

UC Santa Barbara

Spatial Data Science Symposium 2021 Short Paper Proceedings

Title

Deriving Spatio-Temporal Geographies of Human Mobility from GPS traces

Permalink

<https://escholarship.org/uc/item/1h9971z5>

Authors

Calafiore, Alessia
Murage, Nombuyiselo
Nasuto, Andrea
et al.

Publication Date

2021-12-01

DOI

10.25436/E26K5F

Peer reviewed

Deriving Spatio-Temporal Geographies of Human Mobility from GPS traces

Alessia Calafiore¹, Nombuyiselo Murage¹, Andrea Nasuto¹, and Francisco Rowe¹

¹ Geographic Data Science Lab, University of Liverpool, Liverpool, United Kingdom
² a.calafiore@liverpool.ac.uk

Abstract. This paper leverages on the opportunities presented by individual-level GPS data to study human mobility. It develops a methodology to understand the spatio-temporal properties of collective movements using network science. Through a spatially-weighted community detection approach, we derived functional neighbourhoods from human mobility patterns from GPS data and analyse the extent to which they vary across time. The results show that while the overall city structure remains stable, functional neighbourhoods tend to contract and expand over the course of the day. This work proposes a methodological framework and emphasises the importance of detecting short-term structural changes in cities based on human mobility.

Keywords: Human Mobility · GPS Digital Traces · Community Detection

DOI: <https://doi.org/10.25436/E26K5F>

1 Introduction

Time is a fundamental dimension to understand human behaviours across space. Time geography [1–3] is a widely recognised framework to capture how human mobility is constrained by space and time. It conceives an individual’s spatial behaviours as a “continuous temporal sequence of activities in the geographical space” [4], where the space and time of each activity is determined by its specific spatio-temporal constraints: the amount of available time and the nature of the activity.

Building upon the ideas behind time geography, we move our focus from individual’s activities to collective movements of people and we investigate how everyday life in cities is marked by human mobility across space and time. In particular, we aim to identify the extent to which the relationship between places varies across different daytime intervals as inferred from individual mobility patterns.

The increasing availability of data at high spatio-temporal resolution poses challenges, as well as novel opportunities, to elaborate methods which uncover

human mobility patterns [5]. Alessandretti et al. [6] analysed anonymised GPS traces to identify mobility containers. Containers are modelled from each person’s empirical mobility pattern and encapsulate locations at different levels so that each individual trace results in a unique hierarchical structure. Their results show that day-to-day human mobility contains meaningful scales, as containers tend to have typical sizes across, on average, four hierarchical levels. While these results shed light on the geographical ‘scales’ of human mobility, the extent to which time plays a role in determining spatial scales is still unrecognized. Advancing our knowledge of time geography is a matter of great concern in fields, such as transportation planning [7] and to capture places as complex systems [8, 9].

In this work, we leverage on GPS-trajectories describing people movements to derive how cities are spatio-temporally structured. Specifically, we employ network analysis to identify the interdependence of places in relation to human mobility and derive functional neighbourhoods [10] across different parts of the day, and evaluate their spatial stability. Functional neighbourhoods have been defined as “cohesive zones of interconnected agglomerations of activity emerging from human mobility dynamics” [10]. We extend Calafiore et al. [10] study by looking at how functional neighbourhoods vary across different time windows.

2 Data

We used mobile phone GPS data from Tamoco, a business intelligence company based in the UK. We had access to pre-processed non-public, non-identifiable data collected via mobile network applications from different vendors, mobile phone carriers and operating system platforms and versions. The data had the following attributes: device ID, timestamp, latitude and longitude. To avoid the identification of exact home-locations and improve computational efficiency, individual exact GPS locations data point were converted by Tamoco from point to into Uber H3 hexagon geometries using Ubers’ H3 Spatial Indexing library at resolution 9 [11], corresponding to $\sim 0.1 \text{ Km}^2$. For each individual trajectory the data points falling within the same hexagon have been aggregated, providing start and end time of staying within the cell. This process also included removing noise in the trajectory: To remove noise when data points belonging to H3 cells too far from each other given the time they were recorded have been removed. This process aimed at hiding exact home locations and reducing data size. From a dataset of ~ 1.19 billion mobile phone GPS traces, we obtained approximately 80.27 million trajectories as sequence of H3 cells, generated by ~ 6.8 million unique users in New York City in the United States during October 2019. While data binning reduced the accuracy in detecting localised movements, we still rely on a quite detailed geographical scale.

3 Methods

The objectives of this study are to: 1) identify functional neighbourhoods [10] at different time intervals; 2) determine how stable the shape and size of these neighbourhoods remain across time. To achieve this, we implemented the following methods.

Classifying work-home-leisure places: We classified places based on the activities that are more likely to happen there. Specifically, we distinguished between three classes - work, home and leisure places. To do this, we first identified places that have a duration of stay longer than 15 minutes, and can therefore be discerned from areas where people are only passing by. Secondly, we applied a set of heuristics to determine which places for each individual trajectory are more likely home, work or leisure places. Following [12], home places are detected when the activity time spans more than 2 hours within the period of 22:00 pm to 08:00 am the next day and have visits for more than 15 days. Also, all places where the user spends more than 10 hours consecutively are considered possible home locations. Work locations are detected when the total activity time spans more than 2 hours within the period of 08:00 am to 12:00 pm or 14:00pm to 18:00 pm on a weekday³ and the number of visits are more than half of the working days of the month. Leisure places are defined as where all activities which exclude work and staying at home happens. In this study, we decided to focus on leisure places as we centered our interest in activities that tend to have less regular temporal and spatial patterns than commuting to the workplace, therefore, all home and workplaces are removed.

Deriving Space-Time Networks from Human Mobility: Following the classification and detection of leisure places, we generated 4 networks corresponding to leisure activities happening during 4 time intervals: Morning [07AM-11AM], Midday [11AM-03PM], Afternoon [03PM-07PM], Evening [07PM-10PM]. These networks connect people to various leisure places and take the form of two-mode networks. In order to obtain information of places interdependence, we projected these networks from linking individuals to the leisure places they visited, into weighted networks where leisure places are linked to each other and the link's weight is the number of people that have visited both places. These initial weights measure the strength of the relation between certain places. However, in order to account for the role of proximity in determining people movements we also include a distance decay function: $P(d) \sim d^{-n}$ which measures the likelihood of a connection between two nodes and weight the network accordingly, where d is the Euclidean distance between nodes and n would depend on the size and compactness of the network.

Functional Neighbourhoods and Spatio-Temporal Stability Analysis: To identify how New York City is structured based on people movements at dif-

³ Differently from [12, 13] we excluded a time interval within the working day - from 12:00 to 14:00 - in the workplaces classification to identify where possible non-working activities, i.e. lunch break, might be happening in that time interval which could be better classified as leisure places.

ferent times of the day, we implemented a spatially-weighted community detection algorithm [14] on the networks capturing relations between places. The best partitioning of network nodes with a community detection approach is obtained by maximising modularity [15, 16]. To do this on spatially-weighted networks, two parameters can be tuned: (1) the distance decay function exponent n – employed to weight edges, and (2) the resolution parameter of the Louvain community detection algorithm. The resolution determines the communities’ scale, whereby increasing the resolution produces a larger number of smaller communities [17]. The resulting communities can be ascribed as functional neighbourhoods of New York city, insofar functional refers to the interdependence between places derived from people movements [10, 18]. As a result we have 5 different outputs capturing city’s neighbourhoods based on mobility patterns emerging from: the whole dataset, which we refer as the general neighbourhoods, and movements happening in the morning, midday, afternoon and evening. We then developed metrics to establish the extent to which these neighbourhoods change in shape and size across the day in comparison with the general, which is used as the baseline. Specifically, two measures are computed to quantify: 1) how much of the general neighbourhoods has remained stable (stability); and, 2) how much of the general neighbourhoods have been split into different partitions (change).

4 Results

Table 1 and Figure 1 reveal the functional neighbourhoods identified for the five boroughs of NYC. Based on the full data, the general signature comprises 38 distinctive neighbourhoods. Those identified in the boroughs of Staten Island, Manhattan and Bronx are largely geographically contained in their administrative boundaries. This may reflect the way geographical features, such as the East River, contribute to shape human mobility in these boroughs. These patterns contrast with those observed in boroughs of Queens and Brooklyn where communities expand across administrative boundaries.

The general signature, however, masks notable variations in the size, number and stability of communities over the course of a day. The results reveal a greater number of micro communities suggesting smaller spatial living containers, with people moving over relatively shorter geographic distances. Midday displays the largest number of, and smallest communities, perhaps reflecting an increase in the intensity of population movement and diversity in the purpose of those movements for lunch, grocery shopping, school pickups and drops within a small spatial radius. In fact, many of these small communities represent a single hexagon cell suggesting spatial concentration of people around particular venues or points of interest.

We also observe physical spaces not being part of any community, predominantly on the north west of Staten Island, north of Bronx, and coastal areas of Brooklyn and Queens. These spaces correspond to non-residential spaces, such as leisure parks, industrial parks, terminals, storage centres, wetlands and highways, with relatively less visits as temperatures dropped from summer time

Table 1: Functional neighbourhoods (or Communities) identified across the entire dataset are referred to as General and are used as the baseline to assess both the stability and change of Morning, Midday, Afternoon and Evening neighbourhoods. Area is the average area in km squared of the detected communities. Stability is described by: the average H3 cells’ percentage that remain within the same community as the General, its standard deviation and average area in km squared. Change is described by: the average H3 cells’ percentage that has changed community with respect to the General, its standard deviation and average area in km squared.

	Number of Communities	Area (km2)	Stability (%)	Stability (Std Dev)	Stable Area (km2)	Change (%)	Change (Std Dev)	Change Area (km2)
General	38	21.15	-	-	-	-	-	-
Morning	62	10.96	81.6	15.6	1,415	5.75	7.33	94.8
Midday	87	7.73	71.5	20.0	1,232	5.61	7.77	76.3
Afternoon	62	11.33	81.2	17.6	1,381	5.55	7.13	88.9
Evening	72	9.67	74.1	19.4	1,279	6.13	5.48	90.0

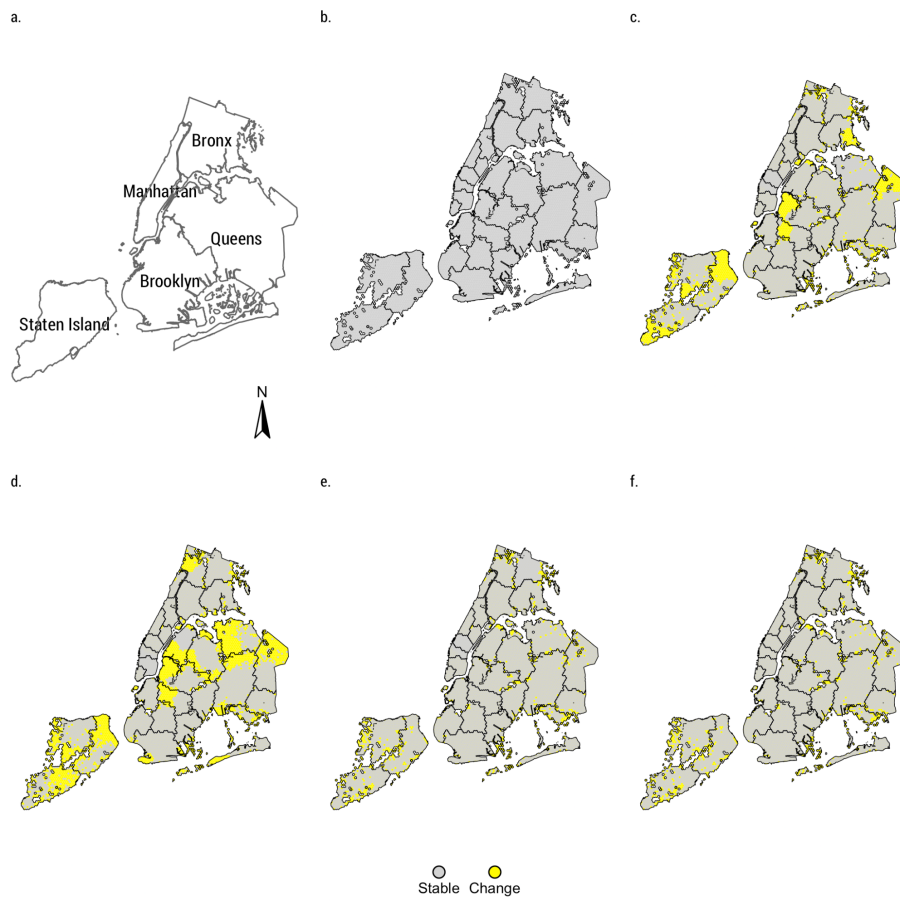
during October. While communities tend to be remarkably stable, they also contract and expand over the course of the day. For instance, 1 indicates that 81.6% of all hexagons in the morning are part of the same communities captured in the general signature. Yet, new communities emerge particularly across Bronx, Queens, Brooklyn and Staten Island. Communities in Manhattan remain largely unchanged representing well established living spaces. The largest extent of change occurs in the evening. Yet, it is not the largest in terms of area cover as many of the communities during this time of the day comprise a very small number of hexagons or a single cell suggesting more localised movements overall. The largest change in geographic area cover occurs in the morning as mobility patterns reflect activities which might happen at farther distances (i.e. with movements being first around home and then around workplaces).

5 Discussion

Human mobility is suggested to follow a predictable functional shape [19] and form spatial containers of regular sizes, representing the geographical scales of human mobility [6]. Yet, such containers neglect temporal variations in visitation patterns, exploration and structural connectedness of geographical spaces through human mobility. We used a dataset of ~ 1.19 billion mobile phone GPS traces to analyse the extent of temporal variations in daily human mobility patterns across NYC.

We found evidence of remarkable stability in the number of spatial containers, but we also observed a considerable degree of change in the number and size of these containers throughout the course of a typical day. We presented evidence of smaller spatial containers compared to those that can be captured if the dimension of time is ignored. We find a remarkable pattern of stability in community formation, reflecting that individuals tend to routinely move within the same geographical spaces. This finding is consistent with prior work on the predictability of daily human mobility as individual mobility trajectories tend to return to a few highly frequented locations and thus show a high degree of

Fig. 1: Maps of the communities at the general level and across day profiles. The map shows day profile community changes (yellow area) with respect to the general communities (boundaries). Map of the New York City boroughs (a), map of the general communities (b), morning communities (c), midday communities (d), afternoon communities (e) and evening communities (f).



temporal and spatial regularity in their mobility signatures [19]. This may signal to the perdurable potential impact of policy interventions influencing the ways in which people move and use local spaces.

Our findings suggest that functional neighbourhoods comprise very well defined small geographic containers of varying size during the day. While people seem to use the same physical spaces throughout the day, others modify their mobility patterns to meet their daily needs, including shopping and school pick ups/drop offs, giving rise to the formation of micro functional neighbourhood spaces. These neighbourhoods emerge from the connection between leisure places, reflecting the availability and accessibility of local urban features [20]. These attributes thus shape the extent to which areas promote localised mobility (smaller communities) or travel over long distances, and the ways in which these patterns of mobility remain consistent over time.

Additionally, our findings contribute to understanding how human mobility and the functional use of space change over short-time intervals while traditional data, such as censuses and surveys, have enabled us to gain an understanding of the ways functional spaces change over long windows of time. So we have developed a comprehensive knowledge of structural changes in urban land use, population restructuring, city decay and economic processes, such as de-industrialisation (e.g. [21–23]).

Our findings based on new forms of data from mobile phone applications have allowed us to understand how human mobility defines small functional neighbourhood containers over short-term temporal scales. Such knowledge is key to understand urban dynamics, and further research in this direction, has the potential improve our understanding of the resilience of cities to fast-paced changes, such as pandemics and extreme weather events, and hence, improve preparedness and planning in urban spaces.

6 Conclusions

This work leverages on the use of mobility data at high spatio-temporal resolution to detect short-term structural changes in cities. Our study opens up a number of avenues for future research. Human mobility is known to differ across population subgroups, particularly across socio-economics, age and gender [20]. Forthcoming analysis will explore how functional neighbourhoods change across population subgroups, investigating how differences in accessibility to employment, education and services amenities promote or deter long-distance mobility. Additionally, future research is required to assess representativeness biases in mobile phone GPS data [24]. We could not assess associated biases as we could not access information on the mobile phone applications and providers for privacy and confidentiality reasons. More broadly, we lack a standard framework to assess and validate biases in mobile phone GPS data. Work is needed in this area.

References

1. Miller, Harvey J. "A measurement theory for time geography." *Geographical analysis* 37.1, 17-45 (2005).
2. Miller, Harvey J. "Time geography and space-time prism." *International encyclopedia of geography: People, the earth, environment and technology*, 1-19 (2017).
3. Miller, Harvey J. "Time geography." *Handbook of Behavioral and Cognitive Geography*. Edward Elgar Publishing, (2018).
4. Kwan, Mei-Po. "Time, information technologies, and the geographies of everyday life." *Urban geography* 23.5, 471-482 (2002).
5. Green, M., Pollock, F.D. and Rowe, F., 2021. New Forms of Data and New Forms of Opportunities to Monitor and Tackle a Pandemic. In *COVID-19 and Similar Futures* (pp. 423-429). Springer, Cham.
6. Alessandretti, Laura, Ulf Aslak, and Sune Lehmann. "The scales of human mobility." *Nature* 587.7834, 402-407 (2020).
7. Patterson, Zachary, and Steven Farber. "Potential path areas and activity spaces in application: A review." *Transport Reviews* 35.6, 679-700 (2015).
8. McKenzie, Grant, Krzysztof Janowicz, Song Gao, and Li Gong. "How where is when? On the regional variability and resolution of geosocial temporal signatures for points of interest." *Computers, Environment and Urban Systems* 54, 336-346 (2015).
9. McKenzie, Grant, and Krzysztof Janowicz. "Where is also about time: A location-distortion model to improve reverse geocoding using behavior-driven temporal semantic signatures." *Computers, Environment and Urban Systems* 54, 1-13 (2015)
10. Calafiore, Alessia, Gregory Palmer, Sam Comber, Daniel Arribas-Bel, and Alex Singleton. "A geographic data science framework for the functional and contextual analysis of human dynamics within global cities." *Computers, Environment and Urban Systems* 85, 101539 (2021).
11. Uber (2018) H3: Uber's Hexagonal Hierarchical Spatial Index. Available at: <https://eng.uber.com/h3>
12. Jundee, Thanisorn, et al. "Inferring commuting flows using CDR data: A case study of Lisbon, Portugal." *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers*. (2018).
13. Phithakkitnukoon, Santi, et al. "Inferring social influence in transport mode choice using mobile phone data." *EPJ Data Science* 6 (2017): 1-29
14. Chen, Yu, Jun Xu, and Minzheng Xu. "Finding community structure in spatially constrained complex networks." *International Journal of Geographical Information Science* 29, no. 6, 889-911 (2015).
15. Newman, Mark EJ, and Michelle Girvan. "Finding and evaluating community structure in networks." *Physical review E* 69.2, 026113 (2004).
16. Lancichinetti, Andrea, and Santo Fortunato. "Community detection algorithms: a comparative analysis." *Physical review E* 80.5, 056117 (2009).
17. Lambiotte, Renaud, J-C. Delvenne, and Mauricio Barahona. "Laplacian dynamics and multiscale modular structure in networks." *Physical Review* (2008).
18. Dunford, Michael. "Regional development models." *International Encyclopedia of Geography: People, the Earth, Environment and Technology: People, the Earth, Environment and Technology*, 1-15 (2016).
19. Gonzalez, Marta C., Cesar A. Hidalgo, and Albert-Laszlo Barabasi. "Understanding individual human mobility patterns." *nature* 453.7196, 779-782 (2008).

20. Graells-Garrido, Eduardo, et al. "A city of cities: Measuring how 15-minutes urban accessibility shapes human mobility in Barcelona." *PloS one* 16.5, e0250080 (2021).
21. Hamnett, Chris. "Gentrification and the middle-class remaking of inner London, 1961-2001." *Urban studies* 40.12, 2401-2426 (2003).
22. Andersen, Hans Skifter. *Urban sores: On the interaction between segregation, urban decay and deprived neighbourhoods*. Routledge (2019).
23. Patias, Nikos, Francisco Rowe, and Dani Arribas-Bel. "Trajectories of neighbourhood inequality in Britain: Unpacking inter-regional socioeconomic imbalances, 1971-2011." *The Geographical Journal* (2021)
24. Rowe, F. *Big Data and Human Geography*. (pre-print) In: Demeritt, D. and Lees L. (eds) *Concise Encyclopedia of Human Geography*. Edward Elgar Encyclopedias in the Social Sciences series (2021).