Social Science & Medicine 161 (2016) 134–142

Contents lists available at ScienceDirect

Social Science & Medicine

journal homepage: www.elsevier.com/locate/socscimed

Using global positioning systems to study health-related mobility and participation

Eugene Brusilovskiy*, Louis A. Klein, Mark S. Salzer

Department of Rehabilitation Sciences, Temple University, Philadelphia, PA, USA

ARTICLE INFO

Article history:
Received 31 December 2015
Received in revised form
4 May 2016
Accepted 1 June 2016
Available online 2 June 2016

Keywords:
USA
GPS
Global positioning systems
Mobility
Community participation
Spatiotemporal data mining

ABSTRACT

Community participation, as indicated by mobility and engagement in socially meaningful activities, is a central component of health based on the International Classification of Health, Functioning, and Disease (WHO, 2001). Global positioning systems (GPS) technology is emerging as a tool for tracking mobility and participation in health and disability-related research. This paper fills a gap in the literature and provides a thorough description of a method that can be used to generate a number of different variables related to the constructs of mobility and participation from GPS data. Here, these variables are generated with the help of ST-DBSCAN, a spatiotemporal data mining algorithm. The variables include the number of unique destinations, activity space area, distance traveled, time in transit, and time at destinations. Data obtained from five individuals with psychiatric disabilities who carried GPS-enabled cell phones for two weeks are presented. Within- and across- individual variability on these constructs was observed. Given the feasibility of gathering data with GPS, larger scale studies of mobility and participation employing this method are warranted.

1. Introduction

The International Classification of Health, Functioning, and Disease (ICF) (WHO, 2001) offers a groundbreaking and unique framework for understanding health and functioning. According to the ICF, health and functioning consists of three distinct components: body function and structure, which refer to bodily impairment or symptoms; activities, which refer to the ability to perform a specific task or action; and participation, which is “involvement in a life situation” as indicated by mobility and natural engagements, typically with others, in one of four social life domains: (a) domestic life; (b) interpersonal life (including formal relationships as well as informal social relationships, family relationships, and intimate relationships); (c) major life activities consisting of education (informal, vocational training, and higher education) and employment (remunerative and nonremunerative, excluding domestic work); and (d) community, civic, and social life, including religion, politics, recreation and leisure, hobbies, socializing, sports, arts, and culture. Participation is highly desired by people with health impairments and disabilities, their family members, and society (Whiteneck, 2006), and the ICF recognizes that participation can enhance physical health and wellness, and activities of daily living. Data about participation is needed for assessing the efficacy and cost-effectiveness of treatment and rehabilitation interventions (Ustun et al., 2003). To date, participation has primarily been measured using self-report, including recall approaches (Salzer et al., 2014) and time use logs (Eklund et al., 2009; Yanos and Robilotta, 2011). While such approaches have been found to be reliable and valid (Salzer et al., 2014, 2015), additional approaches could be valuable, especially if they expand the types of mobility and participation constructs that can be assessed.

Global positioning systems (GPS) technology offers one promising approach for expanding measurement of health through a greater understanding of mobility and participation. Developed by the US Department of Defense as a navigation tool with both military and civil applications, GPS relies on a network of at least 24 satellites orbiting the Earth, from which signal receivers on the ground determine the position (Trimble, 2004). And while GPS technology is not new, decreased costs and increased accessibility have made it a promising method of data collection, and a growing body of research from around the world demonstrates the potential utility of GPS technology. A number of recent studies using GPS have examined various participation-related constructs.

Many studies have used GPS to examine factors associated with

* Corresponding author. 1700 N. Broad St., Suite 313C, Philadelphia, PA 19121, USA.
E-mail address: eugeneby@temple.edu (E. Brusilovskiy).

http://dx.doi.org/10.1016/j.socscimed.2016.06.001
0277-9536/

physical activity. One study combined the use of GPS with accelerometers and found that 40% of light, medium and vigorous physical activity undertaken outdoors took place within 800 m of the individuals' home (Hillsdon et al., 2015). Another study (Yin et al., 2013) found that youth tend to be physically active within a quarter-mile radius around their homes. Collins et al. (2015) used GPS with adolescents to determine how they got around in the community and found that those who walked or biked had higher levels of moderate-to-vigorous physical activity than those who traveled by car or bus during the commute to school, but not during the afterschool hours. Clark et al. (2014) tracked the outdoor activity of individuals with severe traumatic brain injury compared to a non-impaired group. They used customized software that evaluated the location, velocity of movement and number of satellites with fixes to differentiate between time spent indoors, time doing outdoor physical activity and time engaged in non-physical activity-related transportation, and found that the traumatic brain injury and control groups did not differ significantly in terms of the amount of time spent doing outdoor physical activity.

Other studies have examined mobility-related questions. Wettstein et al. (2013) used GPS to track 100 cognitively healthy older adults. They measured several variables related to mobility: identified nodes (destinations), maximum and mean distance traveled from home, walking speed, number of walking tracks per day, walking duration per walking track, and walking distance per walking track. One key finding was that walking-based mobility could be predicted by physical functioning, but not by demographics. Wettstein et al. (2012) conducted a second study using GPS with 257 older adults to assess walking distance, duration, and speed; time spent out of home per day; and the number of destinations per day. They found that individuals without cognitive impairment had a significantly higher number of destinations per day than individuals with Alzheimer's disease. Bostelmann et al. (2015) used GPS-enabled phones to measure total distance walked, average walking speed, and the total walking duration per day of four neurosurgical spine patients in Germany a week before and three months after they had surgery. They reported increases in these areas post-surgery in three out of the four patients.

GPS combined with other technologies have been particularly useful for tracking mobility, activity and participation of individuals with physical impairments who use wheelchairs. Harris et al. (2010) developed a Participation and Activity Measurement System (PAMS) to assess levels of activity and participation of five individuals who used electric wheelchairs. The PAMS employed GPS devices to track each individual's activity outside, and a wheel revolution counter monitored the extent of inside activity. A study by Lankton et al. (2005) presents results from the Wheelchair Activity Monitoring Instrument, which employs GPS, odometers and seat occupancy sensors to examine in detail the daily activities of three individuals who use wheelchairs. Jayaraman et al. (2014) used a GPS and a step activity monitor to examine community mobility and social interaction (variables included the number of steps taken in and out of home, wheelchair use, driving trips, and amount of time spent on social and community trips) of an individual who used both a wheelchair and a prosthesis for mobility. The authors also attempted to identify the participant's destinations with the use of Google Maps and ArcGIS. Hordacre et al. (2015) used GPS to identify destinations, which they categorized as employment, residential, commercial, health service, recreational, social and "other" with the help of the QTravel software, for 46 individuals with transtibial amputations who fall frequently or infrequently. They reported that those who fall frequently had significantly lower participation in recreational roles. Finally, as mobility along a steep slope or uneven pavement is challenging for wheelchair users and individuals with vision impairments, another study has attempted to calculate sidewalk slope and identify the optimal route in real-time by combining Google Street View image data and GPS trajectory data (Lu and Karimi, 2015).

While GPS is being increasingly used to address mobility and participation-related questions, the current literature is lacking. Current studies have examined only a few constructs that can be derived using GPS data. For example, simply assessing whether the person is out of the home or within a certain distance out of the home only minimally utilizes GPS technology, and data appear to be examined manually rather than using any advanced data mining algorithms. Current published studies also provide limited detail on their analytical approaches, which may be problematic, as GPS data are numerous and complex. For example, data collection at one-minute intervals could yield 1440 time-stamped data points with latitude and longitude values over the course of a 24-h period, and using these data in a meaningful way is very complicated. However, there are often inaccuracies and errors associated with GPS data, due to factors including, but not limited to, atmospheric conditions, satellite and receiver errors, and multipath errors (i.e., the GPS signal reflecting off tall objects before it reaches the receiver). Certain types of error may be reduced by post-processing GPS data, however these corrections aren't helpful with receiver or multipath errors, with the latter being especially likely in urban environments (Trimble, 2004). Error also depends on the GPS device that is being used. For instance, a recent study showed that positions obtained by GPS-enabled mobile phones had a median error ranging from 5 to 8.5 m, whereas the error for stand-alone GPS units was substantially smaller, ranging between approximately 2–3 m (Zandbergen and Barbeau, 2011). Finally, there has been little to no discussion about how missing data are addressed. As mentioned above, GPS data collection is sensitive to weather (e.g., Hillier, 2008) and satellite access, which can be blocked when indoors (e.g., Harris et al., 2010), and instead of obtaining an incorrect location at certain time points, data may be absent altogether. Human factors, such as not keeping the tracking device charged, can also create missing data. Therefore, missing data are common in such data collection approaches, and require attention.

This paper offers a broad explication of mobility and participation constructs that can be examined using GPS data, followed by a description of how variables measuring each construct can be calculated. The science of using GPS can be furthered by identifying detailed methodologies that can be commonly used by researchers. Examples of results using these methods with a sample of individuals with psychiatric disabilities are then presented. Other than providing a sense of findings that can be generated with these methods, variability within and across individuals on the various constructs is discussed, which further underscores the likely utility of the constructs and proposed methods.

2. Methods

2.1. Procedures

2.1.1. Participants

Participants were recruited from a community mental health center (CHMC) in the suburbs of Philadelphia, PA, which serves individuals with public insurance (e.g., Medicaid). The following inclusion criteria were used: 1) Pennsylvania residents; 2) 18 years of age or older; 3) able to speak and understand English; 4) an axis I psychotic disorder (DSM 295.XX) or affective disorders such as major depression or bipolar disorder (DSM 296.XX) using agency records (co-occurring substance use disorders were included); 5) currently receiving outpatient services; 6) with a stable residence for at least the past six months with no plans to move in the next month; and 7) able to provide informed consent as assessed by research staff using procedures similar to those discussed by Carpenter et al. (2000). The following individuals were excluded from this study: 1) those with a legal guardian; 2) those with identified co-occurring neurological impairment, intellectual disability, or significant communication-related disorders that would likely affect their ability to provide informed consent or participate in the data collection procedures. The study obtained Institutional Review Board approval from the authors' university.

Recruitment took place between February and April of 2013, and a total of 5 participants enrolled. Each participant received a cellular phone which had GPS tracking software installed on it and was instructed by research staff on the basic functions of the phone. Participants received phone calls every several days reminding them to carry and charge the phone. A follow up face-to-face interview was scheduled 18–20 days after the initial interview. Participants received $15 for their enrollment interview and a $65 honorarium at the second interview when they returned the cell phone.

2.1.2. GPS tracking

AccuTracking (www.accutracking.com) software was used for tracking the locations and movement of study participants over the two weeks that they had the phones. The software is compatible with most modern cellular phones, and does real time tracking of the phones' locations, which are stored as latitude and longitude coordinates, at one minute intervals. AccuTracking has been used by individuals, government agencies, businesses, and researchers in order to track the location of family members, vehicles, employees, or subjects in real time. The software has been used in several published studies (Zhou et al., 2004, 2005). No action was required from the individual being tracked other than keeping the phone on. AccuTracking automatically started whenever the phone was turned on, and ran in the background. However, if the phone was turned off, tracking would stop. Real time transfer of location data to the AccuTracking online database occurred whenever cellular coverage was available, assuming that the GPS device was outdoors; at times when there was no coverage, GPS data were cached and transferred at a later time.

Because location was recorded once every minute for the two week duration of the study, the total number of points that could be obtained was 1 record/min $\times$ 1440 min in a 24-h period $\times$ 14 24-h periods $=$ 20,160 records. As tracking only occurred when the phone was outside or close to a window, it was expected that the number of records obtained would be less than 20,160. GPS data collection could be monitored by logging into the AccuTracking website, and at the end of the tracking period, each participant's data was downloaded as a .csv file.

2.2. Data analysis

A number of mobility and participation-related constructs and variables can be examined using GPS data. These include community participation, geographic scope of mobility, and temporal scope of mobility, and are described in detail below.

2.2.1. Community participation

Total Number of Destinations. Destinations, such as home, a store, a gym, or a park, are places where community participation takes place. As previously mentioned, GPS is sensitive to short-distance movement, and measured locations will not be the same if the individual is moving more than a few feet in either direction, which might happen even once he has reached his destination (e.g., walking along a plaza, playing a sport outside). This, coupled with the satellite, receiver, multipath and other errors that might occur with data collection, makes the identification of destinations from GPS data challenging.

In the past several years, a number of data mining tools which take into account the spatial and temporal proximity of GPS points to identify spatiotemporal clusters, or destinations, have been developed. These include Time-Clustering-based Behavior Analysis (TCBA) (Ji et al., 2012), Spatial-Temporal GRID (ST-GRID) (Wang et al., 2006), and Tree-Based Hierarchical Graphing (TBHG) (Khetarpaul et al., 2011). The data mining tool used here is a point density-based algorithm called Spatio-Temporal Density-Based Spatial Clustering of Applications with Noise (ST-DBSCAN) (Birant and Kut, 2007).

ST-DBSCAN has been used in numerous studies with different foci (e.g., Yue et al., 2012; Rogalsky, 2010). It has several advantages over other spatiotemporal clustering algorithms: for instance, it is a more effective clustering technique than ST-GRID as it does not require a grid, creates more precise clusters, and is more space efficient (Wang et al., 2006). Similarly, when compared with a number of other methods (Parimala et al., 2011), ST-DBSCAN was shown to have the same clustering power. Lastly, like DBSCAN, the ST-DBSCAN algorithm can easily be implemented in the open source RapidMiner package, and the code used here was written by Alex Fechner.

ST-DBSCAN is an extension of DBSCAN, a clustering algorithm which only accounts for spatial distance, and has two parameters – minpts, the minimum number of points needed to form a cluster, and eps, the maximum distance between a pair of points within a cluster (Birant and Kut, 2007; Yue et al., 2012). Because ST-DBSCAN adds a temporal dimension, the number of clusters – or, as we refer to them, destinations – identified by ST-DBSCAN depends on three parameters, the values of which are set by the user. These parameters are spatial distance (eps1), non-spatial distance (esp2, which can be distance in time, temperature, etc.), and the minimum number of points necessary to form a cluster (minpts). For this paper, the minpts parameter was set to 10, the spatial parameter eps1 to 200 m, and the non-spatial (i.e., temporal) parameter eps2 to 20 min because these were believed to provide the most stable and parsimonious destination estimates. Roughly speaking, this means that when there are at least 10 points which are all within 200 m and 20 min of each other, the individual was in a cluster, or at a destination. It is noteworthy that the number of destinations, and consequently, the values of the other variables discussed below, would change if different values for ST-DBSCAN parameters had been used, and that the need to specify these parameters is one of the major limitations of ST-DBSCAN (Maitrey and Vaghela, 2014). As an example, Fig. 1 demonstrates how the number of clusters identified with ST-DBSCAN for each of the five participants would change if the eps1 parameter was held constant at 200 m, the eps2 parameter was varied between 15 and 30 min, and the minpts parameter took on values of 10 and 15. While only the second participant seems to have a substantial decrease in the number of ST-DBSCAN clusters as eps2 increases, for every participant, increasing the minpts parameter from 10 to 15 results in a reduction in the number of clusters that are identified. Similarly, if the eps1 parameter were varied for each individual, the ST-DBSCAN algorithm would yield a different number of clusters. However, this parameter was set at a relatively large value of 200 m for two reasons. First, because of the various errors associated with GPS data that were mentioned above, it could be the case that points might be displaced, sometimes even by tens of meters. Secondly, an individual might continue moving even once she reaches a destination (e.g., moving along a plaza).

To identify clusters (i.e., destinations), each participant's GPS data were sorted by time in RapidMiner and subjected to the ST-DBSCAN algorithm. The parameter values specified above were used to calculate destination clusters. Points which didn't fall

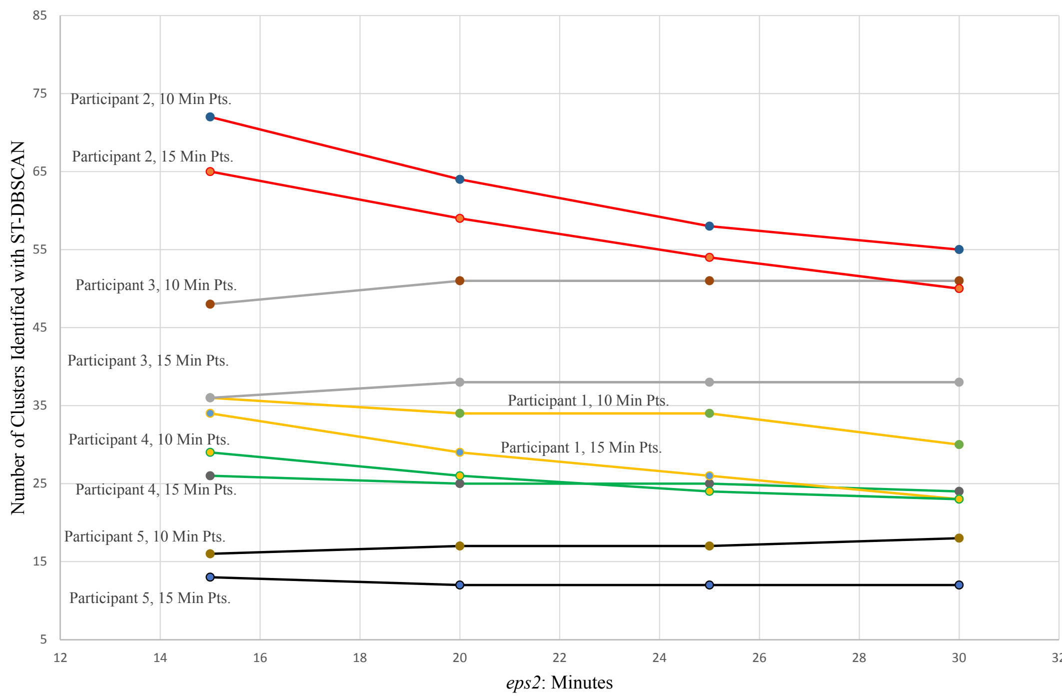

Fig. 1. Number of ST-DBSCAN Clusters by participant.

within a destination cluster were grouped into a separate cluster, called "cluster 0" in ST-DBSCAN, and were designated in-transit points. Because of missing data, the number of clusters identified through ST-DBSCAN is generally an underestimate of the total number of destinations that an individual visited during the study period. As an attempt to correct for this bias, an algorithm was developed to supplement ST-DBSCAN and enable for the identification of additional clusters and modify existing ones in the presence of missing data. This missing data algorithm, written in SAS but easily implementable in other data management packages such as R, focuses on four distinct scenarios.

In the first scenario, there are two consecutive transit points, called a and b, whose distance from one another is less than eps1 distance units (200 m), but is more than eps2 time units (20 min) apart. Although these points are not placed within a cluster by ST-DBSCAN due to the value of the eps2 parameter (and possibly the minpts parameter), the supplementary algorithm puts these two points into a cluster if they are less than 24 h apart. This is because it is likely that participants went inside a building directly after point a, lost satellite reception, and were inside the building (i.e., at a destination) until reappearing at point b. Twenty-four hours as the maximum time was selected because larger values might increase the likelihood that an individual left the building between points a and b, but either did not carry or charge the phone.

The remaining three scenarios do not identify new clusters, but either merge existing clusters with transit points, or merge two existing clusters into one. This is needed not for the identification of destinations, but for a more accurate calculation of the temporal variables, such as time spent outside of home or in transit, described below. In the second scenario, two consecutive points are identified: c is a transit point and d is part of an ST-DBSCAN destination cluster. As in scenario 1, these points are within a distance of eps1 units, but more than eps2 time units apart. An assumption of this scenario is that the person went inside after point c and did not have good satellite reception until point d. Therefore, if c and d are less than 24 h apart, transit point c is placed in the cluster which contains d. The third scenario is similar to the second, except the roles of points c and d are reversed: c is part of an ST-DBSCAN cluster and d is a transit point. If c and d are less than 24 h apart, transit point d is placed in the cluster which contains c. In the fourth scenario, two consecutive points e and f are within a distance of eps1, but fall within two separate ST-DBSCAN clusters, Cluster E and Cluster F, because they are more than eps2 time units apart. If points e and f are within 24 h of one another and the centroids of clusters E and F are within eps1 (200 m), the algorithm combines clusters E and F into a single cluster, EF. ST-DBSCAN and the supplementary algorithm described above were used to calculate the total number of destinations for the entire study period and for each of the 13 full calendar days (beginning at 12:00 a.m. and ending at 11:59 p.m.) that the individuals had the phone.

Total Number of Non-Home Destinations. After geocoding each individual's home address in ArcGIS, any destinations that had centroids within 200 m of each individual's home address were considered to be "home destinations." The total number of non-home destinations for the study period was calculated by subtracting the number of home destinations from the total number of destinations.

Total Number of Unique Destinations. When calculating the total number of destinations, it is important to remember that the destinations that were obtained were not necessarily unique. For example, if a participant returned to her house five times during a day, each of these five times at home would count as a separate destination cluster. However, the number of unique destinations may also be of interest. To create this variable, each destination cluster's centroid is calculated by averaging the latitude and longitude of all points within that cluster. Because it is likely that the centroids of the destination clusters that were very close to each other represented the same destination, each individual's destination cluster centroids are subjected to a DBSCAN analysis in order to calculate the number of "super-clusters", or unique destinations, over the study period.

As mentioned earlier, unlike ST-DBSCAN, DBSCAN does not utilize time between two points, and computes unique destinations using two parameters: spatial distance between input points (destination cluster centroids) and minimum number of points needed for a cluster to be formed. For the purposes of this paper, the spatial distance parameter was set to 200 m and the minpts parameter to 2. Thus, in the example above, DBSCAN would place the five destination clusters that corresponded to the same location (e.g., home) into a single unique destination, assuming that the centroids were all within 200 m of each other. Likewise, any destination cluster centroid that does not fall into a super-cluster after DBSCAN analysis is also representative of a unique destination.

2.2.2. Geographic scope of mobility

2.2.2.1. Activity space area. Activity spaces may be defined as "local areas within which people move or travel during the course of their daily activities" (Axhausen et al., 2006; Rai et al., 2007). Specifically, consistent with numerous other studies which have looked at individuals' activity spaces (e.g., Fan and Khattak, 2008; Villanueva et al., 2012; Buliung and Kanaroglou, 2006), the activity space was generated as the minimum convex polygon containing all of the individual's GPS points over the study period. To create each participant's activity space, the Convex Hull operation in ArcGIS was used. In addition, every individual's daily activity space was calculated for each of the 13 full days he or she was in the study. The area of the overall activity space and each of the daily activity spaces was then computed, in square kilometers. A limitation of this measure of mobility is that if an individual traveled along a road to a single destination and then went back home along the same path, the activity space would be a line, which has an area of 0.

2.2.2.2. Distance traveled. In addition to activity space area, total distance traveled for each participant was also calculated in SAS. The distances, in kilometers, between all temporally consecutive points were computed, and added up. This total included the small distances between points within a destination, which often substantially increased the distance estimate. To correct for this bias, the sum of distances between temporally consecutive points that were within destination clusters was subtracted from this estimate. These calculations were done for the entire study period and for each individual day. Here, Vincenty's formula for calculating distances between any two points on the surface of an ellipsoid was used (Vincenty, 1975).

2.2.3. Temporal scope of mobility

The total amount of time an individual is tracked, 20160 min, can be thought of as the sum of: 1) the time an individual spends at home, 2) the time the individual spends at other destinations, 3) the time the individual spends in transit, and 4) missing data time. This framework was used to calculate the following variables using SAS.

2.2.3.1. Time at home. The time elapsed between the first and last points at each home destination was computed; these time amounts were then summed across all home destinations to calculate the total amount of time the individual spent at home.

2.2.3.2. Time at other destinations. Similarly, the amount of time between the first and last points at each non-home destination was calculated. These time amounts were then added up to create the total amount of time at other destinations variable – calculated as the sum of amount of time elapsed between the first and last points within each non-home destination cluster.

2.2.3.3. Missing data time. All the calendar days which did not have a single point that was part of a destination cluster were identified. That is, the day may have had no points at all, or may only have had transit points which were not merged with any destination cluster in our destination determination process. Such days were considered to be "missing days." Each of these missing days corresponds to 60 min/hour $\times$ 24 h = 1440 min. In addition, all scenarios in the data where two consecutive points on non-missing days were more than 200 m and 2 h apart were identified. Each time such a scenario occurred for an individual, the amount of missing data as the time elapsed between these two consecutive points was calculated. The total amount of missing data time (in minutes) for an individual was then computed as the sum of 1) 1440 min for each missing day, and 2) the amount of time (in minutes) elapsed between each pair of consecutive points on non-missing days which are more than 200 m and 2 h (120 min) apart.

2.2.3.4. Time in transit. The total amount of time an individual spent in transit could then be calculated as follows: 20,160 min – Time at home (in minutes) – Time at other destinations (in minutes) – Missing data time (in minutes).

2.2.3.5. Time outside of home. his variable was computed as the sum of: 1) the time in transit and 2) time at non-home destinations.

3. Results

3.1. Sample description

Of the five participants, three were male and two were female. Three were white, one was black, and one was biracial. Participants ranged in age from 23 to 64 years (mean = 42.0, s.d. = 15.5). Three were single and never married, one was divorced and one was separated. One participant had less than a high school education, one had some college, one a bachelor's degree, and two postgraduate or professional school training. The monthly income ranged between $451 and $2000, and was, on average, $955 (s.d. = $618). Two of the participants lived alone in their own apartment or home, two with parents or other relatives, and one in a boarding home. Only participant 1owned a car.

3.2. Across-individual differences

Table 1 presents the values of all the variables described above for each of the five participants. On average, individuals had a total of 36.6 (s.d. = 15.0) destination clusters over the course of the two-week period, ranging between 18 for participant 5 and 55 for participant 3. Of these total destinations, an average of 13.0 (s.d. = 6.2) were non-home destinations, and the remaining 23.6 were home destinations. The average number of unique destinations across the 5 participants was 8.8 (s.d. = 2.3).

The total activity space area ranged between 35.6 and 6393.6 km^2 (average (s.d.) = 1333.6 (2530.2)), and the total distance traveled ranged between 100.4 km and 1167.9 km (average (s.d.) = 384.6 (400.0)). The first participant took a trip from Philadelphia to Washington, D.C. over the course of the study, which greatly skewed his average activity space upward. The average activity space area and distance traveled for the four remaining participants was substantially smaller: 68.7 (43.7) km^2 and 188.7 (105.2) km, respectively.

Only one individual (participant 3) had a day determined to be a *missing day*. On average, over the course of the study, participants had 1404.2 (s.d. = 1587.8) minutes of missing data time. Most participants spent the vast majority of their time at home – on average, the amount of time at home was 1064.8 min (17.7 h) per

Table 1
Across-individual variability in community participation and mobillity.

	Participant					Summary statistics across participants			
	1	2	3	4	5	Mean	S.D.	Min	Max
Community participation									
Total number of destinations	33	53	55	24	18	36.6	15.0	18.0	55.0
Total number of non-home destinations	14	23	15	8	5	13.0	6.2	5.0	23.0
Total number of unique destinations	8	10	12	9	5	8.8	2.3	5.0	12.0
Geographic scope of mobility									
Total activity space area (km^2)	6393.6	57.8	48.5	35.6	132.7	1333.6	2530.2	35.6	6393.6
Total distance traveled (km)	1167.9	183.5	337.9	100.4	133.1	384.6	400.1	100.4	1167.9
Temporal scope of mobility									
Missing data									
Number of missing days	0	0	1	0	0	0.2	0.4	0.0	1.0
Missing data time for the study period (in minutes)	260	2474	4026	261	0	1404.2	1587.8	0.0	4026.0
Time at home destination									
Total minutes for the study period	15,746	13,699	8360	17,820	18,911	14,907.2	3729.4	8360.0	18,911.0
Average minutes per day	1124.7	978.5	597.1	1272.9	1350.8	1064.8	266.4	597.1	1350.8
Time at other destinations									
Total minutes for the study period	2646	3431	7382	1866	854	3235.8	2241.7	854.0	7382.0
Average minutes per day	189.0	245.1	527.3	133.3	61.0	231.1	160.1	61.0	527.3
Time in transit									
Total minutes for the study period	1508	556	392	213	395	612.8	460.6	213.0	1508.0
Average minutes per day	107.7	39.7	30.2	15.2	28.2	44.2	32.7	15.2	107.7
Time outside of home									
Total minutes for the study period	4154	3987	7775	2079	1249	3848.8	2254.7	1249.0	7775.0
Average minutes per day	296.7	284.8	598.1	148.5	89.2	283.5	176.1	89.2	598.1

day (s.d. = 266.4 min or 4.4 h). The only exception was participant 3, who spent an almost equal amount of time at home on an average day (597.1 min = 10.0 h) as he did at other destinations (527.3 min = 8.8 h). The average time at other destinations was 231.1 min per day (3.9 h) and the standard deviation was 160.1 min (2.7 h). Participant 5 had the lowest amount of time at other destinations per day, which was just 61 min (1.0 h).

Participant 1 spent the most time in transit – on average, 107.7 min per day. The remaining participants spent only between 15.2 and 39.7 min per day in transit on an average day. The mean time in transit across all 5 participants was 44.2 min (s.d = 32.7). Lastly, the average daily amount of time individuals spent outside of home was 283.5 min (4.7 h; s.d. = 176.1 min, or 2.9 h).

3.3. Within-individual differences

The within-participant variability on two community participation and two geographic scope of mobility variables are shown in Table 2. Not surprisingly, each of these variables ranges widely for each individual, indicating that day-to-day participation and mobility vary substantially. Most noteworthy is the first participant, for whom the total activity space area and distance traveled ranged between 0 and 4666 km^2 and between 0 and 357.1 km, respectively. Because the participant did a long-distance drive on one of the days, the average (768.8 km^2) and median (71.5 km^2) daily activity space area were vastly different from one another, as were the average (83.9 km) and median (51.9 km) daily distance traveled. The second participant had the biggest range (1–11) when it came to the number of destinations per day. The variability in activity space area may be visualized with the help of Fig. 2. In this figure, generated with ArcScene, activity spaces for each day are stacked on top of one other, such that the vertical axis represents time and the activity space for the first day is at the very bottom and the activity space for the last day is at the very top. In-transit points, as well as points representing the home and other destinations, are also shown for each day.

4. Discussion

This paper makes a number of important contributions to the literature by describing how GPS data can be analyzed to measure various aspects of community mobility and participation. First, to the best of our knowledge, it is the most comprehensive study for any population in terms of the number of variables related to these constructs that were successfully generated from GPS data. Specifically, we have calculated the activity space area and distance traveled as measures of the geographic scope of mobility; the amount of time individuals spent at home, at none-home destinations, and in transit as measures of the temporal scope of mobility; and the number of destinations, the number of non-home destinations and the number of unique destinations as measures of community participation.

Second, detailed descriptions of how each variable has been generated are offered as a starting point to developing standardized approaches to how GPS data are used. An innovative approach employing ST-DBSCAN, a recently developed algorithm, has been used for the identification of destinations. In addition to the total number of destinations, unique destinations were also determined, which is a construct that has been minimally discussed in the existing literature. Unlike the total number of destinations, indicative of the overall extent of community participation, the unique destinations variable is an indicator of variety of participation – a related but separate construct, which when measured with self-report data, has been found to be positively associated with psychosocial outcomes such as quality of life and recovery in the psychiatric disabilities population (Burns-Lynch et al., in press).

Third, earlier studies lack a thorough explanation of how missing data are addressed when computing the different variables from GPS data. Missing data are inevitable with GPS data collection, and if not dealt with appropriately or in a standardized way, may lead to unreliable or invalid conclusions. A new algorithm which supplements ST-DBSCAN was created to account for missing data in the process of identifying the destination variables. Similarly, a detailed description was offered for how missing data have been handled in the process of generating the temporal variables.

Fourth, results from five participants using the aforementioned

Table 2
Within-individual variability in community participation and mobillity.

	Participant				
	1	2	3	4	5
Community participation					
Number of destinations					
Daily average ± SD	3.3 ± 1.9	4.5 ± 2.9	4.5 ± 2.6	2.8 ± 2.3	2.1 ± 1.9
Daily range [min, max]	[1, 7]	[1, 11]	[1, 10]	[1, 8]	[1, 7]
Median	4	5	4	2	1
Number of unique destinations					
Daily average ± SD	2.3 ± 1.6	2.7 ± 1.4	3.3 ± 1.5	2.1 ± 1.6	1.6 ± 1.2
Daily range [min, max]	[1, 6]	[1, 5]	[1, 7]	[1, 5]	[1, 5]
Median	2	3	3	1	1
Geographic scope of mobility					
Activity space area (km^2)					
Daily average ± SD	768.8 ± 1709.6	4.2 ± 8.3	17.3 ± 21.5	4.0 ± 7.6	4.5 ± 9.1
Daily range [min, max]	[0, 4666.0]	[0, 30.6]	[0.7, 64.6]	[0, 26.8]	[0, 30.1]
Median	71.5	1.0	11.5	0	0
Total distance traveled (km)					
Daily average ± SD	83.9 ± 118.7	7.1 ± 10.4	18.1 ± 15.3	4.0 ± 6.2	6.9 ± 11.5
Daily range [min, max]	[0, 357.1]	[0, 37.4]	[3.7, 59.9]	[0, 18.7]	[0, 36.4]
Median	51.9	2.5	15.5	0	0

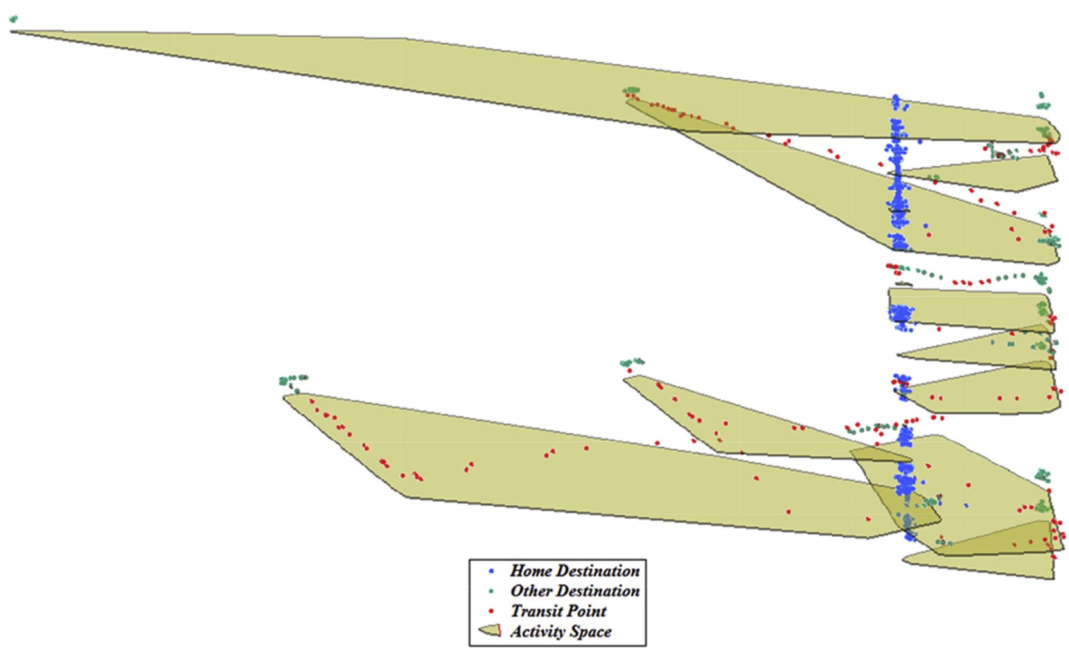

Fig. 2. Daily activity spaces of participant 2.

methods exemplify the types of results that can be obtained and demonstrate variability in the constructs, suggesting that these measures may be effectively used as independent and dependent variables in future research. In particular, the five individuals tracked in this paper differed substantially from one another on all variables measured. Some of the participants left their homes on most of the days and had relatively high numbers of destinations, amount of time spent out of the house, activity space area, and total distance traveled; others, like participants 4 and 5, generally chose to stay in, and had lower values on all of these variables. This across-individual variability observed in the different mobility and participation variables is not surprising and is consistent with prior research where subjective measures of participation have been used. For example, Salzer et al. (2014) have shown that there is a wide range in both the extent and variety of participation in a sample of individuals with psychiatric disabilities.

In terms of within-individual results, the values of the variables differ for each individual from one day to the next. In particular, the first participant indicated in a post-data collection interview that on one of the days, he drove from his home in the suburbs of Philadelphia to Washington, DC, and returned home in the evening. On that day, his activity space area was 4666 square kilometers, and the distance he traveled was 357.1 km. However, there were several days when this participant did not leave his house, and on those days, the activity space area and distance traveled were both zero. While the trip to Washington might be an atypical event, it serves as a reminder that variability in community participation and mobility should be examined not only across individuals, but also within individual and that a single day may substantially skew average values of these variables. For this reason, future studies reporting daily levels of mobility and participation should present median values of those variables in addition to the means, as is done here.

The within-individual variability in mobility and participation observed in this paper is consistent with findings in studies that have used subjectively reported data on participation. For instance, a recent study by Salzer et al. (2015) has found similar variability in daily participation levels among individuals with psychiatric disabilities who, over the course of 30 days, have used a daily activity checklist to indicate which areas they participated in on that day.

Nonetheless, for some participants, mobility and participation levels looked alike on several of the days. For instance, as can be seen in Fig. 2, the activity space polygons looked virtually identical on some days for participant 2. This may be due to the presence of a routine, where the individual goes to particular places on certain days; however a longer tracking period is needed to gain a better understanding of what this routine might be.

Lastly, although the methods presented in this study may be used for any population, to the best of our knowledge, this is the first study using GPS technology to assess community mobility and participation of individuals with psychiatric disabilities. In this population in particular, this methodology may prove to be an innovative and useful way to complement self-report data, which are usually gathered with diaries or retrospective community participation measures, such as the Temple University Community Participation Scale (Salzer et al., 2014). There have been other suggested uses of GPS with this population. For instance, a recent study has proposed – but not yet performed – GPS tracking of individuals diagnosed with bipolar disorder to assess mania (i.e., increased community movements) and depression (i.e., decreased movements) (Prociow and Crowe, 2010). Ben-Zeev et al. (2015) were the first to examine the utility of GPS-enabled cell phones for the actual detection of mental health problems. Using a sample of 47 young adults ages 19–30, they found that the extent of geospatial activity captured with the GPS was associated with levels of daily stress and with changes in depression levels. They suggest that smartphones could be used for “close-to-invisible psychiatric assessment” with unprecedented efficiency.

4.1. Limitations

Because GPS data show where people are at any given point in time, this method of data collection may be viewed as overly intrusive. However, none of the five participants expressed any concerns about having their location tracked; on the contrary, at the follow up interviews, they were very interested in looking at the maps that showed their mobility over the course of their study. Furthermore, AccuTracking only recorded the location of the individual and not what the individual was doing at that location, and if participants wanted to stop tracking at any time, they could simply turn off the phone. Nevertheless, while no protected health information is collected, GPS data are highly personal, and clearly outlined ways to ensure privacy and confidentiality of participants would be required in a study which uses this method. For this paper, several such safeguards were in place. First, data were sent in real time to a secure, password protected AccuTracking database, from which data over 30 days old were automatically purged. Second, the researchers' institution was the owner of the cell phone and AccuTracking accounts; that way, only the researchers had the information linking a participant to a phone number, so in the unlikely case that the security of the AccuTracking database was breached, there would be no link between the phone number and the participant. Third, in case the participant would lose the phone or have it stolen, GPS data could not be accessed on the phone without a password. Fourth, the phone was restored to factory settings after each use, before it was given to a new participant. No problems or concerns related to privacy were reported.

An additional possible limitation is that we may have been able to achieve better measurement precision and less uncertainty if we had opted to use a stand-alone GPS unit, especially one which would enable us to record participants' locations every 10 or 30 s. However, there could be problems with such an approach: it may be the case that participants would be more reluctant to carry the stand-alone units because they don't provide all the capabilities of a phone. Furthermore, if the participants were to lose these GPS units, all the data would be lost as well, which is not the case with AccuTracking. And while it is possible to record location at shorter time intervals, this is likely to be a problem regardless of whether a GPS-enabled phone or a stand-alone GPS unit is used, as the battery of the device would be drained more quickly.

Lastly, it is important keep in mind that the data from five participants presented in this paper are not intended for making inferences about community mobility and participation of the overall population of individuals with psychiatric disabilities, and that studies with a larger sample size would need to be conducted for that purpose. Rather, the data are included here in order to demonstrate the feasibility of tracking individuals with various health conditions with GPS-enabled mobile phones, and to show the different variables that may be generated from GPS data and how they vary within and across individuals.

4.2. Conclusion

This paper used global positioning technology to track five individuals over a course of two weeks and employed a modified version of the spatiotemporal data mining algorithm ST-DBSCAN to create several variables related to community participation and the geographic and temporal scopes of mobility. Results showed that there was substantial within- and across- individual variability on each of the variables measured when the aforementioned ST-DBSCAN parameter values were used and certain assumptions about missing data were made. Future research is needed to determine whether alternative parameter values and missing data assumptions would be more appropriate than the ones used here, and whether they should vary for different populations or individuals. One way to shed more light on this would be to ask participants about the amount of time they would need to spend at various destinations to achieve meaningful participation, as well as what happened during the periods for which there was no GPS data (e.g., left the phone at home; phone wasn't charged; didn't leave the house, etc.), and incorporate this information in the analysis. Participants could also be asked to verify the destinations that were detected with ST-DBSCAN, which could be achieved by asking them to record all destinations at which they were over the course of the study in a diary; similarly, they could record the amount of time they spent at each destination to validate the temporal variables that were described in this paper. However, as mentioned earlier, one should not necessarily trust the self-reported data over GPS data, as self-report data have their own limitations, including recall bias. Furthermore, additional studies with a larger sample size are needed in order to examine how the different variables related to mobility and participation are associated with various indicators of the natural, built and social environment. For instance, individuals whose activity spaces include more resources, or have lower poverty rates based on Census data, might have greater mobility and participation. Finally, these studies could also examine the relationship between the different variables discussed in this paper and subjective measures of community participation and various physical and mental health outcomes.

Acknowledgments

The contents of this publication were developed under a grant from the U.S. Department of Health and Human Services, National Institute on Disability, Independent Living, and Rehabilitation Research (NIDILRR; Grant #901F0065-02-00; Salzer, Principal Investigator). However, the contents do not necessarily represent the policy of the U.S. Department of Health and Human Services, and endorsement by the Federal government should not be assumed. The authors are grateful to Andrea Bilger and David O'Malley for assistance with data collection and analysis, and to Alex Fechner for providing his implementation of the ST-DBSCAN algorithm in RapidMiner.

References

Axhausen, K.W., Rai, R.K., Rieser, M., Balmer, M., Vaze, V.S., 2006. Capturing human activity spaces. ETH, Eidgenössische Technische Hochschule Zürich, IVT, Institut für Verkehrsplanung und Transportsysteme. http://dx.doi.org/10.3929/ethz-a-005237115.

Ben-Zeev, D., Scherer, E.A., Wang, R., Xie, H., Campbell, A., 2015. Next-generation psychiatric assessment: using smartphone sensors to monitor behavior and mental health. Psychiatr. Rehabil. J. 38 (3), 218–226.

Birant, D., Kut, A., 2007. ST-DBSCAN: an algorithm for clustering spatial–temporal data. Data Knowl. Eng. 60 (1), 208–221.

Bostelmann, R., Schneller, S., Cornelius, J.F., Steiger, H.J., Fischer, I., 2015. A new possibility to assess the perioperative walking capacity using a global positioning system in neurosurgical spine patients: a feasibility study. Eur. Spine J. 1–6.

Buliung, R.N., Kanaroglou, P.S., 2006. Urban form and household activity-travel behavior. Growth Change 37 (2), 172–199.

Burns-Lynch, W., Brusilovskiy, E., Salzer, M.S., 2016. Community participation facilitates recovery and quality of life: an empirical study. Isr. J. Psychiatry (in press).

Carpenter, W.T., Gold, J.M., Kovnick, J., Appelbaum, P.S., 2000. Decisional capacity for informed consent in schizophrenia research. Arch. General Psychiatry 57 (6), 533–538.

Clark, R.A., Weragoda, N., Paterson, K., Telianidis, S., Williams, G., 2014. A pilot investigation using global positioning systems into the outdoor activity of pepole with severe traumatic brain injury. J. Neuroeng. Rehabil. 19, 11–37.

Collins, P., Al-Nakeeb, Y., Lyons, M., 2015. Tracking the commute home from school utilizing GPS and heart rate monitoring: establishing the contribution to free-living physical activity. J. Phys. Act. Health 12, 155–162.

Eklund, M., Leufstadius, C., Bejerholm, U., 2009. Time use among people with psychiatric disabilities: implications for practice. Psychiatr. Rehabil. J. 32 (3), 177–191.

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

Harris, F., Sprigle, S., Sonenblum, S.E., MSaurer, C.L., 2010. The participation and activity measurement system: an example application among people who use wheeled mobility devices. Disabil. Rehabil. Assist. Technol. 5 (1), 48–57.

Hillier, A., 2008. Childhood overweight and the built environment: making technology part of the solution rather than part of the problem. Ann. Am. Acad. Political Soc. Sci. 615 (1), 56–82.

Hillsdon, M., Coombes, E., Griew, P., Jones, A., 2015. An assessment of the relevance of the home neighbourhood for understanding environmental influences on physical activity: how far from home do people roam? Int. J. Behav. Nutr. Phys. Act. 12 (100).

Hordacre, B., Barr, C., Crotty, M., 2015. Community activity and participation are reduced in transtibial amputee fallers: a wearable technology study. BMJ Innov. 1, 10–16.

Jayaraman, A., Deeny, S., Eisenberg, Y., Mathur, G., Kuiken, T., 2014. Global position sensing and step activity as outcome measures of community mobility and social interaction for an individual with a transfemoral amputation due to dysvascular disease. Phys. Ther. 94 (3), 401–410.

Ji, Y., Zhang, C., Zuo, Z., Chang, J., 2012. Mining user daily behavior based on location history. In: IEEE 14th International Conference on Communication Technology (ICCT).

Khetarpaul, S., Chauhan, R., Gupta, S.K., Subramaniam, L.V., Nambiar, U., 2011. Mining GPS data to determine interesting locations. In: Proceedings of the 8th International Workshop on Information Integration on the Web: in Conjunction with WWW 2011. ACM, Hyderabad, India, pp. 1–6.

Lankton, S., Sonenblum, S.E., Sprigle, S., Wolf, J., Oliveira, M., 2005. Use of GPS and Sensor-based Instrumentation as a Supplement to Self-report in Studies of Activity and Participation. Paper presented at RESNA 28th International Annual Conference, Atlanta, GA.

Lu, Y., Karimi, H.A., 2015. Real-time sidewalk slope calculation through integration of GPS trajectory and image data to assist people with disabilities in navigation. Int. J. Geo-information 4 (2), 741–753.

Maitrey, N., Vaghela, D., 2014. Survey on different density based algorithms on spatial dataset. Int. J. Adv. Res. Comput. Sci. Manag. Stud. 2 (2).

Parimala, M., Lopez, D., Senthilkumar, N.C., 2011. A survey on density based clustering algorithms for mining large spatial databases. Int. J. Adv. Sci. Technol. 31 (1).

Prociow, P.A., Crowe, J.A., 2010. Towards personalized ambient monitoring of mental health via mobile technologies. Technol. Health Care 18 (4–5), 275–284.

Rai, R., Balmer, M., Rieser, M., Vaze, V., Schönfelder, S., Axhausen, K., 2007. Capturing h*uman* activity spaces: new geometries. Transp. Res. Rec. J. Transp. Res. Board 2021, 70–80.

Rogalsky, J., 2010. The working poor and what GIS reveals about the possibilities of public transit. J. Transp. Geogr. 18 (2), 226–237.

Salzer, M.S., Brusilovskiy, E., Kottsieper, P., Prvu-Bettger, J., 2014. Measuring community participation of adults with psychiatric disabilities: reliability of two modes of data collection. Rehabil. Psychol. 59 (2), 211–219.

Salzer, M.S., Kottsieper, P., Brusilovskiy, E., 2015. Intermethod reliability and factors affecting recall with the Temple university community participation measure. J. Ment. Health 24 (4), 189–195.

Trimble Navigation Limited, 2004. Why Postprocess GPS Data? Mapping and GIS White Paper Trimble Navigation Limited, Westminster, CO.

Ustun, T.B., Chatterji, S., Bickenbach, J., Kostanjsek, N., Schneider, M., 2003. The international classification of functioning, disability and health: a new tool for understanding disability and health. Disabil. Rehabil. 25 (11–12), 565–571.

Villanueva, K., Giles-Corti, B., Bulsara, M., McCormack, G.R., Timperio, A., Middleton, N., Trapp, G., 2012. How far do children travel from their homes? Exploring children's activity spaces in their neighborhood. Health & Place 18 (2), 263–273.

Vincenty, T., 1975. Direct and inverse solutions of geodesics on the ellipsoid with application of nested equations. Surv. Rev. 23 (176), 88–93.

Wang, M., Wang, A., Li, A., 2006. Mining spatial-temporal clusters from geo-databases. In: Li, X., Zaïane, O., Li, Z.-H. (Eds.), Advanced Data Mining and Applications, vol. 4093. Springer, Berlin Heidelberg, pp. 263–270.

Wettstein, M., Wahl, H.-W., Diehl, M., 2013. A multidimensional view of out-of-home behaviors in cognitively unimpaired older adults: examining differential effects of socio-demographic, cognitive, and health-related predictors. Eur. J. Ageing 1–13.

Wettstein, M., Wahl, H.-W., Heinik, J., Landau, R., 2012. Out-of-home behavior and cognitive impairment in older adults: findings of the SenTra project. J. Appl. Gerontol. 34 (1), 3–25.

Whiteneck, G., 2006. Conceptual models of disability: past, present, and future. In: Field, M., Jette, A., Martin, L.M. (Eds.), Workshop on Disability in America: a New Look—summary and Background. The National Academies Press, Washington, DC.

World Health Organization, 2001. ICF: International Classification of Functioning, Disability, and Health (Geneva).

Yanos, P.T., Robilotta, S.A., 2011. An examination of time-use among adults diagnosed with severe mental illnesses using daily interviews. Psychiatr. Rehabil. J. 34 (3), 243–247.

Yin, L., Raja, S., Epstein, L., Roemmich, J., 2013. Neighbourhood for playing: using gps, gis and accelerometry to delineate areas within which youth are physically active. Urban Stud. 2013, 1–18.

Yue, Y., Chen, J., Zuo, X., Xie, X., 2012. Labeling personal characteristics from mobile phone traces. In: 2nd International Workshop on Mobile Sensing, IPSN'12 and CPSWeek.

Zandbergen, P.A., Barbeau, S.J., 2011. Positional accuracy of assisted GPS data from high-sensitivity GPS-enabled mobile phones. J. Navig. 64 (3), 381–399.

Zhou, C., Frankowski, D., Ludford, P., Shekhar, S., Terveen, L., 2004. Discovering personal gazetteers: an interactive clustering approach. In: Proceedings of the 12th Annual ACM International Workshop on Geographic Information Systems. ACM, pp. 266–273.

Zhou, C., Ludford, P., Frankowski, D., Terveen, L., 2005. How do people's concepts of place relate to physical locations?. In: INTERACT Conference. Rome, Italy.