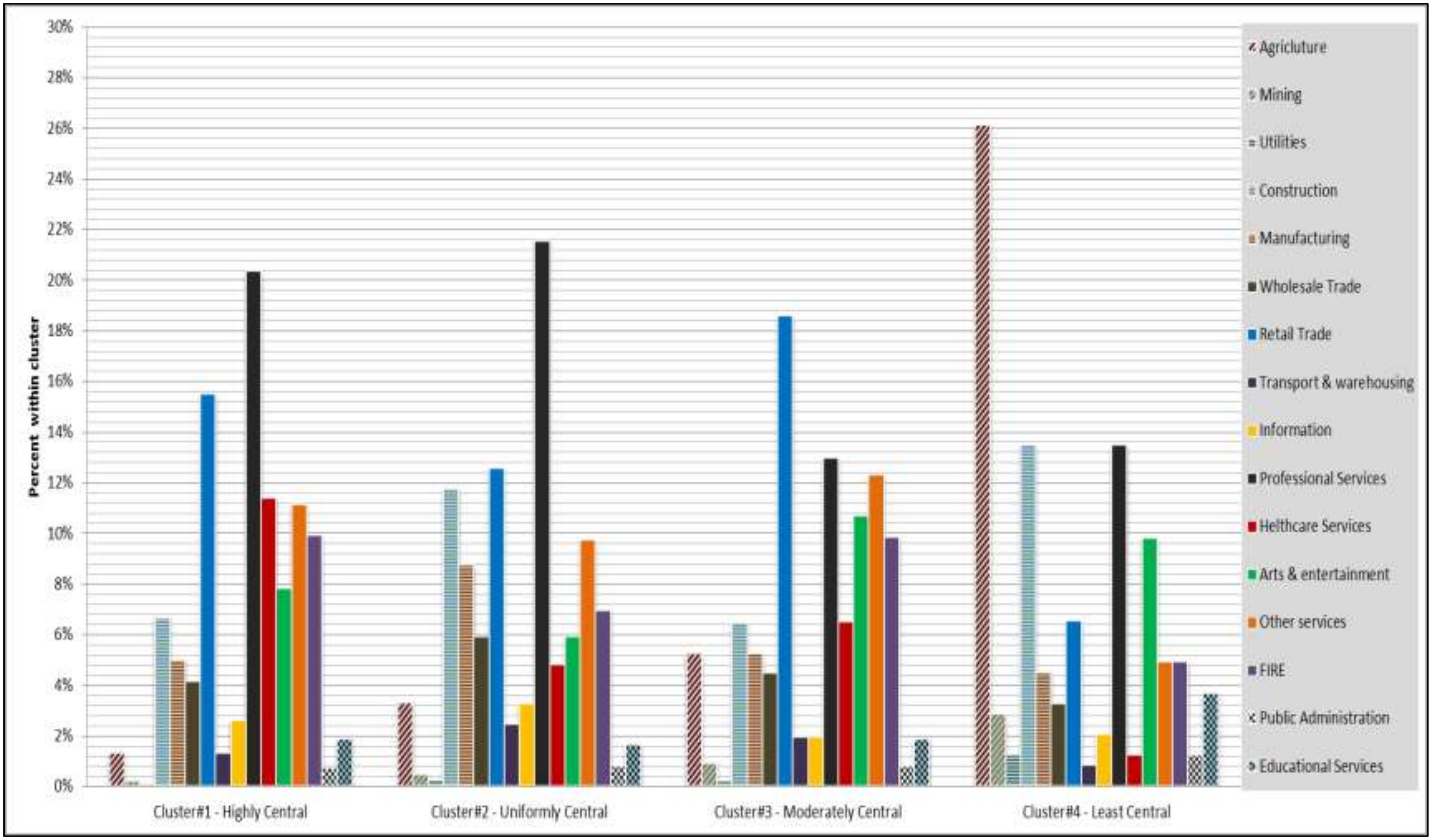

FIGURE 4 Composition of economic activities by network centrality clusters (we represent primary sector economic activities in an inclined 45^0 line pattern; secondary sector economic activities are represented in an horizontal line pattern; tertiary sector economic activities are represented with no patterns and quaternary sector of economic activities are represented in a point pattern)

SUMMARY AND CONCLUSIONS

In this paper we capture the significance of a location in a city in terms of its closeness, intermediacy, directness and accessibility to all other locations. We achieve this using link-based multiple centrality indices (L-MCI). The link-based network centrality is measured across four dimensions – closeness, distance remoteness, betweenness, straightness and reach. These centrality indices exhibit unique geometric properties delineating network regions and critical locations to guide urban planning and design. In addition, using Latent Class Cluster Analysis (LCCA) we classify the region into highly central and least central places and further study the spatial distribution and composition of economic activities within these clusters.

The results show that locations with high centrality values (Cluster#1) tend to sustain high intensity of economic activities with a major proportion being professional services and retail trade activities. Furthermore, Cluster#3 (moderately central) has a unique property in the region, with locations exhibiting high betweenness centrality values with significantly high intensities of retail trade activities. These locations function as pass-through connections to generate great business opportunities. Also, there exists a core-periphery type of a city model in Santa Barbara County, in which spatial configuration of the network and its centrality measures play a significant role. For example, downtown Santa Barbara serves as the core (or major center being Cluster#1) with city of Goleta serving as a periphery (or sub-center being Cluster#2). These results reinforce the notion of strong interdependence of urban street network and spatial patterns of locational activities in a region. Further research includes a more detailed statistical analysis of business composition in each cluster, the study of location and relocation patterns, and testing of the methods used here in other regions.

ACKNOWLEDGEMENTS

Funding for this research was provided by The University of California Transportation Center. This paper does not constitute a policy or regulation of any local, state or federal agency.

REFERENCES

1. Krugman, P. (1998). What's new about the new economic geography?. Oxford Review of Economic Policy, vol.14, no.2, pp.7-16.
2. Hillier, B. and Hanson, J. (1984). The Social Logic of Space. Cambridge University Press, Cambridge, U.K.
3. Hillier, B. (1996). Space is the Machine: A Configurational Theory of Architecture. Cambridge University Press, Cambridge, U.K.
4. Jiang, B. and Claramunt, C. (2004). Topological analysis of urban street networks. Environment and Planning B: Planning and Design, vol.31, pp.151-162.
5. Porta, S., Crucitti, P. and Latora, V. (2006). The network analysis of urban streets: a primal approach. Environment Planning B: Planning and Design, vol. 33, pp. 705-725.
6. Cruccitti, P., Latora, V. and Porta, S. (2006). Centrality in networks of urban streets. Chaos, vol.16, 066107.
7. Batty, M. (2005). Network geography: relations, interactions, scaling and spatial processes in GIS. Unwin D.J. and Fisher, P. (eds.), Re-presenting Geographical Information Systems. John Wiley and Sons, Chichester.
8. Ratti, C. (2004). Space syntax: some inconsistencies. Environment and Planning B: Planning and Design, vol.31, pp.513-516.

9. Freeman, L.C. (1977). A set of measures of centrality based on betweenness. Sociometry, vol.40, pp.35-41.
10. Freeman, L.C. (1979). Centrality in social networks: conceptual clarification. Social Networks, vol.1, pp.215-239.
11. Porta, S., Crucitti, P. and Latora, V. (2009). Street centrality and densities of retail and services in Bologna, Italy. Environment Planning B: Planning and Design, vol.36, pp.450-465.
12. Porta, S., Latora, V., Wang, F., Rueda, S., Strano, E., Scellato, S., Cardillo, A., Belli, E., Cardenas, F., Cormenzana, B. and Latora, L. (2012). Street centrality and the location of economic activities in Barcelona. Urban Studies, vol.49, no.7 pp.1471-1488.
13. Wang, F., Antipova, A. and Porta, S. (2011). Street centrality and land use intensity in Baton Rouge, Louisiana. Journal of Transport Geography, vol.19, pp.285-293.
14. Erath, A., Löchl, M. and Axhausen, K. (2009). Graph-Theoretical Analysis of the Swiss Road and Railway Networks over Time. Networks and Spatial Economics, vol.9, pp.379-400.
15. Foti, F., Waddell, P. and Luxen, D. (2012). A Generalized Computational Framework for Accessibility: From the Pedestrian to the Metropolitan Scale. Presented at 4th Conference on Innovations in Travel Modeling (ITM), April 30-May 2, Tampa, FL.
16. Huang, A. and Levinson, D. (2013). The structure and evolution of a skyway network. The European Physical Journal Special Topics, vol.215, pp.123-134.
17. Wang, P., Hunter, T., Bayen, A.M., Schechtner, K. and González, M.C. (2012). Understanding Road Usage Patterns in Urban Areas. Scientific Reports, vol.2, 1001.
18. Walls, D. (2007). National Establishment Time Series Database: Data Overview. Presented at 2007 Kauffman Symposium on Entrepreneurship and Innovation Data.
19. Sabidussi, G. (1966). The centrality index of a graph. Pschometrika, vol.31, pp.581-603.
20. Bhat, C., Handy, S., Kockelman, K., Mahmassani, H., Chen, Q. and Weston, L. (2000). Development of an Urban Accessibility Index: Literature Review, Center for Transportation Research, University of Texas at Austin.
21. Svetsuk, A. and Mekonnen, M. (2012). Urban Network Analysis: A new toolbox for ArcGIS. Revue Internationale de Geomatique, vol.22, pp.287-305.
22. Borgatti, S.P. and Everett, M.G. (2005). A Graph-theoretic perspective on centrality. Social Networks, vol.28, pp.466-484.
23. Vermunt, J.K. and Magindson, J. (2005). Technical guide for latent gold 4.0: basic and advanced. Statistical Innovations Inc., Belmont, MA.
24. Hoch, I. and Waddell, P. (1993). Apartment rents: another challenge to the monocentric model. Geographical Analysis, vol.25, pp.20-34.
25. Berry, B.L.J. and Kim, H. (1993). Challenges to the monocentric model. Geographical Analysis, vol.25, pp.1-4.
26. Newman, P. and Kenworthy, J. (1999). Sustainability and Cities: Overcoming Automobile Dependence. Island Press, Washington D.C.
27. van Ness, A. (2005). Spatial conditions for a typology of shopping areas. Proceedings of the 3rd Spatial Syntax Symposium, Atlanta, GA.
28. Heikkila, E.P., Gordon, P., Kim, J., Peiser, R., Richardson, H. and Dale-Johnson, D. (1989). What happened to the CBD-distance gradient? Land values in a polycentric city. Environment and Planning A, vol.21, pp.221-232.

29. Wang, F. (2000). Modeling commuting patterns in Chicago in a GIS environment: a job accessibility perspective. Professional Geographer, vol.52, pp.120-133.
30. Lang, R.E. (2003). The Edgeless City: Exploring the Elusive Metropolis. Washington D.C. The Brookings Institution Press.
31. Fujita, M. and Thisse, J.F. (2002). Economies of agglomeration: cities, industrial location and regional growth. Cambridge University Press, Cambridge, U.K.
32. Ravulaparthy, S.K., Goulias, K.G., Sweeney, S.H. and Kyriakidis, P.C. (2013). Exploring the spatial and temporal patterns of business concentration and dispersion: A case-study analysis for County of Santa Barbara. Presented at 52nd Annual Meeting of Western Regional Science Association, February 24-27, 2013, Santa Barbara, CA.