Article

Characteristics of Jetters and Little Boxes: An Extensibility Study Using the Neighborhood Connectivity Survey

Xiaofan Liang [1,*], Seolha Lee [2], Hanzhou Chen [3], Benjamin de la Peña [4], and Clio Andris [1,5]

[1] School of City and Regional Planning, Georgia Institute of Technology, USA
[2] Department of Informatics, University of California – Irvine, USA
[3] Department of Geography, Pennsylvania State University, USA
[4] Shared-Use Mobility Center, USA
[5] School of Interactive Computing, Georgia Institute of Technology, USA

* Corresponding author (xiaofan.l@gatech.edu)

Abstract

Individuals connect to sets of places through travel, migration, telecommunications, and social interactions. This set of multiplex network connections comprises an individual's "extensibility," a human geography term that qualifies one's geographic reach as locally-focused or globally extensible. Here we ask: Are there clear signals of global vs. local extensibility? If so, what demographic and social life factors correlate with each type of pattern? To answer these questions, we use data from the Neighborhood Connectivity Survey conducted in Akron, Ohio, State College, Pennsylvania, and Philadelphia, Pennsylvania (global sample N = 950; in model n = 903). Based on the location of a variety of connections (travel, phone call patterns, locations of family, migration, etc.), we found that individuals fell into one of four different typologies: (a) hyper-local, (b) metropolitan, (c) mixed-many, and (d) regional-few. We tested whether individuals in each typology had different levels of local social support and different sociodemographic characteristics. We found that respondents who are white, married, and have higher educational attainment are significantly associated with more connections to a wider variety of places (more global connections), while respondents who are Black/African American, single, and with a high school level educational attainment (or lower) have more local social and spatial ties. Accordingly, the "urban poor" may be limited in their ability to interact with a variety of places (yielding a wide set of geographic experiences and influences), suggesting that wide extensibility may be a mark of privileged circumstances and heightened agency.

Keywords

community sociology; extensibility; geography; social support; social ties; spatial social networks

1. Introduction

Humans are connected to a set of places through a variety of mechanisms. These places can be childhood home cities, other places they have lived for an extensive period, locations where they have extended family, regions from where they receive information, or locales where they are members of institutions. Some individuals are connected to many places, while others are connected to few; some have distant connections, and some have nearby connections. Colloquially, we might call someone a "jetter" if they connect to a variety of places (Chen & Wellman, 2009) or perhaps living in "little boxes" if their ties and their energies are invested in local places (Wellman, 1999a). These behaviors can be encompassed under the scholarly term "extensibility," defined

as the reciprocal of time-space convergence (Adams, 1995; Janelle, 1973), the geographic spread or reach of an agent (Adams, 2009), or the geographic reach of a place or event (Kwan, 2000). It is important to study individuals' extensibility because it can tell us more about the places (i.e., communities) that may have influenced an individual and the forces that continue to shape their cultural, political, and world views. A challenge, however, is how to measure and codify extensibility so it can be used as a descriptor variable for individuals (and, in turn, for the places where these individuals reside).

Extensibility, in its simplest form, can be captured by the number of places or people one has ties to, and the geographic separation (the distance) between the ego (an individual) and alters (their contacts; see Janelle, 1973). Social science researchers have used travel diaries or surveys to capture the locations of social ties, communications, and travel patterns (Fischer, 1982; Hampton & Wellman, 2003; Stutz, 1973). These spatial and social ties are distributed differently and can interact with socio-demographic attributes. For example, when distance increases, the likelihood of forming weak or strong ties also reduces (Hipp & Perrin, 2009); kin ties can be distant in urban communities (Illenberger et al., 2011; Kowald et al., 2013), yet local in rural communities (Fischer, 1982); and social friendships tend to be more spatially compact than core ties (Boessen et al., 2018).

Research has shown that localized, tight-knit, or small networks are often associated with individual characteristics such as low income, gender (male), single status, low educational attainment, and with the African American community (Small, 2007; van Eijk, 2010). Furthermore, those with greater educational attainment and higher income tend to have connections from varied ethnic backgrounds (Marsden, 1987), perhaps suggesting a relationship with multiple geographies. While significant correlation exists, socioeconomic indicators can be weak predictors of personal network size, composition, or contact frequency (Kowald et al., 2013; van den Berg et al., 2009). Extensibility patterns also correlate with levels of social support, travel behavior, and disaster resilience (Klinenberg, 2015). Extensive social networks can provide social capital in the form of emotional or material aid (Lin et al., 2001; Wellman, 1999b) and motivate travel (Picornell et al., 2015; van den Berg et al., 2013). Conversely, individuals with less residential mobility tend to have locally concentrated contacts (Viry, 2012). Socially isolated individuals are also less likely to abandon their homes in disastrous events due to a lack of support and exposure to others' decisions (Sadri et al., 2017). These correlations may underline the role of greater social factors (e.g., racial discrimination) in including or excluding individuals in developing local or far-reaching ties (Sibley, 1995).

The emergence of big data has provided researchers with large volumes of individual behavior, that, for privacy reasons, is aggregated as place-to-place connectivity, effectively expanding the concept of extensibility to define groups of interconnected places (as described in Neal, 2012). Studies using these types of datasets have found that places with wealthier, more educated, and more resourceful populations tend to have more far-reaching ties. For instance, Facebook friendship data tells us that for a resident of Kentucky, the probability of having a Facebook friend outside 500 km is much lower than for a resident of Los Angeles (Bailey et al., 2018). Furthermore, people in counties with higher average income and education have wider, more extensive networks (Bailey et al., 2018). Relatedly, a study of British telephone calls found that wealthy locales have connections to many places, whereas poorer locales have fewer connections (Eagle et al., 2010).

There are many ways we can connect to a place: through movement, information transfer, social ties, belongingness to organizations, etc. Yet, in many survey-based studies, social network researchers often solely focus on social relationships or travel, not both. Implications are typically drawn for single variables (such as "places where kin live") rather than a collective set of places (e.g., where one has a vacation home, where they grew up, and where they make calls to), despite ample evidence that travel and relationships are intertwined (Chen & Wellman, 2009). Big data sets do not provide the full story of individual extensibility and its interactions with other social and behavioral factors because one individual is rarely found and linked between datasets. If we had such datasets, we could capture a larger swath of an individual's place-based connections and thus, use this extensibility profile as an independent or dependent variable with more confidence than if we had only one mode (locations of friends *or* cities visited). Thus, in this study, we aimed to leverage the advantages of both survey data and computational methods to characterize individuals' extensibility. We created a new dataset of ego-centric and multi-modal spatial social networks through a survey deployed in multiple US cities and characterized individuals through a data-driven machine learning model.

Our research questions are twofold. First, do individuals have common extensibility patterns (that is, does a typology emerge) that match theories of local ("little boxes"), glocal, and global reach (Wellman, 2001)? Secondly, do individuals in each type have similar demographic and behavioral attributes? To answer these questions, we clustered 903 individuals (a subsample of the 950 respondents that were suited for the model) with more than 20,000 connections into four groups (i.e., clusters, profiles, categories, classes, types). These groups are distinctive in the distances of the locations they connect with and the types of connections. Clustering was done using the K-means clustering algorithm. Then we used post-hoc tests of ANOVA and Chi-square to reveal whether these groups can be distinguished by a priori sociodemographic and behavioral factors.

We find four major types that reflect the "jetters" and "little boxes" tropes, after Chen and Wellman (2009) and Wellman (1999a), and two groups who have

characteristics of each. Our results suggest correlations between connectivity patterns and race, education, relationship status, local social support, and the security of having alternative places to stay. However, individuals within the same group do *not* have similar political orientation, age, gender, household size, or employment status. Our findings that certain demographic variables lead to more connections and more interaction with a wide variety of places can help create rules of thumb for questions such as: Which groups are more likely to travel between cities? Who may lack ties outside of communities? Who may have been exposed to different types of cultures and environments throughout their lives? Also, since individuals' connectivity data are difficult to source consistently, this study makes a conceptual advancement in data collection.

2. Data and Methods

2.1. Neighborhood Connectivity Survey

Our study uses data collected from the Neighborhood Connectivity Survey, a large mail-based survey conducted in 2017 and 2018. A mail was sent to participants selected from cities near three major locales: Akron, in the Ohio Metropolitan Area (pop. 700,000 as per the 2018 US census); State College, in the Pennsylvania Metropolitan Area, home to the large Pennsylvania State University (pop. 158,000 as per the 2018 US census); and Philadelphia County, Pennsylvania, i.e., "urban Philadelphia" (pop. 1.6 million as per the 2018 US census). These cities were chosen because they were of interest to our partners at the John S. and James L. Knight Foundation.

In 2017 and 2018, we mailed a total of 20,000 addresses and received 1,023 surveys, 950 of which were sufficiently completed. The survey includes four modules: connectivity, social life, behaviors, and demographic metrics, which, combined, took roughly 30 minutes to finish. Participants could answer the survey on paper or online and were rewarded with a gift card to nationwide retailers for their participation (see the Supplementary File for a copy of the survey).

Using data from the 2018 US census from the American Community Survey, we compared the demographics (relationship status, educational attainment, age, race/ethnicity) of our sample to those in the same set of tracts where any respondents lived. We found that our respondents have lower educational attainment rates, higher average age, and fewer people in the 18–24 range than the population in the study area. Our sample also has fewer Black and Latino members of the population than the study area.

2.2. Variables: Connections, Demography, and Behavior

We define connectivity as individuals' connections to geographic locations. To protect privacy, locations are reported at the city level (and some international links are reported as countries). We asked thirteen relational questions and grouped them into five categories: migration (i.e., where people have lived for an extensive period of time), social ties (e.g., close friends/families, communication, financial/legal supports, etc.), affiliated institutions (e.g., school, affiliated organizations), news (i.e., subscriptions to non-local news), and travel (i.e., where people have visited). These responses could be presented as a network centered at a respondent's home location and connected to geographic locations to which the individual has connections: 950 responses out of 1,023 total responses reported more than two connections and 10 out of 950 responses were missing sociodemographic information. However, we report findings for only 903 subjects because 47 subjects were not able to be effectively classified using our method (see Section 2.3).

Demographic variables include age, race, employment status, gender, relationship status, political orientation, and education level. Of the 903 respondents, 592 identified as female and 278 as male (33 reported "other" or did not disclose their gender). About 80% (n = 719) of respondents were White/Caucasian, 12.6% (n = 112) were Black/African American, and 6.8% (n = 61) were Hispanic/Latino, Asian, bi-racial, or other. Most respondents were employed (n = 523) and most described their political orientation as neutral, left, or very left. About one-third attained a bachelor's degree or higher, 48.2% were married (n = 436), and 50.3% did not have children in the home (n = 454).

Behavioral factors include a derived local social support index, intercity travel frequency, and the percentages of people who could evacuate to locations of close friends and families during emergencies. We generated a local social support index based on questions about people's social life, such as how often they have lunch with coworkers and how many friends they feel comfortable inviting to dinner (as in Stewart et al., 1988). The index scales from 0 to 1, representing low to high levels of local social support. We derived an estimate of people's intercity travel frequency based on how often they used intercity modes of transport (e.g., flights, intercity buses, etc.). Respondents also listed locations they would go to if they had to evacuate the area for two weeks, two months, and indefinitely. We then compared those locations to locations of their close friends and families to calculate the percentages of people evacuating to locations where they had close ties.

2.3. Choosing a Clustering Algorithm

We next classified individuals into different groups based on their spatial connections to find common types of extensibility profiles. In prior work, the direction, magnitude, and distance of flow patterns successfully revealed typologies of places with different compositions of social groups and spatial interactions (Andris & Hardisty, 2011; Chen et al., 2021; Liu et al., 2018; Prestby et al., 2020).

The goal is to sort each survey respondent into one of *n* number of groups that help us find common types of extensibility patterns (e.g., near, far, mixture, etc.).

We chose unsupervised learning to overcome the limitations of a priori assumptions of connectivity patterns. Machine learning techniques have been widely used to study network-based data for different purposes, such as finding a prevalent subgraph pattern (Cook & Holder, 2006), classifying or identifying different members (nodes) from a communication/social network (Alsayat & El-Sayed, 2016; Nurek & Michalski, 2020), or measuring dynamics in networks (Agarwal & Bharadwaj, 2015).

There are several advantages of using unsupervised learning in this network study. First, the algorithms allow us to input many data attributes into the classifier, and second, they suggest an optimal number of clusters (i.e., typologies/profiles) to fit our data. The unsupervised learning algorithm iterates assigning clusters to samples until the sum of the feature attribute distances between the samples in each cluster is minimized.

We tested and compared the results from three prominent algorithms: nearest-neighbor algorithms (e.g., K-means), decision tree algorithms (e.g., hierarchical clustering), and model-based clustering, in the R statistical computing environment. We ultimately chose K-means clustering for our data analysis since the algorithm resulted in an adequate number of clusters and had better within-cluster consistency (i.e., those within a single group had similar characteristics) compared to the results of other algorithms, as calculated by Silhouette scores for each cluster. The Silhouette score is a standard method to evaluate the internal consistency of K-means clusters. We calculated the 95% confidence intervals ($[-0.0176, -0.0172]$) of Silhouette scores by assigning individuals to random clusters 1,000 times. Though the clusters were moderately homogenous (Silhouette scores ranging from 0.25 to 0.32), they still provided groupings that were significantly better than random assignment. We excluded individuals with negative Silhouette scores in the K-means clustering analysis, as a negative score indicates that they were misclassified or are best classified between clusters. Thus, in this study, we used 903 responses for connectivity classification and statistical analyses.

2.4. Applying the K-Means Algorithm

We input eight variables into the K-means algorithm to characterize each individual's network. Five are the distance distributions of all places (nodes) that the individual connects with, and the other three are the total number of links that the individual reports via relationship questions, the number of unique place connections, and the number of connection types (i.e., migration, social ties, affiliated institutions, news, and travel). They represent the network structure's spatial scales, magnitudes, and diversity, respectively.

To convert the distance distribution into a vector, we divided the distribution into five distance bins: <5 km, 5–50 km, 50–1,300 km, >1,300 km, and non-US. The distance is measured as Euclidean distance, which closely approximates the travel distance (Boscoe et al., 2012). The thresholds were selected based on the observed distribution, i.e., visually distinctive troughs (5, 50 km) or natural breaks (1,300 km), and can be interpreted as connections in the neighborhood, city, and regional scale (as in Boessen et al., 2014; see Figure 1). To avoid any single feature dominating the classification process, we used the percentage of links that fall in each distance bin instead of the absolute numbers, and we used min–max scaling to transform the three other features into ranges of 0 to 1.

2.5. Statistical Tests With Chi-Square and ANOVA

To examine whether the resulting clusters have statistically distinctive demographic and behavioral characteristics, we used Chi-Square post-hoc tests for all categorical (demographic) variables and ANOVA post-hoc tests (Tukey HSD) for continuous (behavioral) variables. We calculated the standardized residuals in Chi-Square post-hoc tests for each cluster. The residuals represent the extent to which the observed counts of a demographic category in a cluster deviate from the expected counts (i.e., total counts divided by the number of clusters) normalized by the residual cell variance V (Agresti, 2018):

$$\text{Std Residuals} = \frac{\text{Observed} - \text{Expected}}{\sqrt{V}}$$

We also used Bonferroni correction for the p-values to account for the multiple comparisons. We used ANOVA post-hoc tests to compare each cluster to each other cluster for each of the variables. We use the Tukey HSD statistic to define the statistical significance of the mean differences, as it accommodates groups with unequal sample size, which is the case in our survey.

3. Results

3.1. Classification of Extensibility

The K-means clustering returned four clusters, each with a distinct feature distribution. We call the first cluster "hyperlocal" (n = 195) because most connections are concentrated within 5 km of the respondent's home location (Figure 2). These connections tend to represent social and institutional ties and contain scant non-local news subscriptions or travel outside of the local areas, indicating a close-knit local social circle, or "little boxes" (Figure 3). The 195 people in this category are marked with an overlap between the distribution of their spatial ties and local social ties. Consistent with this interpretation, the number of unique places they are connected to
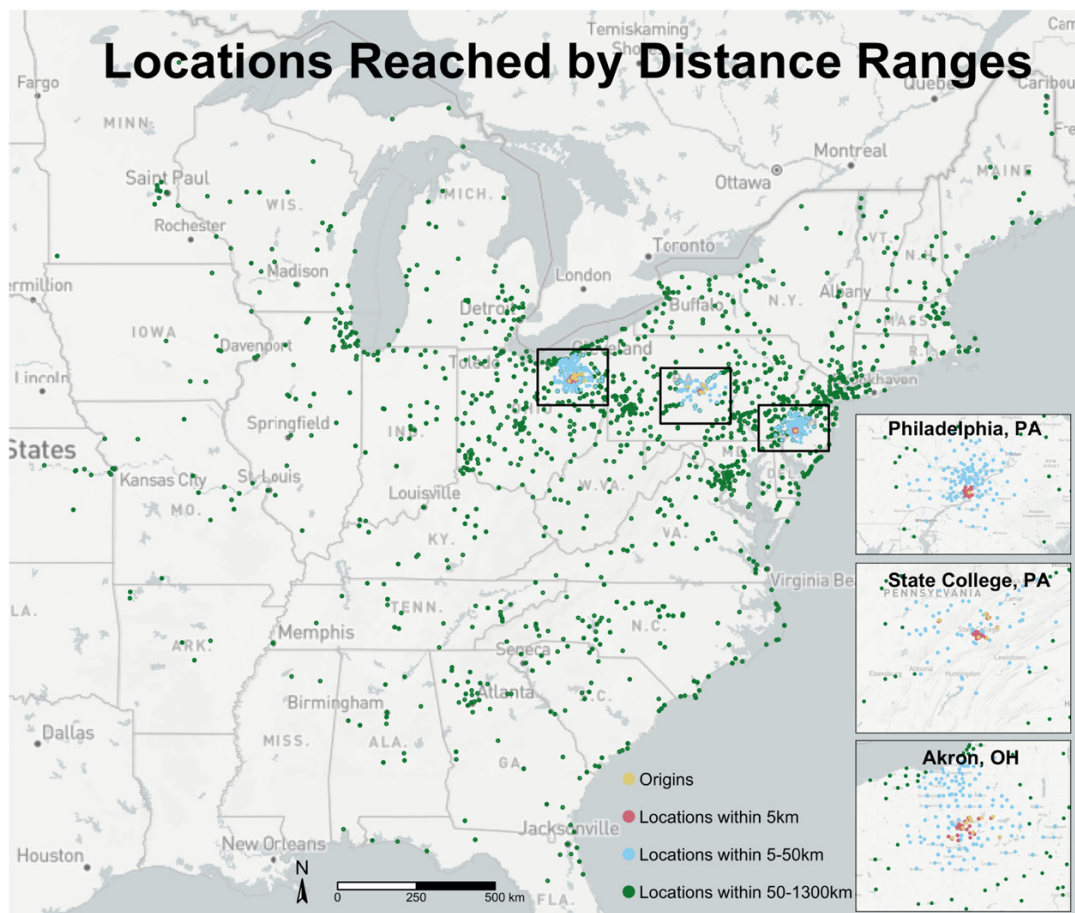
Figure 1. Destinations in different distances range from the respondents' origin cities.

is also the lowest compared with people from other clusters (Figure 2). Among all cities, Philadelphia has the highest percentage of people identified as hyperlocal (41%), which reflects Boessen et al. (2018)'s observation that people living in denser neighborhoods are more likely to have a restricted geographic reach, as well as prior findings that deprived populations have smaller social networks (see Small, 2007).

The second cluster is called "metropolitan" (n = 213), named after the concentration of links within a metropolitan area (i.e., within 50 km; see Figure 2). The distance distribution of people's migration history closely follows their social and institutional ties (Figure 3) in both the neighborhood (0–5 km) and the city (5–50 km) range. People in this cluster have many total connections and connection types, as with those in the hyperlocal cluster. Cities with the most respondents under this category are Cuyahoga Falls (51%) and Barberton (48%), two periphery cities in Akron.

The third cluster, called "mixed-many" (n = 273), has the highest average number of total connections and mixed-distance ties (Figure 2). Individuals in this cluster have local connections through institutions, while at the same time, maintain extensive social networks and spatial footprints (migration and travel; see Figure 3). The respondents in this category have the most connections to international destinations and the most

diverse ties in terms of connection types and the number of unique places. Many individuals in this category (47%) are from the university town, State College (Pennsylvania), and we expect that being affiliated with a university and academic system may encourage international ties and movement patterns.

Finally, "regional-few" (n = 222) has the fewest number of total links, most of which extend across regions (Figure 2). Respondents tend to lack local ties and have the least diverse connection types. While their institutional connections are mostly local, their spatial, social, information (news), travel histories, and networks are generally found within the (regional) range between 50–1300 km (Figure 3). The overlap may suggest that a respondent recently moved to their current city but still maintained social contacts from former places. Accordingly, State College has the highest percentage of regional-few individuals (43%), which may indicate that university affiliates have been to a distant city but are not deeply rooted in their local area.

3.2. Statistical Correlation With Sociodemographic and Behavioral Characteristics

To associate extensibility patterns with sociodemographic characteristics, we report the standardized residuals from Chi-square post-hoc tests (Table 1).
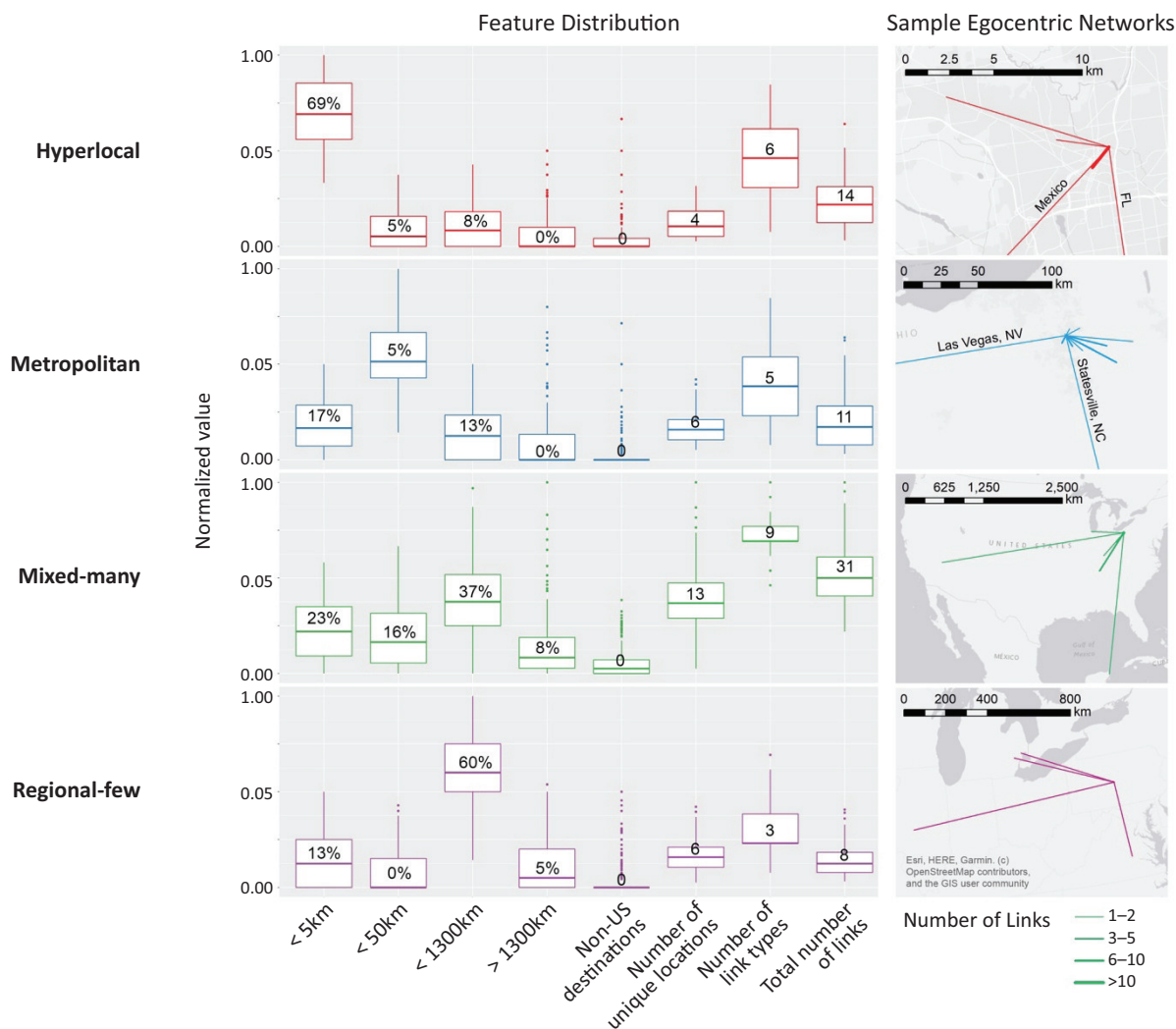
Figure 2. Boxplots and sampled ego-centric networks from each of the four resulting clusters. Notes: The boxplot shows the descriptive statistics of each feature in each cluster; for non-distance features, the boxplots report the distribution of absolute counts; for distance-related features, the boxplots report the distribution of the percentage of links in the distance ranges rather than absolute counts. The labeled statistics on the boxplots are the unscaled median values for each feature; the y values are then scaled between 0 and 1 for relative comparison. The sampled networks on the right provide examples of scale, wherein the edges are weighted by the number of links connected to each location.

We found that respondents with a high school education level or lower are statistically more likely to have locally concentrated ties, as featured by the hyperlocal and metropolitan patterns. Conversely, respondents with a Bachelor's degree or higher are more likely to have a mixed-many network pattern. We also observed that pursuing an associate's degree is correlated with a spatial social network that expands beyond one's local context. We postulated that education beyond high school may have a significant impact on people that meet others from distant places or visit places outside of their hometown areas.

The metropolitan and mixed-many clusters were comprised of many white individuals, while Black or African American individuals were often found in the hyperlocal category (with a residual of 6.41). Black or African American respondents were more likely than

white respondents (32% vs. 11%) to be in the hyperlocal category. Forty-six percent (n = 51) of Black or African American respondents were classified as hyperlocal, which exceeds the expected 25% if the population was evenly split across four patterns. In addition, race and education levels were correlated: 76% (n = 39) of Black or African American respondents in the hyperlocal category also had educational attainment at the high school level or lower. This group may also have close-knit relationships at the neighborhood level.

Respondents who identified as single seemed to concentrate in the hyperlocal cluster, but this effect may be explained by education levels. Most single people in the hyperlocal cluster have an education level of high school or lower. In contrast, people who are married tend to have a mixed-many type of connectivity pattern. Three percent of married mixed-many individuals are Black or
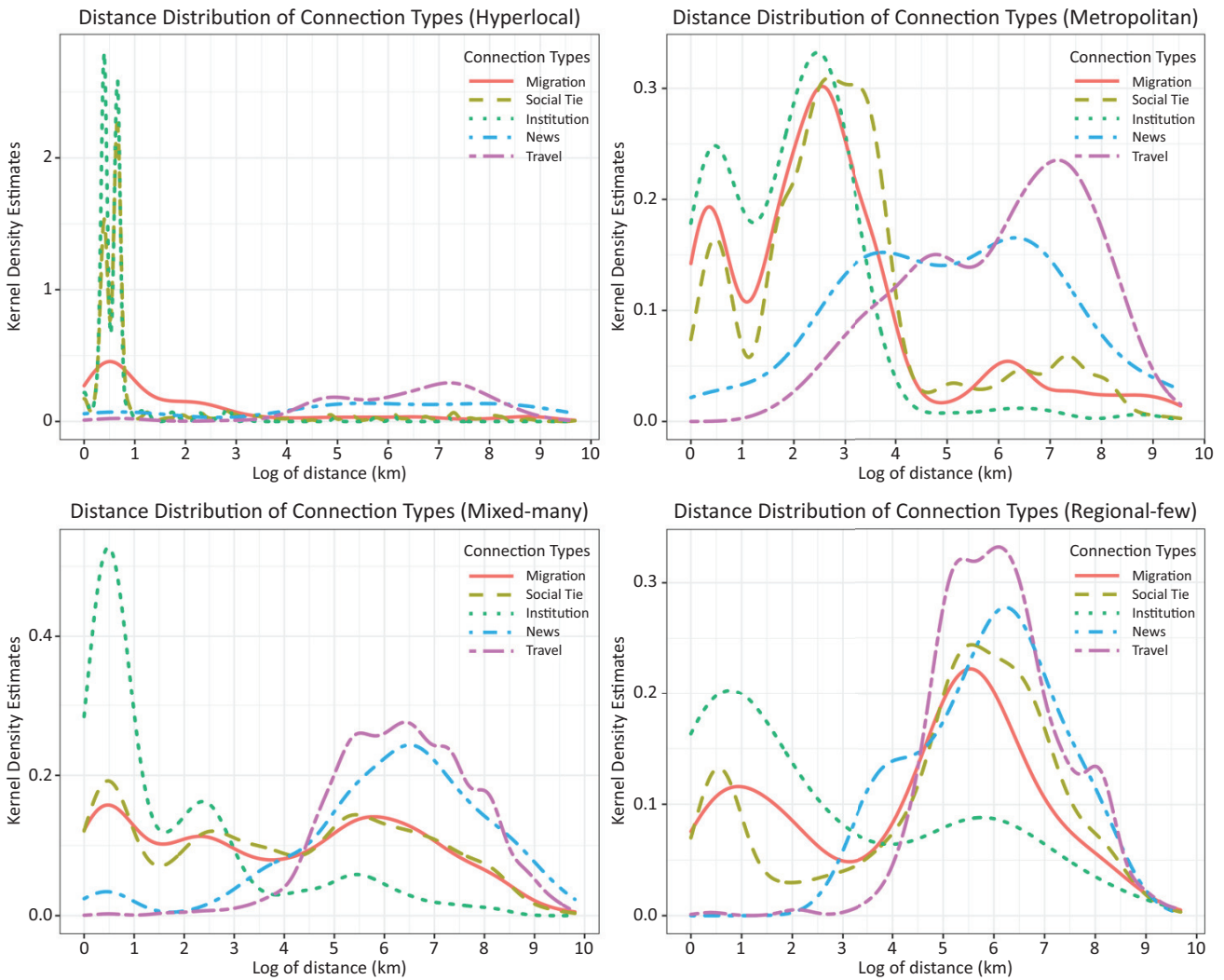
Figure 3. Distance distribution of various connection types for each of the four clusters. Notes: The Y-axis shows the kernel density estimates of the count of a type of connection at various distances; higher y values indicate a greater likelihood/frequency of occurrence.

African American, which is significantly lower than the overall percentage (12%) in the total respondent population. Black or African American respondents who are single and have lower levels of educational attainment are often found in the hyperlocal pattern.

In terms of behavioral characteristics, the ANOVA post-hoc tests report statistically-significant mean differences between two clusters (Figure 4). Respondents with more long-distance connections (in the mixed-many and regional-few categories) travel more often between cities. This correlation is reasonable because connections provide motivations for (and evidence of) past travel, perhaps visiting family or attending alumni events.

People with hyperlocal and metropolitan styles of extensibility also reported less local social support than people in the mixed-many group, despite the former having a high concentration of local ties. Since the local social support index only measures the quality of social life locally, the result indicates that people in the

mixed-many group are more likely to receive social support from their local networks than people in hyperlocal and metropolitan clusters, even if they share a similar number of total connections.

Lastly, we tested whether people with different extensibility patterns have more or few options regarding alternative places to stay (which is especially useful in emergencies). Eighty-four percent of mixed-many respondents identified plausible evacuation locations, while only 45%, 43%, and 27% of people in hyperlocal, metropolitan, and regional-few groups, respectively, described destination cities for evacuations. The hyperlocal group had the fewest percentage of people (36%) that said they would evacuate to locations where they also had friends and families (inferred), perhaps because their ties are nearby (and likely to be impacted by the same evacuation events due to co-location). Still, many respondents in the mixed-many cluster appeared to be at an advantage, as they could supply more scenarios with support during evacuation events.

Table 1. Standardized residuals from Chi-square post-hoc tests.

Sociodemographic Variables	Hyperlocal	Metropolitan	Mixed-many	Regional-few	Count
Age: 18–24	1.02	−1.25	−0.65	0.95	39
Age: 25–34	1.05	−0.70	−0.77	0.50	162
Age: 35–54	−0.65	0.80	−0.93	0.83	267
Age: 54–65	0.45	0.85	−0.85	−0.36	184
Age: 65+	−1.12	−0.41	2.68	−1.39	246
Employment: Unemployed	1.96	0.78	−2.34	−0.12	46
Employment: Retired or Disabled	1.15	1.55	−0.52	−2.06	268
Employment: Student	0.31	−2.00	0.72	0.91	39
Employment: Employed	−2.10	−0.97	1.25	1.60	523
Gender: Female	0.08	−0.15	2.55	−2.63	592
Gender: Male	−0.08	0.15	−2.54	2.63	278
Education: High school or less	**6.41*****	**4.04****	**−6.41*****	**−3.12***	376
Education: Associate	**−3.76****	0.21	2.22	0.95	247
Education: Bachelor	−1.76	**−3.40***	**3.74****	0.98	165
Education: Master or above	−2.74	−2.65	2.45	2.54	79
Political Orient: Very right	−0.88	−0.57	−0.14	1.50	50
Political Orient: Moderate right	−1.00	0.53	−0.01	0.39	140
Political Orient: Neutral	1.33	1.57	−1.70	−0.88	222
Political Orient: Moderate left	−0.66	−0.78	1.93	−0.75	215
Political Orient: Very left	0.77	−1.12	−0.15	0.57	133
Race: White or Caucasian	**−7.06*****	**3.37****	**4.22*****	−1.13	719
Race: Black of African American	**6.70*****	**−3.21***	**−3.79****	0.85	112
Race: Other	2.27	−1.07	−1.63	0.65	61
Relationship: Single	**4.37*****	−0.01	**−4.72*****	0.89	177
Relationship: In a relationship	0.13	1.81	−0.65	−1.21	115
Relationship: Married	**−3.73****	−1.81	**3.53****	1.56	436
Relationship: Divorced or separated	−0.70	0.63	0.32	−0.30	107
Relationship: Widowed	1.22	0.40	0.90	−2.51	62
Children below 18: Yes	1.49	0.18	1.67	0.24	204
Children below 18: No	−1.49	−0.18	−1.67	−0.24	454

Notes: $*p < 0.05$; $**p < 0.01$. $***p < 0.001$; p-values are adjusted by Bonferroni correction; the standardized residuals should be compared within the sociodemographic subtypes (e.g., gender, race, education) for a particular cluster; a statistically-significant standardized residual means that a sociodemographic attribute is highly concentrated in a cluster beyond the expected mean (see Section 2.5); the count is the number of people in a sociodemographic subtype.

Age, employment, gender, children status, and political orientation variables are relatively well-distributed across the clusters and thus do not exhibit a significant correlation with one or more patterns.

4. Discussion and Conclusions

This study created a typology of individual connectivity patterns (including hyperlocal, metropolitan, mixed-many, and regional-few) through an extensive mail-based survey called the Neighborhood Connectivity Survey. The survey provided a unique dataset that included a wide range of spatial social connections of individuals and socio-demographic information. We conducted unsupervised clustering of the individual spatial social networks using the K-means algorithm to characterize the individual connectivity with multiple features. Lastly, we examined the tendencies in sociodemographic characteristics, social life, and spatial activities of individuals with each connectivity pattern through ANOVA and Chi-square tests.

We found that the four typologies have distinct extensibility profiles and are only moderately homogenous, indicating that individuals can deviate from the typologies or have mixed profiles of extensibilities. We also found that race, education, and relationship status
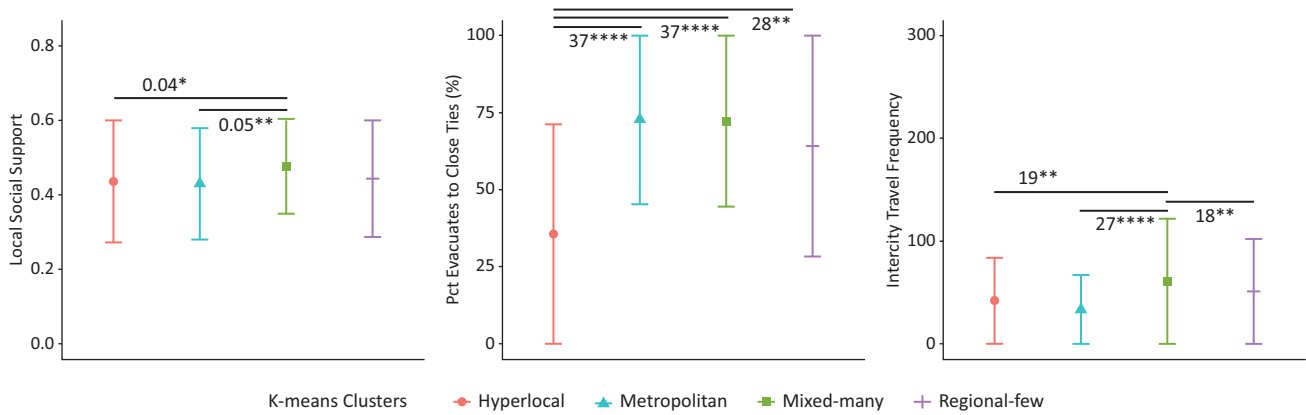
Figure 4. Mean and standard deviations of local social support, percentage of people that could leave to the locations of close friends and families, and intercity travel frequency for each extensibility type. Notes: The y axis scales with the minimum and maximum value in a behavioral factor; a black line between any two clusters signals a statistically-significant relationship and is annotated with the absolute value of the mean difference between those two clusters at either end of the black line; the statistical significance of the mean difference is tested with ANOVA multiple comparisons (Tukey HSD); the p-value is adjusted by Bonferroni correction ($*p < 0.05$; $**p < 0.01$; $***p < 0.001$; $****p < 0.0001$).

correlate with individuals' spatial social network patterns, while age, gender, family size, employment status, and political orientation did not show a significant correlation with the clusters. A notable finding is that residents with low education attainment and residents from Black or African American populations had the smallest networks (by area). This finding triangulates with past research showing that Black individuals tend to have smaller and weaker social networks and maintain fewer social ties outside of families than white individuals (Small, 2007). It also reflects prior findings that education beyond secondary schools is statistically associated with network heterogeneity and levels of resources leverageable from the social networks (van Eijk, 2010). Yet, our findings further reveal that these local ties are likely to be social and institution connections (thus with limited lived experience, news, and travel outside of the home city) and that individuals in this group are least likely to evacuate to locations of closest friends and families during disastrous events. It is also important to note that despite a high concentration of local ties, they may not have the highest level of local social support. Accordingly, more attention and resources should be allocated to this community in terms of community facilities and emergency preparedness.

Our results also speak to the privilege of mixed-range and diverse network patterns. mixed-many and regional-few individuals tended to be White, married, or college-educated, and exhibited frequent travel tendencies between cities, local social support, and resilience during disaster events. Similarly, Viry (2012) found that people's social support (i.e., the number of supporting ties) is not affected by the geographic distribution of their networks and the frequency of moving, though those who move frequently lean toward a sparsely knit and transitive social network.

These results serve as evidence that systemic deprivation and exclusion in terms of race and especially access to education tends to result in a geographically-limited range of social contacts and experiences. While our results associate traveling and having experiences in many places with higher socioeconomic status, we acknowledge that migration can also be forced, as in the case of population displacement during crises. However, a more novel perspective is that these patterns tend to be consistent regardless of the respondent's home location, and urban or rural distinctions. Therefore, we suggest that inferring peoples' experiences given the traditional context of the geographic situation (i.e., hometown location) should also consider the influence of inclusive or exclusive social factors (as in Sibley, 1995).

Finally, this study has a number of limitations. First, the sample population was limited to residents in a few cities in neighboring states in the US. Accordingly, the distance distribution was reasonably consistent across the sample population. Due to the limited sample size, we also did not examine the implications of these cities' characteristics on extensibility patterns, which has been explored in other studies (Boessen et al., 2014, 2018; Mazumdar et al., 2018). Given the differences in our sample characteristics and the population characteristics of our study area, as mentioned, our results may be skewed to represent older people who have less educational attainment and are white. Next, variances persisted between individuals *within* each cluster. Using the mean values of the features for clustering removed important parts of the data distribution (such as anomalies or bimodal trends). Lastly, we lacked a detailed explanatory mechanism for the clusters. Unsupervised clustering captures intrinsic tendencies but does not explain why variables within one group may correlate. Future work should examine direct correlation with fewer variables

from our survey data to provide a more in-depth understanding of how different connections are associated with demographic or lifestyle factors. We suggest that this type of extensibility-driven work be replicated across a wider range of geographic areas to capture communities that differ in terms of density, isolation, etc., and to capture respondents from a wider variety of sociodemographic groups.

Acknowledgments

Conflict of Interests

The authors declare no conflict of interests.

Supplementary Material

Supplementary material for this article is available online in the format provided by the author (unedited).

References

Adams, P. C. (1995). A reconsideration of personal boundaries in space-time. *Annals of the Association of American Geographers*, *85*(2), 267–285.

Adams, P. C. (2009). *Geographies of media and communication*. John Wiley & Sons.

Agarwal, V., & Bharadwaj, K. K. (2015). Predicting the dynamics of social circles in ego networks using pattern analysis and GA K-means clustering. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *5*(3), 113–141.

Agresti, A. (2018). *An introduction to categorical data analysis*. John Wiley & Sons.

Alsayat, A., & El-Sayed, H. (2016). *Social media analysis using optimized K-means clustering* [Paper presentation]. 2016 IEEE/ACIS 14th International Conference on software engineering research, management and applications, Towson, USA.

Andris, C., & Hardisty, F. (2011). Visualizing migration dynamics using weighted radial variation [Paper presentation]. GeoViz Hamburg '11: Linking Geovisualization With Spatial Analysis and Modeling Conference, Hamburg, Germany.

Bailey, M., Cao, R., Kuchler, T., Stroebel, J., & Wong, A. (2018). Social connectedness: Measurement, determinants, and effects. *Journal of Economic Perspectives*, *32*(3), 259–280.

Boessen, A., Hipp, J. R., Butts, C. T., Nagle, N. N., & Smith, E. J. (2018). The built environment, spatial scale, and social networks: Do land uses matter for personal network structure? *Environment and Planning B: Urban Analytics and City Science*, *45*(3), 400–416.

Boessen, A., Hipp, J. R., Smith, E. J., Butts, C. T., Nagle, N. N., & Almquist, Z. (2014). Networks, space, and residents' perception of cohesion. *American Journal of Community Psychology*, *53*(3/4), 447–461.

Boscoe, F. P., Henry, K. A., & Zdeb, M. S. (2012). A nationwide comparison of driving distance versus straightline distance to hospitals. *The Professional Geographer*, *64*(2), 188–196.

Chen, Q., Chuang, I. T., & Poorthuis, A. (2021). Entangled footprints: Understanding urban neighbourhoods by measuring distance, diversity, and direction of flows in Singapore. *Computers, Environment and Urban Systems*, *90*. https://doi.org/10.1016/j.compenvurbsys.2021.101708

Chen, W., & Wellman, B. (2009). Net and jet: The internet use, travel and social networks of Chinese Canadian entrepreneurs. *Information, Communication & Society*, *12*(4), 525–547.

Cook, D. J., & Holder, L. B. (Eds.). (2006). *Mining graph data*. John Wiley & Sons.

Eagle, N., Macy, M., & Claxton, R. (2010). Network diversity and economic development. *Science*, *328*(5981), 1029–1031.

Fischer, C. S. (1982). *To dwell among friends: Personal networks in town and city*. University of Chicago Press.

Hampton, K., & Wellman, B. (2003). Neighboring in Netville: How the Internet supports community and social capital in a wired suburb. *City & Community*, *2*(4), 277–311.

Hipp, J. R., & Perrin, A. J. (2009). The simultaneous effect of social distance and physical distance on the formation of neighborhood ties. *City & Community*, *8*(1), 5–25.

Illenberger, J., Kowald, M., Axhausen, K. W., & Nagel, K. (2011). Insights into a spatially embedded social network from a large-scale snowball sample. *The European Physical Journal B*, *84*(4), 549–561.

Janelle, D. G. (1973). Measuring human extensibility in a shrinking world. *Journal of Geography*, *72*(5), 8–15.

Klinenberg, E. (2015). *Heat wave: A social autopsy of disaster in Chicago*. University of Chicago Press.

Kowald, M., van den Berg, P., Frei, A., Carrasco, J. A., Arentze, T., Axhausen, K., Mok, D., Timmermans, H., & Wellman, B. (2013). Distance patterns of personal networks in four countries: A comparative study. *Journal of Transport Geography*, *31*, 236–248.

Kwan, M. P. (2000). Human extensibility and individual hybrid-accessibility in space-time: A multi-scale representation using GIS. In D. G. Janelle & D. C. Hodge (Eds.), *Information, place, and cyberspace* (pp. 241–256). Springer.

Lin, N., Cook, K. S., & Burt, R. S. (Eds.). (2001). *Social capital: Theory and research*. Transaction Publishers.

Liu, X., Hollister, R., & Andris, C. (2018). Wealthy hubs and poor chains: Constellations in the US urban migration system. In L. Perez, E.-K. Kim, & R. Sengupta (Eds.), *Agent-based models and complexity science in the age of geospatial big data* (pp. 73–86). Springer.

Marsden, P. V. (1987). Core discussion networks of Americans. *American Sociological Review*, *52*(1), 122–131.

Mazumdar, S., Learnihan, V., Cochrane, T., & Davey, R. (2018). The built environment and social capital: A systematic review. *Environment and Behavior*, *50*(2), 119–158.

Neal, Z. P. (2012). *The connected city: How networks are shaping the modern metropolis*. Routledge.

Nurek, M., & Michalski, R. (2020). Combining machine learning and social network analysis to reveal the organizational structures. *Applied Sciences*, *10*(5). https://doi.org/10.48550/arXiv.1906.09576

Picornell, M., Ruiz, T., Lenormand, M., Ramasco, J. J., Dubernet, T., & Frías-Martínez, E. (2015). Exploring the potential of phone call data to characterize the relationship between social network and travel behavior. *Transportation*, *42*(4), 647–668.

Prestby, T., App, J., Kang, Y., & Gao, S. (2020). Understanding neighborhood isolation through spatial interaction network analysis using location big data. *Environment and Planning A: Economy and Space*, *52*(6), 1027–1031.

Sadri, A. M., Ukkusuri, S. V., & Gladwin, H. (2017). The role of social networks and information sources on hurricane evacuation decision making. *Natural Hazards Review*, *18*(3). https://ascelibrary.org/doi/10.1061/%28ASCE%29NH.1527-6996.0000244

Sibley, D. (1995). *Geographies of exclusion: Society and difference in the west.* Routledge.

Small, M. L. (2007). Racial differences in networks: Do neighborhood conditions matter? *Social Science Quarterly*, *88*(2), 320–343.

Stewart, A. L., Hays, R. D., & Ware, J. E. (1988). The MOS short-form general health survey: Reliability and validity in a patient population. *Medical Care*, *26*, 724–735.

Stutz, F. P. (1973). Distance and network effects on urban social travel fields. *Economic Geography*, *49*(2), 134–144.

van den Berg, P., Arentze, T., & Timmermans, H. (2009). Size and composition of ego-centered social networks and their effect on geographic distance and contact frequency. *Transportation Research Record*, *2135*(1), 1–9.

van den Berg, P., Arentze, T., & Timmermans, H. (2013). A path analysis of social networks, telecommunication and social activity–travel patterns. *Transportation Research Part C: Emerging Technologies*, *26*, 256–268.

van Eijk, G. (2010). *Unequal networks: Spatial segregation, relationships and inequality in the city*. IOS Press.

Viry, G. (2012). Residential mobility and the spatial dispersion of personal networks: Effects on social support. *Social Networks*, *34*(1), 59–72.

Wellman, B. (1999a). From little boxes to loosely bounded networks: The privatization and domestication of communities. In J. L. Abu-Lughod (Ed.), *Sociology for the twenty-first century. Continuities and cutting edge*. University of Chicago Press.

Wellman, B. (1999b). *Networks in the global village: Life in contemporary communities*. Routledge.

Wellman, B. (2001). Little boxes, glocalization, and networked individualism. In M. Tanabe, P. Besselaar, & T. Ishida (Eds.), *Digital cities II: Computational and sociological approaches* (pp. 10–25). Springer.

About the Authors

Xiaofan Liang is a city and regional planning PhD student at Georgia Institute of Technology. Her research interests are in spatial social networks, urban analytics, and participatory and engagement methods for spatial planning. She is passionate about leveraging technology to support community and public initiatives in planning. Xiaofan obtained a BA in sociology from the University of California – Berkeley and a BS in computational science from Minerva University.

Seolha Lee is a PhD student in informatics at the University of California – Irvine. Going across urban data analytics and civic activism, her research questions how to use data and information technology to improve social justice in cities using mixed methods.

Hanzhou Chen graduated from Penn State with an MS in geography. His research interests include human mobility, social flow, visual analytics, and urban planning. He is currently working at a start-up company, where he creates map and GIS products that use foot traffic data and location intelligence to help businesses and municipalities gain insight into human mobility and travel patterns.

Benjamin de la Peña is the CEO of the Shared-Use Mobility Center, a public-interest organization dedicated to achieving equitable, affordable, and environmentally sound mobility across the US. He also chairs the Global Partnership for Informal Transportation and is a senior fellow on Mobility for the Canadian Urban Institute. He has a Master's in urban planning from Harvard University's Graduate School of Design and a BA in communication, with a major in journalism, from the University of the Philippines.

Clio Andris is an associate professor with a joint appointment in the School of City & Regional Planning and the School of Interactive Computing at Georgia Institute of Technology. She directs the Friendly Cities Lab and her research is in geographic information science, social networks, urban planning, and information visualization.