Transportation (2017) 44:1383–1402
DOI 10.1007/s11116-016-9706-6

A complex network perspective for characterizing urban travel demand patterns: graph theoretical analysis of large-scale origin–destination demand networks

Meead Saberi[1] · Hani S. Mahmassani[2] · Dirk Brockmann[3] · Amir Hosseini[1]

Published online: 14 May 2016

Abstract Urban travel demand, consisting of thousands or millions of origin–destination trips, can be viewed as a large-scale weighted directed graph. The paper applies a complex network-motivated approach to understand and characterize urban travel demand patterns through analysis of statistical properties of origin–destination demand networks. We compare selected network characteristics of travel demand patterns in two cities, presenting a comparative network-theoretic analysis of Chicago and Melbourne. The proposed approach develops an interdisciplinary and quantitative framework to understand mobility characteristics in urban areas. The paper explores statistical properties of the complex weighted network of urban trips of the selected cities. We show that travel demand networks exhibit similar properties despite their differences in topography and urban structure. Results provide a quantitative characterization of the network structure of origin–destination demand in cities, suggesting that the underlying dynamical processes in travel demand networks are similar and evolved by the distribution of activities and interaction between places in cities.

Keywords Complext networks · Network science · Travel demand · Melbourne · Chicago

Electronic supplementary material The online version of this article (doi:10.1007/s11116-016-9706-6) contains supplementary material, which is available to authorized users.

✉ Meead Saberi
meead.saberi@monash.edu

1 Department of Civil Engineering, Institute of Transport Studies, Monash University, Melbourne, Australia

2 Transportation Center, Northwestern University, Evanston, USA

3 Institute for Theoretical Biology, Humboldt Universität zu Berlin, Berlin, Germany

Introduction

With the growing global urban population, many cities around the world are under pressure because of the increasing demand for activities and travel. Activity- and agent-based models of travel demand are helping understand mechanisms underlying travel choices in their spatial context (Axhausen and Garling 1992; Bowman and Ben-Akiva 2001; Carrasco and Miller 2006; Timmermans and Zhang 2009); however, travel demand has not typically been analyzed from a network perspective. Advances in transportation network modeling have focused on the supply side, mainly modeling within-day equilibria and day-to-day evolution of traffic flows (Mahmassani 2001; Watling and Hazelton 2003), and exploring dynamics of network traffic flow (Daganzo 2007; Geroliminis and Daganzo 2008; Mahmassani et al. 2013). More recently, availability of vehicle trajectory data from GPS devices has helped further characterize driver route choices and network flow properties (Saberi et al. 2014; Joubert and Axhausen 2013; Kim and Mahmassani 2015). In this study, we are motivated by the recognition that travel demand in cities, consisting of thousands or millions of origin–destination trips, can be viewed as a large-scale complex weighted directed graph or network. Such networks exhibit additional degree of complexity because of the link weight representing intensity of interaction between nodes.

Complex networks theory is a highly active interdisciplinary research area inspired by numerous empirical studies of computer and social networks (Faloutsos et al. 1999; Yook et al. 2002; Siganos et al. 2003; Watts 2003). A system consisting of several non-identical elements connected by diverse interactions can be viewed as a complex network where the nodes are the system elements and the links are the interactions between the elements. Examples of such complex networks are the internet (Faloutsos et al. 1999), friendship networks (Eagle et al. 2009), business relationship networks (Costa and Baggio 2009), scientific collaboration networks (Newman 2001), virtual gaming networks (Keegan et al. 2010), airline and cargo ship networks (Woolley-Meza et al. 2011), genetic interaction networks (Kelley and Ideker 2005), and protein–protein interaction networks (Rual et al. 2005).

More recently, understanding human mobility patterns in different scales from a complex network perspective has become an attractive research topic for statistical physicists, applied mathematicians, and social scientists. Brockmann et al. (2006) analyzed the circulation of bank notes in the United States as a proxy to the traveling behavior of humans and showed that human traveling patterns can be described by a two-parameter continuous time-random walk model. Several other studies have used phone call, social network and mobile data to explore and verify the scaling laws of human mobility from urban to global scale (Brockmann et al. 2006; González et al. 2008; Jiang et al. 2009; Bazzani et al. 2010; Song et al. 2010; Roth et al. 2011; Woolley-Meza et al. 2011; Noulas et al. 2012; Simini et al. 2012; Liang et al. 2012; Kang et al. 2012; Peng et al. 2012; Schneider et al. 2013). Studies by Colak et al. (2013), Wang et al. (2012) and Hasan et al. (2013) applied complex network-driven measures to study mobility characteristics in urban areas. Iqbal et al. (2014), Toole et al. (2015), Colak et al. (2015), and Widhalm et al. (2015) used cell phone data to explore patterns in urban activities, mainly inferring origin–destination matrices.

The idea of viewing individuals' travel patterns as a tree, graph, or network is not new. Numerous studies in the past have looked at human activity spaces to study the spatial behavior of individuals (Theriault et al. 2002; Schonfelder and Axhausen 2003; Fan and Khattak 2008; Kamruzzaman and Hine 2012; Rai et al. 2007). Activity space is an environment where a person travels between destinations for his or her daily activities (Golledge and Stimson 1997; Schonfelder and Axhausen 2003). It is also known as a traveler cognitive or mental map, personal world, activity repertoire, and expectation space (Lynch 1960; Gould and White 1974; Downs and Stea 1973). In a recent series of studies, Betty (2013) argues that to better understand cities, one must understand flows and networks to demonstrate relationships between people and places. Notwithstanding the growing number of studies on traveler-activity space, the collective structure and properties of combined individuals' activity spaces in an urban environment have not been thoroughly studied before from a network perspective. Traditional methods of analyzing and predicting urban travel demand focus more on the attributes of individuals and land use characteristics of locations with less emphasis on understanding and predicting sets of interactions between different elements in an urban system. Gravity based models of travel demand are perhaps the simplest representation of spatial interactions between locations. Alternatively, utility-based destination choice models take into account the behavioral factors underlying travel demand. In this paper, we try to demonstrate the further potential insights that can be obtained by extending existing demand analysis methods to complexity theory driven approaches. The idea is that urban travel demand can be better understood through analyzing its network structure. Network-based analysis of travel demand provides a deeper understanding of the spatial interdependencies and interactions between locations. A challenging next step is to couple various types of networks. An example is the recent study by Chen et al. (2015) exploring the influence of social network on travelers' destination choice. They show that social interactions play a role in travelers' choice of destination by observing a possible correlation between travelers' behavior and the influence from their friends.

Here we introduce a complex network-driven approach to understand and characterize urban travel demand as a complementary tool to existing advanced activity- and agent-based models, building upon recent studies by González et al. (2008) and Schneider et al. (2013). The analysis investigates whether the patterns observed in previous complex network studies are reproducible specifically in the urban transportation context using household travel survey data. We compare selected network characteristics of travel demand in two cities, presenting a comparative network-theoretic analysis of urban travel demand in Chicago and Melbourne. The approach provides an interdisciplinary and quantitative framework to understand and characterize statistical properties of the complex network of urban trips. We show that travel demand networks in these two cities exhibit similar properties despite the differences in topography and urban structure. The resulting insights from viewing travel demand as a complex network uncovers interesting spatial phenomena in cities. Results suggest that the underlying dynamical processes in travel demand networks are driven by the interaction strength between places (or nodes) as previously observed in other types of networks (Newman 2010). These results provide a first step towards a new methodological basis for calibration and validation of activity- and agent-based travel demand models, which still requires further research. Overall, the study addresses two main research questions: (a) What can we learn from characterizing urban travel demand using complex network-theoretic measures? (b) Does travel demand networks in different cities exhibit similar statistical characteristics?

The remainder of the paper is organized as follows. In "Background on human mobility characteristics" section, we provide a summary of previous studies including their data, scale, and key findings. In "Complex network of urban travel demand: concept and data" section, we briefly explain applicable concepts from complex network theory and describe the data used in the paper. "Statistical properties of urban travel demand networks" section provides a comparative statistical analysis of the complex network of trips in Melbourne and Chicago. In "Spatial analysis of travel demand network properties" section, we study the effect of spatial form and geographical context on the selected network properties. "Shortest path trees and effective distance" section further explores the properties and structure of the complex network of trips from the perspective of a node in the network, using shortest path trees and effective distances. The "Conclusion" section concludes the paper.

Background on human mobility characteristics

Numerous studies have analyzed human mobility characteristics in different scales using different data sources. Several studies suggested that mobility patterns represented by distance traveled l, follows a power law $P(l) \propto l^{-\beta}$ (Brockmann et al. 2006; González et al. 2008; Jiang et al. 2009; Song et al. 2010) while others proposed that an exponential law $P(l) \propto e^{-\beta l}$ provides a better fit (Bazzani et al. 2010; Roth et al. 2011; Liang et al. 2012; Peng et al. 2012; Kang et al. 2012; Noulas et al. 2012). The observed differences are known to be dependent on the travel mode and spatial scale of the data. Other studies expanded the use of mobile phone calls and taxi GPS data to characterize urban travel demand and to properly represent individuals' daily travel patterns (Ratti et al. 2006; Calabrese et al. 2011; Çolak et al. 2015). Table 1 provides a summary of selected previous studies.

Complex network of urban travel demand: concept and data

Urban transportation can be viewed as a complex densely connected network of individuals' activity spaces. In this section, we provide a quantitative description of such networks using complex network theory (Newman 2010) where pairs of nodes i and j represent origins and destinations which are connected by links with non-negative weights $w_{ij} > 0$ if one or more trips are made between the nodes. If no trip is made between a pair of nodes, $w_{ij} = 0$. In such networks, one could possibly find an indirect path between any pair of nodes; w_{ij} quantifies the number of trips between pairs of nodes per unit of time. See Fig. 1 for an illustration of daily activity spaces of randomly selected individuals in a household as a small sample network and the spatial distribution of origins and destinations (nodes) in the Melbourne metropolitan area.

The complex network of trips in Chicago is constructed using the Chicago household travel survey data. The network includes 78,681 trips made between 1868 nodes covering the entire northeastern Illinois region, USA. Nodes in Chicago represent Census blocks that are the smallest statistical subdivisions of a county, having on average a population of 4000. To construct the travel demand network for Melbourne, we use the Victorian integrated survey of travel and activity (VISTA) data with 133,938 trips and 9310 nodes covering greater Melbourne, Geelong and regional centers in Victoria, Australia. Similarly, nodes in Melbourne represent census collection districts (CCD) that are the smallest geographical areas defined in the Australian standard geographical classification. Each CCD contains an average of about 250 dwellings. Nodes in both datasets may contain a variety of land uses including residential, commercial, institutional, industrial, parks, etc. In fact, they could represent a location or destination of any trip.

Table 1 A summary of selected recent literature on human mobility characteristics

Study	Data source	Scale	Key findings
González et al. (2008)	Mobile phone	Urban, regional	Human trajectories show a high degree of temporal and spatial regularity, each individual being characterized by a time independent characteristic travel distance and a significant probability to return to a few frequented visited locations.
Brockmann (2009, 2010) and Thiemann et al. (2010)	Wheresgeorge.com (money circulation)	National	High degree of symmetry in multi-scale mobility networks. Larger nodes are strongly connected to a larger number of other nodes. Travel of bank notes exhibits a Lévy-like pattern. Human mobility borders are very different to administrative borders.
Song et al. (2010)	Mobile phone	Unknown	- Human mobility patterns are highly predictable. - Most individuals are well localized.
Woolley-Meza et al. (2011)	World-wide air transportation network and global cargo-ship network	Global	Degree, flux and weight distributions do not fit power law. The effective distance approach gives a better representation of a large network.
Nguyen and Szymanski (2012)	Gowalla (social network)	Unknown	Network congestion is dramatically affected by friendship mobility patterns. A friendship mobility model (FMM) is presented.
Noulas et al. (2012)	Foursquare	Global	Human urban movements do not follow power law.
Riccardo et al. (2012)	GPS	Regional	Distribution of average daily trips follows an exponential law. Urban short trips distribution does not fit a power law. A power law only fits the distribution of long distance urban travels.
Liang et al. (2012, 2013)	Beijing taxis GPS, Chicago and Los Angeles household travel tracker survey	Urban, regional	Geographic origins and destinations distributions follow similar patterns. Distribution of trip distances depends of geographic distribution of human travel demands. Trip length distribution is best fitted by exponential law.
Schneider et al. (2013)	Individuals' travel survey, mobile phone	Urban	Data from two cities exhibit the same set of ubiquitous networks that can reveal general human mobility characteristics.

Both networks represent a relatively small sample (near 0.5 %) of the total number of trips made daily in those regions. Here we assume that the obtained sample data provide a true representation of travel demand of the entire population in the selected cities. The size

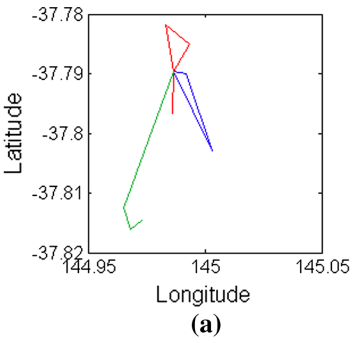


Fig. 1 **a** Illustration of daily activity spaces of three members of a randomly selected household in Melbourne; **b** spatial distribution of origins and destinations in Melbourne

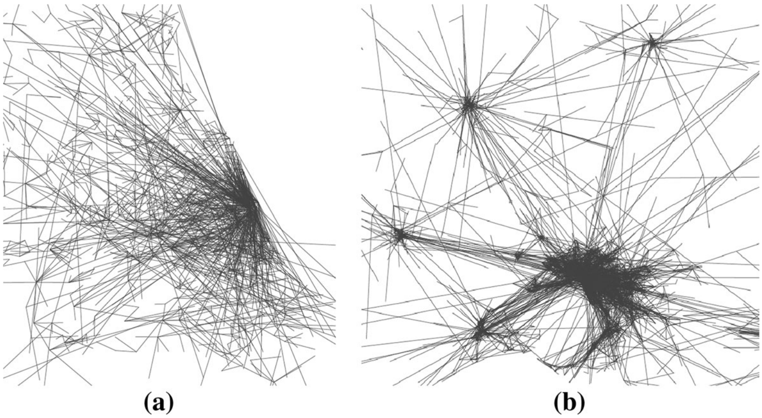


Fig. 2 **a** The complex network structure of a sample of trips in the Chicago metropolitan area, USA. Number of trips = 78,681, number of links = 37,528, and number of nodes = 1868, **b** The complex network structure of a sample of trips in the Melbourne metropolitan area, Australia. Number of trips = 133,938, number of links = 63,916, and number of nodes = 9310. See Figs S1 and S2 in the supplementary material for a higher resolution zoomed in version of the figures

of the sample data from Melbourne is 1.7 times larger compared to Chicago. For both Chicago and Melbourne networks, the resulting weight matrix W is not fully symmetric, $w_{ij} \neq w_{ji}$. Figure 2 shows the regional coverage and qualitative structure of each network. Despite their significant topological and structural differences, both networks exhibit surprising similarities, as discussed in the next sections.

Figure 3 illustrates a comparative analysis of probability density functions of individual's distance traveled l in kilometers per trip and activity duration d in minutes in the selected cities. An earlier study by Brockmann et al. (2006) showed that distance traveled obtained from bank note dispersal follows a power law with a scaling exponent of $\beta = 1.59$. Using mobile phone data, a later study by González et al. (2008) showed that individual's distance traveled follows a truncated power law with a scaling exponent of $\beta = 1.75$ not far from the previously observed value. Here we use actual distance traveled on the road network, or so-called network distance traveled, different from the

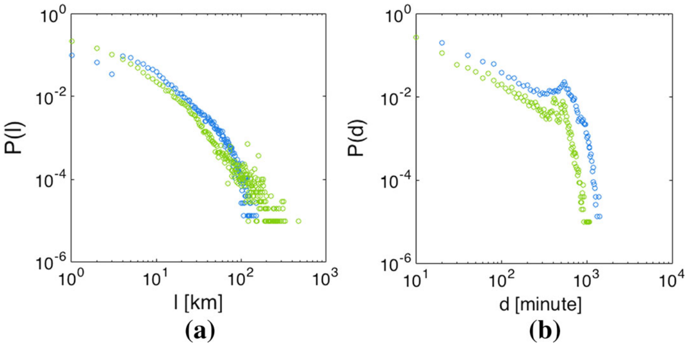


Fig. 3 Comparative analysis of probability density functions of **a** network distance traveled l (km) and **b** individual activity duration d (min) in Chicago (*blue*) and Melbourne (*green*). (Color figure online)

geographical direct line distance between an origin and a destination. Results suggest that a power law fit does not entirely represent the probability distribution of the distance traveled in both cities, consistent with findings of Noulas et al. (2012). We found that a truncated power law function with an exponential cutoff reasonably fits the distributions, confirming findings of González et al. (2008).

$$P(l) = (l + l_0)^{-\beta}\exp(-l/\kappa) \quad (1)$$

with exponents $\beta = 1.03 \pm 0.06$ and 1.29 ± 0.02, $l_0 = 3.6 \pm 0.7$ km and 2.1 ± 0.0 km, and cutoff values $\kappa = 16.6 \pm 2.4$ km and 19.0 ± 1.6 km for Chicago and Melbourne, respectively for $l < 100$ km. However, the estimated βs are smaller than the scaling exponents observed in earlier studies.

More interestingly is the distribution of activity durations that exhibit an abrupt cutoff at $d \sim 600$ min (10h). The activity duration distributions also show two peaks at $d \sim 8$ h and $d \sim 10$ h. Note that activity durations considered in this study exclude the duration if an individual stays home at night until next morning before making the first trip in the day after 3:00 am. Further research should explore travel demand as a temporal network considering time-dependency of trips and activity durations.

Statistical properties of urban travel demand networks

In this section, we provide a comparative quantitative analysis of the statistical properties of the selected networks. The original networks are first reduced to their largest connected component. The resulting connected network of trips in Melbourne comprises more than three times as many nodes N as in Chicago while its number of links L is less than two times larger. This yields to a less densely connected network of trips in Melbourne with $\delta = 3 \times 10^{-3}$ and $L/N = 10.63$ compared to Chicago with $\delta = 21 \times 10^{-3}$ and $L/N = 20.1$ where $\delta \approx 2L/N^2$, representing network connectivity, assuming the weight matrix is almost symmetric, as summarized in Table 2.

The node degree k is the number of links connected to a node in a network where a_{ij} are elements in the adjacency matrix.

Table 2 Travel demand network characteristics of Chicago and Melbourne

	Chicago	Melbourne	Ratio (Melbourne/Chicago)
N	1867	5998	3.21
L	37,527	63,788	1.70
L/N	20.1	10.63	0.53
δ	21e-3	3e-3	0.14
T	78,680	133,754	1.70
$\langle l \rangle$ (km)	11.7	8.2	0.70
$\langle d \rangle$ (h)	3.9	1.9	0.48
$\langle F \rangle$ (trips)	42.1	22.3	0.53
$\langle k \rangle$	20.10	10.63	0.53
$\langle w \rangle$	2.10	2.10	1
$CV(l)$	1.23	2.02	1.64
$CV(d)$	1.07	1.41	1.32
$CV(F)$	1.15	2.01	1.74
$CV(k)$	0.93	1.40	1.50
$CV(w)$	1.73	1.56	0.90
$\langle c \rangle$	0.30	0.36	1.22
$\langle wc \rangle$	0.16	0.18	1.18
C	0.18	0.15	0.86
d_T	3.03	4.05	1.33
wd_T	3.28	4.78	1.45
φ	7	12	1.71
$w\varphi$	11	23	2.09

Number of nodes N, number of edges L, network connectivity $\delta = 2L/N^2$, total number of trips T, mean distance traveled in km l, mean activity duration in hours d, mean node flux F, mean node degree k, mean edge weight w, and associated coefficients of variation, mean clustering coefficient c, mean weighted clustering coefficient wc, network clustering coefficient C, weighted and unweighted average shortest path wd_T and d_T, weighted and unweighted network diameter $w\varphi$ and φ

$$k_i = \sum_j a_{ij} \tag{2}$$

The average node degree in the network of trips in Chicago is $\langle k \rangle = 20.1$ which is near two times greater than the average node degree in Melbourne $\langle k \rangle = 10.6$. This is mainly due to the smaller sample size in Chicago and the larger geographical coverage of data in Melbourne. This could also be interpreted as larger interaction between places in Chicago compared to Melbourne. For the same reason, the variability of $\langle k \rangle$ being measured by the coefficient of variation *CV(k)* is greater in Melbourne by one and a half times compared to Chicago, suggesting a larger heterogeneity in connectivity between nodes in Melbourne.

The node flux F is the number of trips starting or ending at a node where w_{ij} represents the weight or the number of trips between each pair of nodes.

$$F_i = \sum_j w_{ij} \tag{3}$$

The average node flux is also near two times greater in Chicago $\langle F \rangle = 42.1$ compared to Melbourne $\langle F \rangle = 22.3$. While this is perhaps due to the different characteristics of the collected survey data and different geographical coverage of the two samples, it could also suggest that interaction between places are stronger in Chicago. Similarly, the coefficient of variation of node flux *CV(F)* is greater in Melbourne by nearly two times compared to Chicago suggesting a more heterogeneous distribution of interaction strengths in Melbourne.

The average number of trips on each link is given by the mean link weight w, which is identical in both networks, in spite of the larger number of trips in the Melbourne network. We believe this is only a coincidence rather than a general phenomenon since the dispersion of link weights is different in the two cities; coefficients of variation *CV(w)* in Chicago and Melbourne networks are 1.73 and 1.56, respectively.

The clustering coefficient c is a measure of the degree to which the nodes in a network tend to cluster together, and is measured by the fraction of paths of length two in the network that are closed. This is simply the number of triangles that pass through a node.

$$c_i = \frac{(number\ of\ pairs\ of\ neighbors\ of\ i\ that\ are\ connected)}{(number\ of\ pairs\ of\ neighbors\ of\ i)} \tag{4}$$

The mean clustering coefficient c is larger in Melbourne compared to Chicago, suggesting that the complex network of trips in Melbourne is more locally connected despite being globally sparser. The clustering coefficient is also suggested to reflect the formation of groups or communities in networks (Newman and Park 2003). Alternatively, one can calculate the weighted clustering coefficient wc where the number of trips between neighboring nodes is also taken into consideration. The mean weighted clustering coefficient wc is found to be similarly larger in Melbourne compared to Chicago.

The network clustering coefficient C is a global measure of the extent to which nodes in a network are clustered. C is calculated as a ratio of the number of triangles to the number of connected triples of nodes, expressed as

$$C = \frac{(number\ of\ triangles) \times 3}{(number\ of\ connected\ triples)} \tag{5}$$

For Chicago and Melbourne $C = 0.18$ and 0.15, respectively. These values suggest that both cities have very similar network characteristics as typical social networks (Newman 2010). C in Melbourne is slightly larger than Chicago suggesting a lower connectivity between the nodes and larger spatial distribution of nodes.

The term d_T is defined as the average shortest path length between each pair of nodes in the network using the adjacency matrix. Similarly, wd_T is defined as the average weighted shortest path length between each pair of nodes using the weight matrix. Average shortest path is often used to measure network efficiency (Ye et al. 2010). For Melbourne, both d_T and wd_T are larger compared to Chicago, similarly suggesting a lower connectivity between the nodes and larger spatial distribution of nodes in Melbourne.

The network diameter φ is defined as the longest shortest path between each pair of nodes in the network using the adjacency matrix. The weighted network diameter $w\varphi$ is the longest weighted shortest path between each pair of nodes using the weight matrix. The diameter represents the linear size of a network. Both φ and $w\varphi$ are significantly larger in Melbourne, indicating that the network of trips in Melbourne is near two times larger in size compared to Chicago. This could also be interpreted as denser connectivity between nodes in Chicago representing higher interaction between places.

Table 3 Summary of the underlying physical meaning of selected network measures in the context of travel demand

Physical meaning	Network measure
Connectivity or interaction between places	Node degree $\langle k \rangle$
Heterogeneity in connectivity between places	Coefficient of variation of node degree *CV(k)*
Interaction strength between places	Node flux $\langle F \rangle$ Link weight ($\langle w \rangle$) Average shortest path d_T Network diameter φ
Heterogeneity in interaction strength between places	Coefficient of variation of node flux *CV(F)* Coefficient of variation of link weight *CV(w)*
Local interaction between places	Node clustering coefficient *c*
Global interaction between places	Network clustering coefficient *C*

Table 3 provides a summary explaining the underlying physical meaning of selected network measures. In general, network-based measures provide a quantifiable picture of interactions and interaction strengths between origins and destinations. Such comprehensive information about connectivity between places is not fully captured by gravity-based models of trip distribution and utility-based models of destination choice. While network-based characteristics of travel demand present a deeper understanding of interactions between places in a city, how to fully incorporate them in existing modeling approaches remains an open question.

These measures provide a basic quantitative picture of the complex network of travel demand in the selected cities. However, some of the observed differences are mainly due to the difference in sample size and geographical coverage of the data. To overcome this issue, we normalize each measure by the mean value of the same measure in each city and plotting the normalized cumulative distribution functions comparatively. Figure 4 provides a comparative illustration of the statistical properties of the complex network of trips in Chicago and Melbourne. Complementary cumulative distribution functions (CDF) of node degree k, node flux F, and link weight c normalized by the mean in each distribution k_0, F_0, and w_0 are plotted. Complementary CDF is defined as $P(X \geq x) = 1 - F(x)$ where F(x) is the CDF. Figure 4 shows that k, F, and c have, surprisingly, very similar distributions in

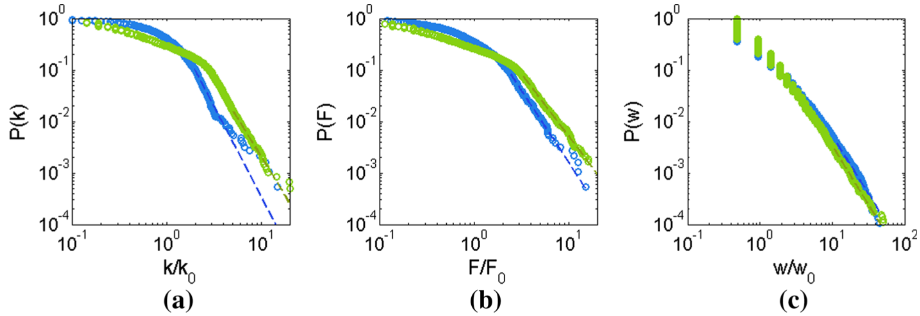


Fig. 4 Statistical properties of the complex network of trips in Chicago (*blue*) and Melbourne (*green*). Complementary CDF of **a** node degree k, **b** node flux F, and **c** link weight c normalized by the mean in each distribution k_0, F_0, and w_0. *Dashed lines* represent power law fits. (Color figure online)

both networks, suggesting that the underlying dynamics of these networks follow a similar fundamental process. The node degree distributions shown in Fig. 4a could be interpreted as a representation of the spatial distribution and accessibility of activities in a city. The probability of having a node with a larger degree k than the average degree k_0 is higher in Melbourne, since Melbourne's curve sits above Chicago for k/k_0 values larger than one (10^0). Also, the variability of k, *CV(k)*, is one and a half times larger in Melbourne than Chicago. Although this can be partially due to a larger size of the study area in Melbourne, it also suggests that activities are more homogeneously distributed in Chicago than in Melbourne. The node flux distribution for both cities is shown in Fig. 4b. Here, we interpret the node flux distribution as a representation of the level of attractiveness for different locations. Also, both k and F in both cities exhibit a pronounced kink in the center of the distribution. Finding an explanation for the observed phenomenon requires further research. Figure 4c shows the distribution of link weights. In the urban travel demand context, link weight w is a representation of travel demand between two locations. We observed that the distribution of link weights are very similar in both cities, despite the small difference in variability of w, *CV(w)*.

Another commonly used network measure is node and link betweenness centrality. The betweenness centrality measure b reflects the importance or centrality of a node or a link in a network. It is computed as the fraction of shortest paths in the entire network that pass through a particular node or link. b indicates the extent to which a node or link falls on the path between other nodes or links. The weighted betweenness centrality is computed

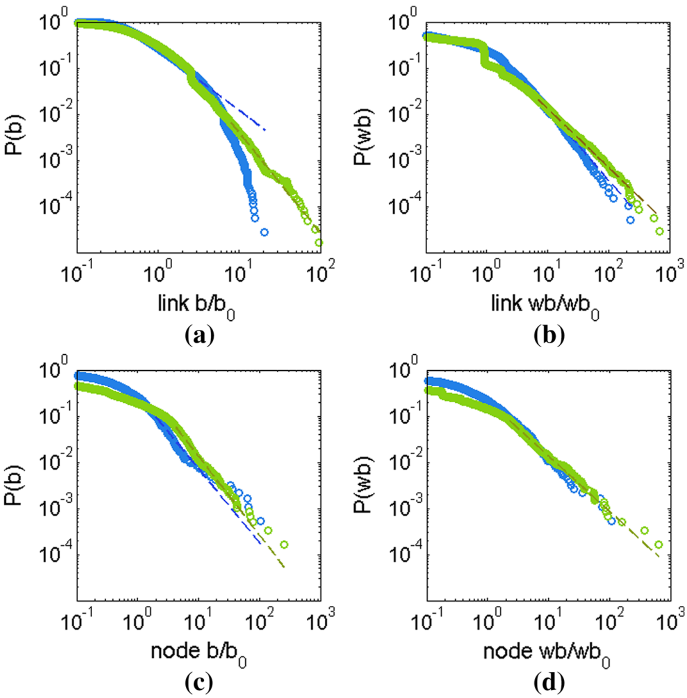


Fig. 5 Betweenness centrality in Chicago (*blue*) and Melbourne (*green*). Complementary CDF of **a** link betweenness centrality b, **b** link weighted betweenness centrality wb, **c** node betweenness centrality b, and **d** node weighted betweenness centrality wb normalized by the mean in each distribution b_0. *Dashed lines* represent power laws. (Color figure online)

similarly. However, the shortest paths are calculated using $1/w_{ij}$ instead of the elements of the adjacency matrix. Figure 5 shows the distributions of node and link betweenness centrality *P(b)* and weighted betweenness centrality *P(wb)* in Melbourne and Chicago.

Both networks exhibit similar behavior in weighted betweenness distributions, suggesting the existence of a fundamental process in urban travel demand despite the structural and topographical differences in those cities. The broad betweenness distributions in Melbourne could be explained by the multi-centric topology of the study region with some nodes having substantial betweenness and potentially a smaller redundancy of paths. The analysis of betweenness centrality provides more insight if conducted in the context of travel supply; rather than travel demand. A more effective way of analyzing betweenness centrality in the travel demand context is to study network measurements based on individuals' activity spaces rather than analyzing the collective characteristics of the entire network which could be an interesting direction for future research.

Consistent with findings from Woolley-Meza et al. (2011), we also found that simple power laws do not reasonably fit to these distributions despite their similarity and existence of a scaling behavior. The tail of the degree distributions exhibits a power law behavior, however, with a relatively large exponent that does not imply a scale-free network. Table 4 summarizes the power law fitting results based on the method proposed in Clauset et al. (2009) and Virkar and Clauset (2014) using a series of open source tools developed by the Santa Fe Institute accessible via http://tuvalu.santafe.edu/~aaronc/powerlaws/.

Table 4 Power-law fitting results for normalized node degree *k*, node flux *F*, link weight *c*, link and node betweenness centrality *b*, and link and node weighted betweenness centrality *wb*

	α	x_{min}
k/k_0		
Chicago	4.4964	1.8905
Melbourne	3.9844	4.8426
F/F_0		
Chicago	3.7962	2.2780
Melbourne	3.4109	3.0942
w/w_0		
Chicago	3.4073	12.8778
Melbourne	3.3112	14.3072
link b/b_0		
Chicago	2.3291	0.6495
Melbourne	3.1807	7.1866
link wb/wb_0		
Chicago	2.5326	3.5712
Melbourne	2.2773	6.1830
node b/b_0		
Chicago	2.6061	1.2011
Melbourne	2.7091	3.7966
node wb/wb_0		
Chicago	2.2294	1.4455
Melbourne	2.2016	2.0849

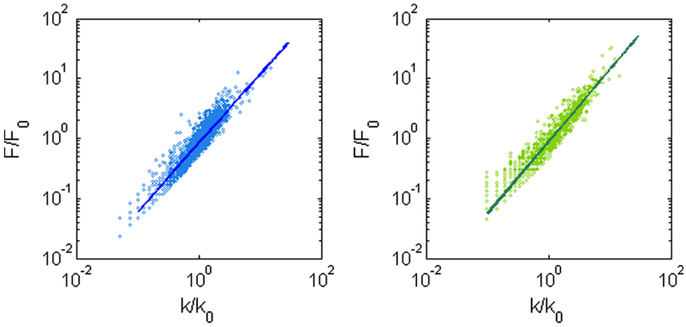


Fig. 6 Correlation between node degree and node flux in the complex network of trips in Chicago (**a**) and Melbourne (**b**)

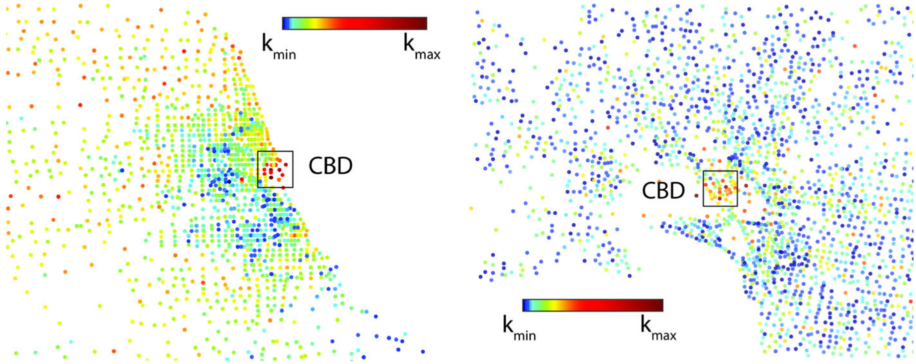
$$p(x) = x^{-\alpha} \quad \text{for} \quad x \geq x_{min} \tag{6}$$

A positive correlation and scaling behavior between node degree and node flux has been previously observed in other types of networks (Woolley-Meza et al. 2011). We found a similar correlation in the complex network of urban trips. The observed behavior can be described as a super-linear scaling relationship $F \propto k^{\beta}$ shown by the solid line in Fig. 6 with the scaling exponent of 1.20 and 1.14 for Chicago and Melbourne, respectively. Furthermore, Fig. 6 shows that as node degree increases, node flux also increases at an almost equal rate, in both cities. This suggests that highly visited nodes are also well connected to other nodes.

Spatial analysis of travel demand network properties

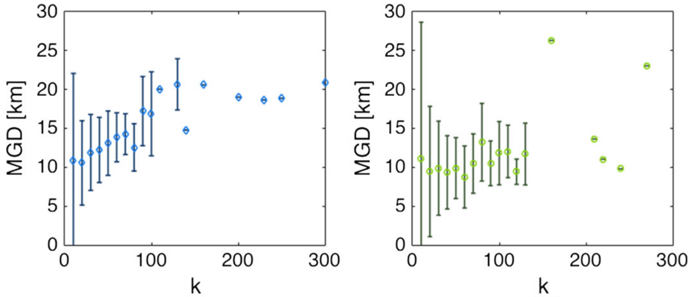
Travel demand networks, unlike some biological or technological networks, have a significant spatial dimension. Figure 7 illustrates the spatial distribution of node degrees in Chicago and Melbourne. The spatial distribution follows a heterogeneous pattern in which a cluster of high degree nodes is located in the central business district (CBD) of both cities, as expected. An interesting observation is that as the radial distance from Chicago CBD increases, nodes tend to have lower degrees to an extent where node degrees slightly increase and remain less variant afterward. We believe this is due to the change in the density of the nodes in space in the outer suburbs because of the larger size of the census tracts and lower population density. The observed change suggests that spatial aggregation of nodes coupled with population density could affect network properties. Further research is required to better understand the effect of spatial aggregation of nodes on the statistical properties of urban travel demand networks.

To further explore the influence of spatial form and geographical extent of both cities on their network characteristics, we define MGD_i as the mean geographical distance of node i to all other nodes connected to node I as represented in the adjacency matrix. Figure 8 illustrates MGD (mean ± SD) as a function of node degree grouped in equal sized bins. As node degree increases in Chicago, MGD also increases with less variability suggesting that places with higher connectivity to other places in Chicago also connect to places that are geographically farther. The trend is, however, different in Melbourne with roughly



(a) **(b)**

Fig. 7 Spatial distribution of node degrees in **a** Chicago and **b** Melbourne following a heterogeneous pattern



(a) **(b)**

Fig. 8 Mean geographical distance (MGD) ±SD as a function of node degree in **a** Chicago and **b** Melbourne

constant MGD with decreasing variability suggesting that as node connectivity increases in Melbourne, nodes do not necessarily connect with other nodes located in a farther geographically distance. This suggests that the travel demand network in Melbourne is more locally connected compared to Chicago.

Shortest path trees and effective distance

In this section, we explore the structure of the network of travel demand from the perspective of a chosen node. Here, we use a different measure of distance, namely effective distance as previously introduced and used in Thiemann et al. (2010), Woolley-Meza et al. (2011), and Brockmann and Helbing (2013). We measure effective distance as the reciprocal of the weight of a link $1/w_{ij}$. One could also use the normalized weight $P_{ij} = \frac{w_{ij}}{\sum_i w_{ij}}$ and measure effective distance as $d_{ij} = 1 + logP_{ij}$ as introduced in Brockmann and Helbing (2013). Based on this concept, places that are connected by larger traffic w_{ij} are

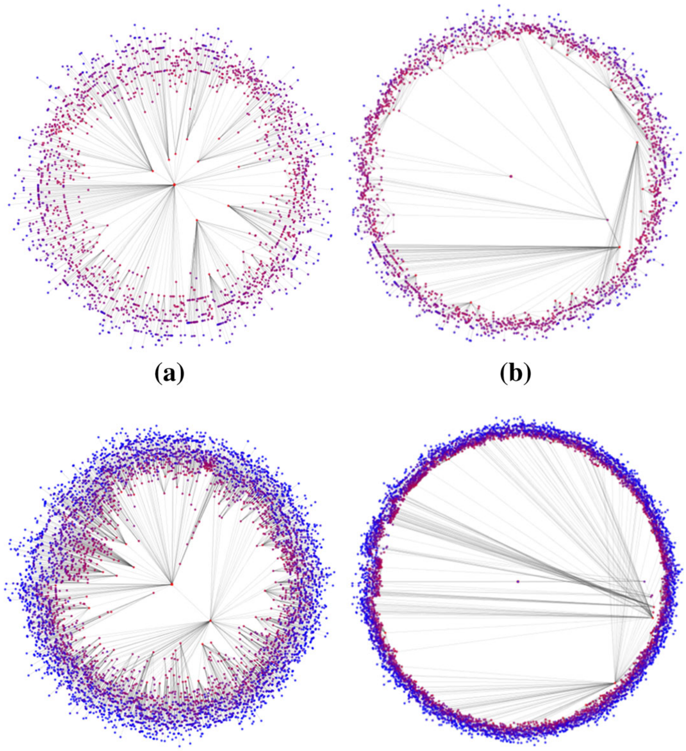

(a) (b) (c) (d)

Fig. 9 Illustration of the shortest path tree structure using the effective distance concept from the perspective of **a** a node in the central business district of Chicago, **b** a node in the western suburbs of Chicago, **c** a node in the central business district of Melbourne, and **d** a node in the northern suburbs of Melbourne

effectively closer. Effective distance of a randomly selected node is best illustrated via a shortest path tree. A shortest path tree approximately visualizes the most probable path from a random node in a network to the other nodes based on effective distances. Figure 9 presents the shortest path tree plots from the perspective of two different nodes in Chicago and Melbourne.

The shortest path tree for Melbourne looks more compact than Chicago's, simply because of the larger number of nodes in the sample data from Melbourne. The selected node in Chicago CBD as shown in Fig. 8a is connected to several other well-connected nodes in a short effective distance, whereas this is not the case in Melbourne. This supports our previous argument that activities seem to be more homogeneously distributed in Chicago. The circumference of the circular illustrations in Fig. 8b, d shows that almost all nodes are located at a roughly same effective distance to the selected nodes. This suggests that the selected nodes are among low attractive or least visited locations in both cities. To better understand the qualitative discussion above, next we plot the distribution of the shortest path distances. The shortest path between two nodes is usually, the geodesic

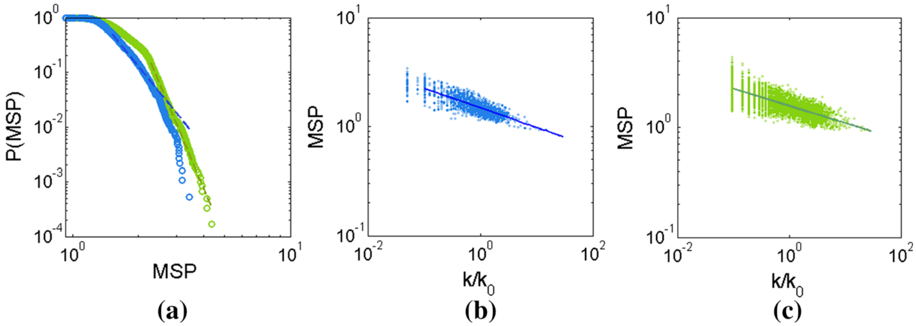


Fig. 10 **a** Complementary CDF of the mean shortest path (MSP) length using the effective distance concept in Chicago (*blue*) and Melbourne (*green*); **b** correlation between MSP and node degree in Chicago; and **c** correlation between MSP and node degree in Melbourne. (Color figure online)

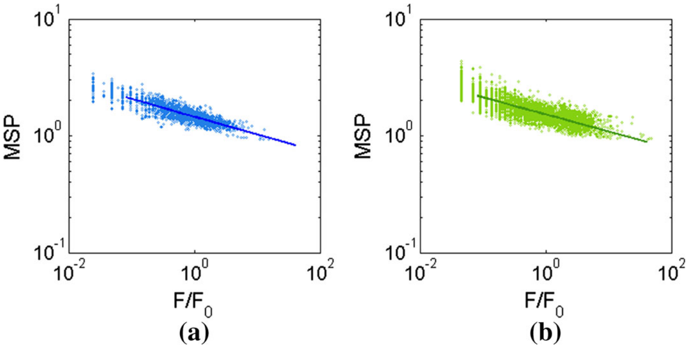


Fig. 11 Correlation between MSP and node flux in **a** Chicago and **b** Melbourne

distance between them, however, for the purpose of this study, the shortest path distributions are plotted using effective distances.

Figure 10a shows the distribution of mean shortest path (MSP) using the effective distance concept, for both cities. MSP generally represents the efficiency of information or mass transport in a network. Chicago is more likely to have a smaller MSP across the network compared to Melbourne. Since the distributions are plotted using the effective distance, the observed pattern suggests that on average there are smaller number of trips between any two locations in Melbourne.

Figures 10b, c show the correlation between MSP and node degree in Chicago and Melbourne respectively. Despite the difference between the urban structure and topology of the selected cities, the distributions, however, exhibit very similar characteristics. In both distributions, there is a larger variability in MSP as node degree or node flux decreases. The observed MSP and node degree correlation follows a scaling relationship with negative slope MSP $\propto k^{\beta}$ shown by the solid line with scaling exponent of -0.18 and -0.16 for Chicago and Melbourne, respectively. The MSP and node flux correlation (Fig. 11) follows a similar relationship with negative slope MSP $\propto F^{\beta}$ shown by the solid line with scaling exponent of 0.15 both for Chicago and Melbourne. Results suggest that locations with greater connectivity (larger node degree) and greater attractiveness (larger node flux) also enjoy smaller MSP lengths.

Conclusion

In this study, we applied a complex network-motivated approach to understand and characterize urban travel demand patterns using origin–destination data. The comparative analysis of the travel demand in Chicago and Melbourne presented here is a first step towards a better understanding of the structure, interactions, and evolution of travel demand networks in cities. Travel demand has been long viewed as a function of socio-economic and land use characteristics of locations and activities in space. In this paper, we suggest that the underlying processes in travel demand, viewed as a network, are also driven by the interaction strength between places (or nodes) consistent with recent arguments by Betty (2013). The new network perspective, as a supplementary approach to traditional travel demand analysis methods, better captures the influence of urban geography on mobility patterns. Network measures provide a clear picture of connectivity or interaction between places, origins and destinations. Such information can be of great value when evaluating travel demand model outcomes. Commonly used traditional methods to evaluate and validate travel demand models include comparison of aggregate measures of travel such as vehicle miles of travel, vehicle hours of travel, mode share, trip length distributions, and total trips or trip rates (Pearson et al. 2002; Pendalaya and Bhat 2006; Yagi and Mohammadian 2010). As a supplementary method, one can also perform a direct comparison of network measures obtained from empirical data and modeled data to measure the "goodness of fit" of a model from a network perspective. Pendalaya and Bhat (2006) highlighted the need to further identify behavioral paradigms and concepts to be incorporated into travel demand and activity models. One of such behavioral concepts is the spatial (location) interdependencies and interactions as analyzed in this paper. However, the full potential of application of network science in analyzing social interaction among people and between people and places in the context of travel demand modeling still requires further research.

We postulate that travel demand networks of the selected cities exhibit similar statistical properties despite their differences in topography and urban structure. This could set a new methodological basis for calibration and validation of travel demand models. Also, some of the observed network statistical properties of travel demand are interestingly similar to properties of other types of networks such as social and technological networks reported in the literature. In summary, results suggest that network of travel demand in Melbourne has larger heterogeneity in connectivity between nodes with a more heterogeneous distribution of interaction strengths compared to Chicago. We observed that Melbourne is more locally connected despite being globally sparser while Chicago enjoys denser connectivity between nodes representing higher interaction between places. This could generally be interpreted as a more homogenous distribution of activities in Chicago compared to Melbourne. We also found that highly visited nodes are also well connected to other nodes in both cities. Locations with greater connectivity (larger node degree) and greater attractiveness (larger node flux) also have smaller MSP lengths. A spatial analysis of both networks showed that the distribution of network properties follows a heterogeneous pattern in space as expected.

Although the observations made in this study were based on actual travel data from household travel surveys, we did not distinguish the trips by mode or purpose. Future research should focus on the network properties of travel demand by trip mode and purpose. Also, this study explored the collective structure and properties of combined individuals' activity spaces. Network measurements based on individuals' activity spaces could further shed light onto the underlying dynamical processes of travel demand as evolved by the interaction between places in cities.

References

Axhausen, K., Garling, T.: Activity-based approaches to travel analysis: conceptual frameworks, models, and research problems. Transp. Rev. **12**(4), 323–341 (1992)

Bazzani, A., Giorgini, B., Rambaldi, S., Gallotti, R., Giovannini, L.: Statistical laws in urban mobility from microscopic GPS data in the area of Florence. J. Stat. Mech. **2010**, P05001 (2010)

Betty, M.: The new science of cities. MIT press, Cambridge (2013)

Bowman, J.L., Ben-Akiva, M.E.: Activity-based disaggregate travel demand model system with activity schedules. Transp. Res. A **35**, 1–28 (2001)

Brockmann, D.: Human mobility and spatial disease dynamics. Reviews of nonlinear dynamics and complexity, pp. 1–24. Wiley-VCH, Weinheim (2009)

Brockmann, D.: Statistical mechanics: the physics of where to go. Nat. Phys. News Views **6**, 720–721 (2010)

Brockmann, D., Helbing, D.: The hidden geometry of complex, network-driven contagion phenomena. Science **342**, 1337 (2013)

Brockmann, D., Hufnagel, L., Geisel, T.: The scaling laws of human travel. Nature **439**, 462–465 (2006)

Calabrese, F., Di Lorenzo, G., Liu, C.L., Ratti, C.: Estimating origin-destination flows using mobile phone location data. IEEE Pervasive Comput. **10**(4), 36–44 (2011)

Carrasco, J.A., Miller, E.J.: Exploring the propensity to perform social activities: a social network approach. Transportation **33**(5), 463–480 (2006)

Chen, Y., Frei, A., Mahmassani, H. S.: Exploring activity and destination choice behavior in social networking data. Proceedings of the transportation research board 94th annual meeting (No. 15-5808), Washington (2015)

Clauset, A., Shilazi, C.R., Newman, M.E.J.: Power-law distributions in empirical data. Soc. Ind. Appl. Math. Rev. **51**(4), 661–703 (2009)

Colak, S., Schneider, C.M., Wang, P., González, M.C.: On the role of spatial dynamics and topology on network flows. New J. Phys. **15**, 113037 (2013)

Çolak, S., Alexander, L.P., Alvim, B.G., Mehndiretta, S.R., González, M.C.: Analyzing cell phone location data for urban travel: current methods, limitations and opportunities. Trans. Res. Rec. **2526**, 126–135 (2015)

Costa, Da F., Baggio, L.R.: The web connections between tourism companies: structure and dynamics. Physica A **388**, 4286–4296 (2009)

Daganzo, C.: Urban gridlock: macroscopic modeling and mitigation approaches. Transp. Res. B **41**(1), 49–62 (2007)

Downs, R.M., Stea, D.: Image and environments. In: Downs, R.M., Stea, D. (eds.) Cognitive maps and spatial behavior: process and products. Aldine, London (1973)

Eagle, N., Pentland, A.S., Lazer, D.: Inferring friendship network structure by using mobile phone data. Proc. Nat. Acad. Sci. USA **106**(36), 15274–15278 (2009)

Faloutsos, M., Faloutsos, P., Faloutsos, C.: On power-law relationship of the internet topology. ACM SIGCOMM Comp. Commun. Rev. **29**, 251–262 (1999)

Fan, Y., Khattak, A.: Urban form, individual spatial footprints, and travel: examination of space-use behavior. Transp. Res. Rec. **2082**, 98–106 (2008)

Geroliminis, N., Daganzo, C.: Existence of urban-scale macroscopic fundamental diagrams: some experimental findings. Transp. Res. B **45**(3), 605–617 (2008)

Golledge, R.G., Stimson, R.J.: Spatial behavior: a geographic perspective. Guilford, New York (1997)

González, M.C., Hidalgo, C.A., Barabási, A.L.: Understanding individual human mobility patterns. Nature **453**, 779–782 (2008)

Gould, P., White, R.: Mental maps. Pernguin Books, Harmondsworth (1974)

Hasan, S., Schneider, C.M., Ukkusuri, S.V., González, M.C.: Spatio-temporal patterns of urban human mobility. J. Stat. Phys. **151**, 304–318 (2013)
Iqbal, M.S., Choudhury, C.F., Wang, P., González, M.C.: Development of origin–destination matrices using mobile phone call data. Trans. Res. **40**, 63–74 (2014)
Jiang, B., Yin, J., Zhao, S.: Characterizing the human mobility pattern in a large street network. Phys. Rev. **80**, 1–11 (2009)
Joubert, J.W., Axhausen, K.: A complex network approach to understand commercial vehicle movement. Transportation **40**(3), 729–750 (2013)
Kamruzzaman, Md, Hine, J.: Analysis of rural activity spaces and transport disadvantage using a multi-method approach. Transp. Policy **19**(1), 105–120 (2012)
Kang, C., Ma, X., Tong, D., Liu, Y.: Intra-urban human mobility patterns: an urban morphology perspective. Physica A **391**, 1702–1717 (2012)
Keegan, B., Ahmed, M.A., Williams, D., Srivastava, J., Contractor, N.: Dark gold: statistical properties of clandestine networks in massively multiplayer online games. IEEE Second Int. Conf. Soc. Comput. (2010). doi:10.1145/1925041.1925043
Kelley, R., Ideker, T.: Systematic interpretation of genetic interactions using protein networks. Nat. Biotechnol. **23**, 561–566 (2005)
Kim, J., Mahmassani, H.S.: Spatial and temporal characterization of travel patterns in a traffic network using vehicle trajectories. Trans. Res. C **59**, 375–390 (2015)
Liang, X., Zheng, X., Lv, W., Zhu, T., Xu, K.: The scaling of human mobility by taxis is exponential. Physica A **391**, 2135–2144 (2012)
Liang, X., Zhao, J., Dong, L., Xu, K.: Unraveling the origin of exponential law in intra-urban human mobility. Sci. Rep. **3**, 2983 (2013)
Lynch, K.: The image of the city. MIT Press, Cambridge (1960)
Mahmassani, H.: Dynamic network traffic assignment and simulation methodology for advanced system management applications. Net. Spat. Econ. **1**(3–44), 267–292 (2001)
Mahmassani, H., Saberi, M., Zockaie, A.: Urban network gridlock: theory, characteristics, and dynamics. Trans. Res. C **36**, 480–497 (2013)
Newman, M.E.J.: The structure of scientific collaboration networks. Proc. Nat. Acad. Sci. USA **98**(2), 404–409 (2001)
Newman, M.: Networks: an introduction. Oxford University Press, Oxford (2010)
Newman, M., Park, J.: Why social networks are different from other types of networks. Phys. Rev. E **68**(3), 036122 (2003)
Nguyen, T., Szymanski, B.: Using location-based social networks to validate human mobility and relationships models. Proc. IEEE/ACM Int. Conf. Adv. Soc. Netw. Anal. Min. (2012). doi:10.1109/ASONAM.2012.210
Noulas, A., Scellato, S., Lambiotte, R., Pontil, M., Mascolo, C.: A tale of many cities: universal patterns in human urban mobility. Plos One **7**(5), e37027 (2012). doi:10.1371/journal.pone.0037027
Pearson, D., Ellis, P., Farnsworth, S.: Calibration of a past year travel demand model for model evaluation. Texas Trans. Inst. (2002). doi:10.1016/j.procs.2012.08.001. (Report No. FHWA/TX-03/4198-2)
Pendalaya, R., Bhat, C.: Validation and assessment of activity-based travel demand model systems. Proc. Innov. Travel Demand Model. **2**, 157–158 (2006)
Peng, C., Jin, X., Wong, K.-C., Shi, M., Liò, P.: Collective human mobility pattern from taxi trips in urban area. Plos One **7**, e34487 (2012)
Rai, R.K., Balmer, M., Rieser, M., Vaze, V.S., Schönfelder, S., Axhausen, K.W.: Capturing human activity spaces: new geometries. Trans. Res. Rec. **2021**, 70–80 (2007)
Ratti, C., Frenchman, D., Pulselli, R.M., Williams, S.: Mobile landscapes: using location data from cell phones for urban analysis. Environ. Plan. **33**, 727–748 (2006)
Riccardo, G., Armando, B., Sanro, R.: Towards a statistical physics of human mobility. Int. J. Mod. Phys. C **23**(9), 1250061 (2012)
Roth, C., Kang, S.M., Batty, M., Barthélemy, M.: Structure of urban movements: polycentric activity and entangled hierarchical flows. Plos One **6**, e15923 (2011)
Rual, J.F., Venkatesan, K., Hao, T., Hirozane-Kishikawa, T., Dricot, A., Li, N., Berriz, G.F., Gibbons, F.D., Dreze, M., Ayivi-Guedehoussou, N., Klitgord, N., Simon, C., Boxem, M., Milstein, S., Rosenberg, J., Goldberg, D.S., Zhang, L.V., Wong, S.L., Franklin, G., Li, S., Albala, J.S., Lim, J., Fraughton, C., Llamosas, E., Cevik, S., Bex, C., Lamesch, P., Sikorski, R.S., Vandenhaute, J., Zoghbi, H.Y., Smolyar, A., Bosak, S., Sequerra, R., Doucette-Stamm, L., Cusick, M.E., Hill, D.E., Roth, F.P., Vidal, M.: Towards a proteome-scale map of the human protein–protein interaction network. Nature **437**(7062), 1173–1178 (2005)

Saberi, M., Mahmassani, H., Hou, T., Zockaie, A.: Estimating network fundamental diagram using three-dimensional vehicle trajectories: extending Edie's definitions of traffic flow variables to networks. Transp. Res. Rec. **2422**, 12–20 (2014)

Schneider, C.M., Belik, V., Couronné, T., Smoreda, Z., González, M.C.: Unravelling daily human mobility motifs. J. R. Soc. Interface **10**, 1–8 (2013)

Schonfelder, A., Axhausen, K.W.: Activity spaces: measures of social exclusion? Transp. Policy **10**(4), 273–286 (2003)

Siganos, G., Faloutsos, M., Faloutsos, P., Faloutsos, C.: Power laws and the AS-level internet topology. IEEE/ACM Trans. Netw. **11**(4), 514–524 (2003)

Simini, F., González, M.C., Maritan, A., Barabási, A.L.: A universal model for mobility and migration patterns. Nature **484**, 96–100 (2012)

Song, C., Koren, T., Wang, P., Barabási, A.: Modelling the scaling properties of human mobility. Nat. Phys. **6**, 818–823 (2010)

Theriault, M., Claramunt, C., Seguin, A.M., Villeneuve, P.: Temporal GIS and statistical modelling of personal lifelines. Adv. Spat. Data Handl. (2002). doi:10.1109/DEXA.1999.795202

Thiemann, C., Theis, F., Grady, D., Brune, R., Brockmann, D.: The structure of borders in a small world. Plos One **5**(11), e15422 (2010)

Timmermans, H.J., Zhang, J.: Modeling household activity travel behavior: examples of state of the art modeling approaches and research agenda. Transp. Res. B **43**(2), 187–190 (2009)

Toole, J.L., Colak, S., Sturt, B., Alexandre, L., Evsukoff, A., González, M.C.: The path most travelled: travel demand estimation using big data resources. Transp. Res. C (2015). doi:10.1016/j.trc.2015.04.022

Virkar, Y., Clauset, A.: Power-law distributions in binned empirical data. Ann. Appl. Stat. **8**(1), 89–119 (2014)

Wang, P., Hunter, T., Bayen, A., Schechtner, K., González, M.C.: Understanding road usage patterns in urban areas. Sci. Rep. (2012). doi:10.1038/srep01001

Watling, D., Hazelton, M.L.: The dynamics and equilibria of day-to-day assignment models. Netw. Spat. Eco. **3**(3), 349–370 (2003)

Watts, D.J.: The science of a connected age. Six degrees. W. W. Norton & Co. Inc., New York (2003)

Widhalm, P., Yang, Y., Ulm, M., Athavale, S., González, M.C.: Discovering urban activity patterns in cell phone data. Transportation **42**(4), 597–623 (2015)

Woolley-Meza, O., Thiemann, C., Grady, D., Lee, J.J., Seebens, H., Blasius, B., Brockmann, D.: Complexity in human transportation networks: a comparative analysis of worldwide air transportation and global cargo ship movements. Eur. Phys. J. B **84**, 589–600 (2011)

Yagi, S., Mohammadian, A.: An activity-based microsimulation model of travel demand in the jakarta metropolitan area. J. Choice Model. **3**(1), 32–57 (2010)

Ye, Q., Wu, B., Wang, B.: Distance distribution and average shortest path length estimation in real-world networks. Adv. Data Min. Appl. **6440**, 322–333 (2010)

Yook, S.H., Jeong, H., Barabasi, A.L.: Modeling the internet's large-scale topology. Proc. Nat. Acad. Sci. USA **99**(22), 13382–13386 (2002)