Activity Change Index (ACI) – new aspects of indoor mobility identified from GPS and lifelogging data

K. Sila-Nowicka*[1,3] and P. Thakuriah[2]

[1]School of Environment, Faculty of Science, The University of Auckland, 23 Symonds Street, 1142 Auckland, New Zealand
[2] Edward J. Bloustein School of Planning and Public Policy, Rutgers University – New Brunswick, 33 Livingston Ave., New Brunswick, USA
[3]Urban Big Data Centre, University of Glasgow, 7 Lilybank Gardens, Glasgow, UK
*Email: Katarzyna.sila-nowicka@auckland.ac.nz

Abstract

Global Positioning System (GPS) technology has changed the world. We now rely on it to navigate our vehicles, source real-time information about chosen locations and to track our movements. GPS is, however, principally limited to receiving signals outdoors and therefore not able to provide reliable information about indoor movements. This paper examines patterns of indoor physical activity and sedentary behaviour using a combination of lifelogging and GPS movement data to develop an index for capturing activity changes. In our preliminary analysis, we explore possible associations between people's indoor physical activity levels and their socio-demographic characteristics.

Keywords: GPS movement data, lifelogging, indoor mobility, sedentary behaviour, physical activity.

1. Introduction

In comparison to our parents' or grandparents' generations, we spend a far greater proportion of time indoors, limiting our levels of physical activity. The average adult now spends more than 90% of their time in an indoor environment (Bruinen de Bruin *et al.*, 2008). Social and technological changes have led us into an era of increasingly sedentary lifestyles that have been linked with decreased productivity at work, poor metabolic health and a greater risk of heart disease (Siddarth *et al.*, 2018). Due to the considerable length of time we spend in various types of enclosed spaces (e.g., homes, workplaces, cars and public transport), a measure of the levels of indoor physical activity would provide valuable behavioural information.

Global Positioning System (GPS) technology provides reliable and high-precision location data of people's whereabouts. We can use this for navigation as well as for tracking our movements for a variety of purposes such as outdoor sporting and leisure activities. However, while GPS can receive signals outdoors, structural interference means that it is not able to provide reliable information about indoor movements.

Without recognising whether an activity takes place indoors or outdoors, a number of studies have looked at pattern recognition methods to determine the mode or type of activity being performed from raw acceleration data (including sitting, standing, walking, running, cycling, vacuum cleaning, Nordic walking, ascending and descending stairs, and lying down (Arif and Kattan, 2015; Twomey *et* *al.*, 2018)). Relating the amount of time spent on each of these activities to medical records data allows us to drive to conclusions about the negative effects of non-exercising behaviours (Loveday *et al.*, 2015). Furthermore, breaks in sedentary time can be shown to have favourable associations with body mass index (BMI), triglyceride levels and glucose levels that are independent of the total time spent sitting or engaged in physical activities (Healy *et al.*, 2008).The integration of accelerometery and GPS in wearable technology gives us the capability to "assess the indoor and outdoor location of physical activity and sedentary behaviour" (Loveday *et al.*, 2015).

With this in mind, in this paper we propose a frequency-based method that uses a count of changes in activity derived from a combination of GPS and lifelogging data as a measure of physical activity.

2. Data and case study

In this research we use the Integrated Multimedia City Data (iMCD) platform that covers the Greater Glasgow Urban Area, UK. The is a multi-modal data platform that consists of seven strands of data: a socio-demographic participant survey with travel and activity diaries (2095 participants), social media with combined local and national news website data, remote sensing data, sensor data (333 participants who collected GPS movement data; 223 participants with additional lifelogging images and sensor readings), specialised private sector datasets and other administrative data such as Census data (detailed information about all the strands is available in Thakuriah, Sila-Nowicka and Paule, 2016).

The data used in this study come from a sample of 160 participants and consist of a set of individual-related records, where each record is a person's single activity derived from their GPS and lifelogging data (Figure 1).

Participant ID	Image ID	Timestamp	Activity derived from lifelogging data	Location and activity derived from GPS
1	1234	12/09/2015 12:37:08	Walking Indoor	Home
1	1235	12/09/2015 12:37:15	Walking Indoor	Home
1	1236	12/09/2015 12:37:22	Walking Outdoor	Walk
1	1237	12/09/2015 12:37:29	Walking Outdoor	Walk

Table 1: Data structure of the used dataset.

In order to derive the type of activity, we used a combination of GPS movement data and lifelogging data (Figure 1) which were pre-processed as follows:

The collected in the project GPS data were semantically enriched using two-step feedforward neural network with a general backpropagation algorithm for trajectory classification. We first distinguished the movement (walk, drive etc.) from the non-movement segments. The stop-related segments were classified based on their location and importance to the user starting with 'home' and then into a set of significant locations such as 'work', 'school', 'third place'. Other stops were compared to a Points/Places of Interest dataset (a combination of Ordnance Survey, OSM and self-created POI dataset) in order to semantically enrich them with their functions such as: shopping, leisure, education, health or transport-related (Siła-Nowicka *et al.*, 2016).

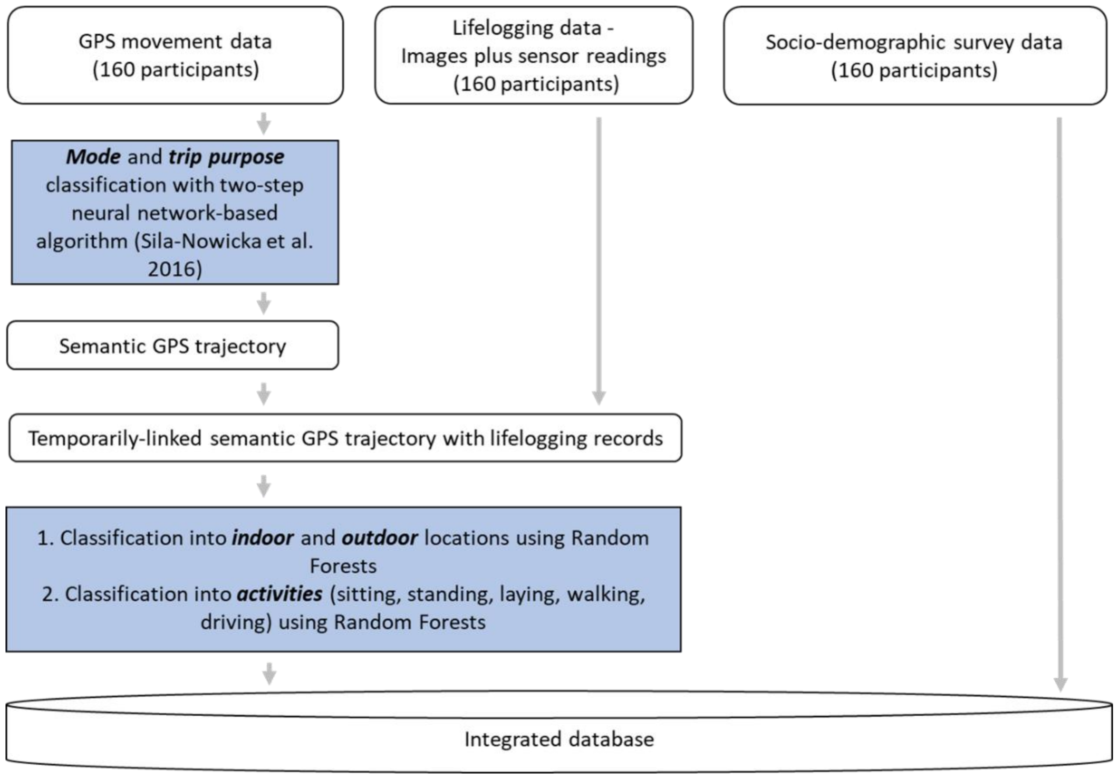

Figure 1: A process of creating the integrated database used in this project.

The Lifelogging data were collected from a wearable camera (Autographer) carried by the project participants for up two consecutive days. These cameras take images at time interval of 7 seconds, and every image generates a set of associated sensor readings (the sensors are accelerometer, motion detector, magnetometer, thermometer, GPS sensor and a brightness detector). The sensor readings were used to separate indoor and outdoor locations as well as to determine the type of activity e.g, driving and walking for outdoors and sitting, standing, walking or lying for indoors. In order to make these classifications we used the Random Forest classifier designed by Sila-Nowicka and Thakuriah, (2019).

The processed GPS and lifelogging data were then temporally linked resulting in an integrated database of individual images with an assigned travel mode or type of activity and its possible indoor or outdoor location. In the resultant dataset, the average daily recorded time of indoor activities was 5:02 hours, with min of 1:23h and max 14h: 54h.

The socio-demographic information about each of the participants consists of age, gender, income, working status and general health information; BMI and participants' attitude to walking. The average age of respondent is about 48 years old, nearly half of the respondents are male, 69% are actively working and more than 70% claim to be relatively healthy.

3. Method

In this research we limited ourselves to only indoor locations as our focus is on the levels of physical activity in enclosed spaces. We designed a frequency-based index which we call Activity Change Index (ACI) that uses a count of the changes in type of activity (e.g. from sitting to standing) derived from a combination of GPS and lifelogging data as a measure of physical activity. ACI can be written as:

$$ACI_i = \frac{C_i}{M * T_i},$$

Equation 1

where ACI is the Activity Change Index, C_i is the count of all activity changes, M is the maximum number of changes possible in one hour and T_i is the total time spent indoors (in hours) by an individual *i*. With the temporal resolution of the integrated database at 7 seconds (corresponding to the lifelogger collection interval), the maximum number of changes in each hour would not exceed 514 changes.

4. Results

The levels of ACI vary from 0.0004 to 0.07875, where the lowest value corresponds to one change of activity within a four-hour period and the highest corresponds to a change every two minutes. The mean value of the ACI across the studied population is equal to 0.01252 which translates to a change of activity every 10 minutes.

For a more in-depth analysis, we used beta regression to study the associations between ACI and socio-demographic and health-related characteristics of the survey participants. The dependent variable was ACI with a range of values from 0 to 1. The results from the modelling are presented in Table 2. Several factors show significant associations with the levels of Activity Change Index.

	β	SE	P-Value	
Intercept	-3.982	0.409	0.000	***
Socio-demographic				
Age	-0.009	0.004	0.027	*
BMI	0.011	0.013	0.414	
Gender (Female as a reference)	-0.228	0.115	0.048	*
Work (Employed as a reference)	0.049	0.143	0.734	
Driving licence (Does not own as a reference)	0.286	0.146	0.050	*
Pet (No as reference)	0.070	0.121	0.561	
Health (Bad as a reference)				
Health-Fair	0.504	0.254	0.047	*
Health-Good	0.011	0.219	0.959	
Income	0.000	0.000	0.731	
Attitudes (For me, walking is something I like (disagree to agree))				
Walking attitude (Agree as a reference)				
Walking attitude - neutral	-0.803	0.272	0.003	**
Walking attitude -disagree	0.065	0.205	0.751	
Pseudo-R^2	0.37			
AIC	417.98			

Table 2: Associations between ACI and socio-demographic factors.

Older people and male participants have the lowest values of ACI meaning they change their activities the least frequently. People who reported their health as fair tend to change modes or types of activities more regularly than people reporting bad health. Drivers tend to be more active in indoor locations. Participants who stated that they have a neutral attitude to walking tend to have longer individual activities in comparison to people who like to walk. The presented results show some significant but in most cases to be expected associations between the levels of activities and different socio-demographical and health-related aspects.

5. Conclusions and future work

Our overall goal in this paper has been to evaluate whether a combination of GPS and lifelogging data can provide a way to detect patterns in the levels of indoor physical activities. The ACI represents an average activity change level which proves to be useful information describing sedentary patterns of individuals. Nevertheless, as the index represents the average values across time it generalizes possible hourly differences. To overcome this problem, in future research we will create hourly-based time series from the data and will look at the changes across the time of day as well as the activity type in different locations (home, work place, school etc.) to understand how different locations, influence possible levels of physical activities. Furthermore, in order to show more dependencies we plan to cluster the survey's participants.

6. . Acknowledgements

This work was supported by the ESRC Grant (grant number ES/L011921/1). The authors would also like to thank all the survey participants.

7. References

Arif, M. and Kattan, A. (2015) 'Physical activities monitoring using wearable acceleration sensors attached to the body', *PLoS ONE*. Edited by A. Lucia. Public Library of Science, 10(7), p. e0130851. doi: 10.1371/journal.pone.0130851.

Bruinen de Bruin, Y. *et al.* (2008) 'Characterisation of urban inhalation exposures to benzene, formaldehyde and acetaldehyde in the European Union', *Environmental Science and Pollution Research*, 15(5), pp. 417–430. doi: 10.1007/s11356-008-0013-4.

Healy, G. N. *et al.* (2008) 'Breaks in Sedentary Time', *Diabetes Care*. American Diabetes Association, 31(4), pp. 661–666. doi: 10.2337/DC07-2046.

Loveday, A. *et al.* (2015) 'Technologies that assess the location of physical activity and sedentary behavior: A systematic review', *Journal of Medical Internet Research*. Journal of Medical Internet Research, 17(8), p. e192. doi: 10.2196/jmir.4761.

Owen, N. *et al.* (2010) 'Sedentary Behavior: Emerging Evidence for a New Health Risk', *Mayo Clinic Proceedings*. Mayo Foundation, 85(12), pp. 1138–1141. doi: 10.4065/mcp.2010.0444.

Siddarth, P. *et al.* (2018) 'Sedentary behavior associated with reduced medial temporal lobe thickness in middle-aged and older adults', *PLoS ONE*. Edited by S. D. Ginsberg. Public Library of Science, 13(4), p. e0195549. doi: 10.1371/journal.pone.0195549.

Siła-Nowicka, K. *et al.* (2016) 'Analysis of human mobility patterns from GPS trajectories and contextual information', *International Journal of Geographical Information Science*. Taylor & Francis, 30(5), pp. 881–906. doi: 10.1080/13658816.2015.1100731.

Sila-Nowicka, K. and Thakuriah, P. (2019) 'Multi-sensor movement analysis for transport safety and health applications', *PLoS ONE*. Edited by A. N. Stephens. Public Library of Science, 14(1), p. e0210090. doi: 10.1371/journal.pone.0210090.

Thakuriah, P., Sila-Nowicka, K. and Paule, J. G. (2016) 'Sensing spatiotemporal patterns in urban areas: Analytics and visualizations using the integrated multimedia city data platform', *Built Environment*, 42(3).

Twomey, N. *et al.* (2018) 'A Comprehensive Study of Activity Recognition Using Accelerometers', *Informatics*. Multidisciplinary Digital Publishing Institute, 5(2), p. 27. doi: 10.3390/informatics5020027.