The exploration of human activity zones using geo-tagged big data during the COVID-19 first lockdown in London, UK

Tongxin Chen *[1], Tao Cheng†[1] and Di Zhu‡[2]

[1]SpaceTimeLab for Big Data Analytics, Department of Civil, Environmental and Geomatic Engineering, University College London, London WC1E 6BT, UK
[2]Department of Geography, Environment and Society, University of Minnesota, Twin Cities, MN 55455, USA

March 27, 2021

Summary

Exploring the human activity zones (HAZs) gives significant insights into understanding the complex urban environment and reinforcing urban management and planning. Though previous studies have reported the significant human activity shifting at the city-level in global metropolises due to COVID-19 containment policies, the dynamic of human activity across urban areas at space and time during such an ever-changing socioeconomic period has not been examined and discussed hitherto. In this study, we proposed an analysis framework to explore the human activities zones using geo-tagged big data in London, UK. We first utilised the activity-detection method to extract visits/stops at space and time as the human activity metric from the mobile phone GPS trajectory data. Then, we characterised HAZs based on the homogeneity of hourly human activity footfalls on the middle layer super output areas (MSOAs). The results show the HAZs not only exhibit declines in human activity but are strongly associated with urban land-use and population variables during the COVID-19 pandemic.

KEYWORDS: Urban functions, human activity, social sensing, geo-tagged big data, COVID-19.

*tongxin.chen.18@ucl.ac.uk
†tao.cheng@ucl.ac.uk
‡dizhu@umn.edu

1 Introduction

As COVID-19 continues to spread around the global cities, urban citizen lives have been changed significantly due to the pandemic containment policy, such as the national lockdown order. Recent studies have addressed human mobility/activity pattern shifting through the analysis of city or regional-level migration data sets (Chinazzi et al., 2020; Kraemer et al., 2020; Sirkeci and Yucesahin, 2020; Iacus et al., 2020; Hadjidemetriou et al., 2020). However, previous studies that exploring and interpreting human mobility pattern change concerning large geospatial district cannot disentangle the complex dynamics of urban areas in the context of social-economic environment change. As the massive amounts of geo-tagged big data (e.g., mobile phone GPS data, WiFi data and social media data) can be collected efficiently with the rapid development of ubiquitous location awareness technologies, large volumes of human activity data regarding the spatio-temporal footprints tied with places, neighbourhoods and urban areas can provide the data-driven perspective to reveal the urban complex dynamics (Hasan et al., 2013; Ahas et al., 2015; Liu et al., 2015). Here, our intention is to explore the human activity zones characterised by temporal human activity patterns at different urban geospatial areas during the COVID-19 pandemic. In this study, we utilised raw mobile phone GPS data to extract human activity metric and aggregate the hourly human activity pattern at MSOAs, then portray the human activity zones (HAZs) based on the similarity of human activity pattern in London metropolis, UK.

2 Methodology

2.1 Human activity detection from mobile phone GPS data

As human activities exhibit regularity of patterns at space and time, it is the acknowledgement that footfall as a metric of human activities can be extracted from mobile phone GPS trajectory data (Sevtsuk and Ratti, 2010; Tu et al., 2018). Then, potential human activity can be described as a single user spending some time in one place, i.e., the consecutive records of user are at the same location during a time period (Zheng, 2015; Tu et al., 2017). Here, we use the stay detection algorithm proposed by Hariharan and Toyama (2004) and Pappalardo et al. (2019), which needs two preset parameters for trajectory data, i.e., Δd (the maximum distance that records of a user move around from a point location to count as a stay/stop) and Δt (the minimum duration time period that the records stay within time distance to qualify as a stay/stop at the location). After the implementation of stay detection for every GPS trajectory in the data sets, we aggregate the counts of stay as the proxy of human activity metric to defined geospatial units (e.g., census block, community) and temporal units (e.g., hourly, daily), respectively.

2.2 Function zoning based on dynamic human activity

In parallel to the urban functional zone that urban space with specific functions for human activities (Tu et al., 2018), human activities zone (HAZ) refers to a set of geospatial unit clusters exhibiting a certain similarity characterized by the human activities. In this study, our functional zoning is to extract N HAZs based on the human activity temporal pattern data from the M geospatial units ($N < M$). We employ the clustering strategy to achieve the functional zoning of HAZs: (1) using silhouette coefficient (the best value is 1 and the worst value is -1) as a clustering evaluation metric for chosen linkage criterion (Ward's criterion) to determine the numbers of clusters (numbers of types of HAZs); (2) utilising agglomerative clustering algorithm [1] to generating the N HAZs based on the human activity temporal pattern from M geospatial units.

3 Case study

3.1 Data and study area

With COVID-19 spreading in global cities, London metropolis continues to undergo the diffusion of the viruses as the city of the highest numbers of confirmed cases in the UK. As an emergency response to pandemic, the first national lockdown announced by the government started Mar. 24, 2020 following a series of restricted measures in the city. In this study, we focus on the urban area of London comprising 983 MSOAs where the hourly human activity patterns generate. In addition, the human activity metric in terms of the footfalls (stays/stops) are calculated based on the millions of anonymous raw mobile phone GPS data provided by Location Sciences. In this study, stay/stop in the stay is defined as a user spend at least 5 mins within a distance of 50m spatial radius from a given GPS trajectory points. Then, the detected footfalls numbers are aggregated to the hourly temporal unit and MSOA spatial unit, respectively. Here, our interest observation time period is four weeks around the first national lockdown (Mar. 24, 2020) in London, i.e., from Mar. 9, 2020 to Apr. 6, 2020. So, the generation of HAZs is employed on the 696 hourly human footfall records on every MSOA in London ($696*983$ in total).

3.2 Human activity zones in London during pandemic

Figure 1 denotes the decline trends of daily human activities (detected footfall numbers) after the first national lockdown (Mar. 23, 2020) in London. It is obvious that the effect of stay-at-home order has enforced a significant effect on the human routine activities in the metropolis.

For ready the clustering step, we use silhouette coefficients as the optimization of cluster numbers from the hourly footfall number records (the numbers of footfalls are standardised), Figure 2 denotes 7 groups of clusters (i.e., 7 types of HAZs) is the optimal for the dataset based on the highest index values.

Then, according to the results generated by agglomerative clustering algorithm, the clustered hourly temporal footfall patterns as 7 different groups are illustrated in Figure 3. It is interesting to find that the significant differences of footfalls tend to decline (i.e., tend to be a similar pattern)across seven groups from before-lockdown to after-lockdown period. One alternative explanation is that the stay-at-home order has normally shaped the regularity of human activity in most urban areas after lockdown which exhibits various human routines before lockdown due to distinctive urban function.

Figure 4 denotes the spatial distribution of HAZs in urban areas of London. The HAZs are visualised as their urban functional types according to their temporal human activity patterns. Specifically,

[1] https://scikit-learn.org/stable/modules/clustering.html#hierarchical-clustering

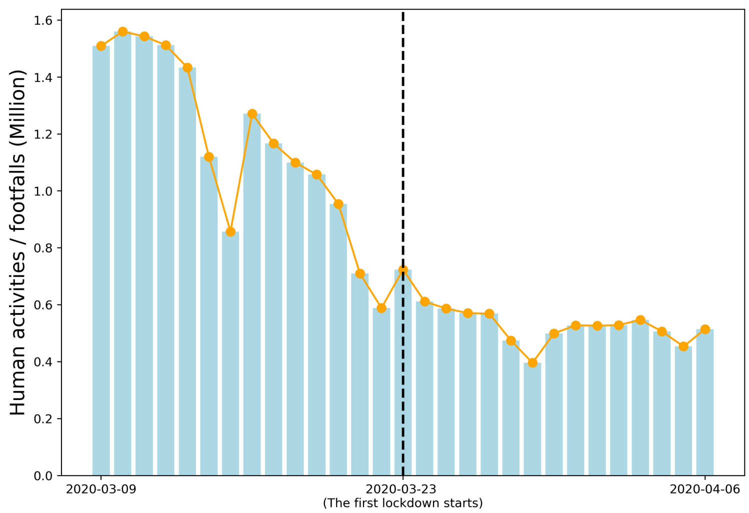

Figure 1: Daily human activities (footfalls) in London from Mar. 9, 2020 to Apr. 6, 2020.

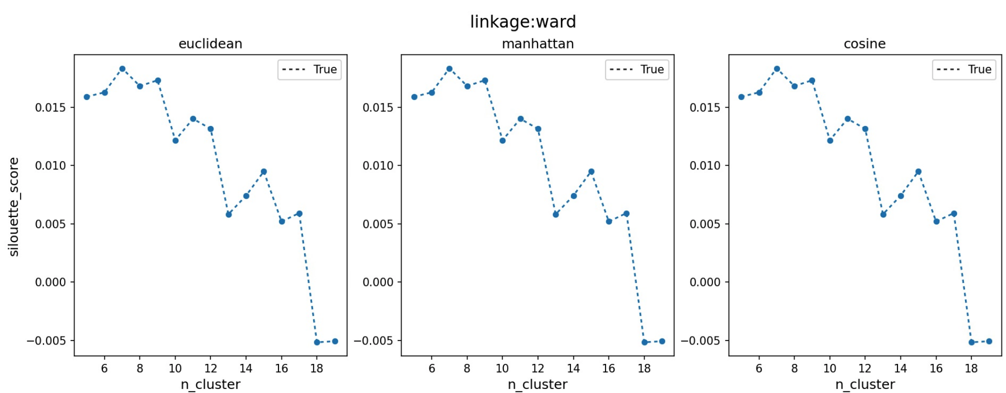

Figure 2: The silhouette coefficients measured by three different distances as the evaluation of optimal cluster numbers.

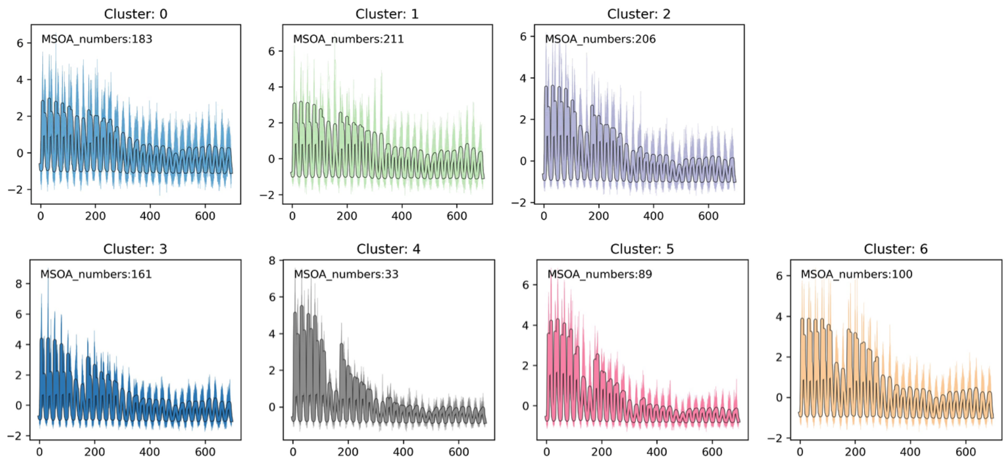

Figure 3: Hourly human activity patterns in seven types of HAZs in London.

the HAZs 5 (red areas) consists of 89 MSOAs mainly concentrated on the urban central areas exhibiting amounts of human footfalls, such as City of London. On the contrary, the HAZs 1 comprising 211 MSOAs (green areas) are distributed on the external areas of London, such as Richmond Park.

4 Discussion and conclusions

Concerning the population density as an explanation of the HAZs distribution in our study area, we employ the census data from London Datastore [2] at MSOA-level to examine such demographic variable's correlation with our generated HAZs. Figure 5 reveals that the distribution of population density in different types of HAZs. Generally, the box-plot shows the population densities of HAZs 1, HAZs 3 and HAZs 6 are less than HAZs 0, HAZs 2, HAZs 4 and HAZs 5. Statistically, population density has significant differences in the seven groups of HAZs based on a non-parametric test called Kruskal-Wallis test ($p < 0.0001$).

Following our proposed analysis framework, human footfalls representing human activity metric aggregated from stays/stops from raw mobile GPS data using stop detection algorithm, and HAZs can be efficiently generated based upon the MSOA using an agglomerative clustering algorithm. The results facilitate our understanding of how human activity patterns could be impacted by containment policies, and thus human routine patterns are associated with demographic variables differently across the urban areas in London. As the complexity of the urban system, our further work will focus on using multi-variables to interpret the spatial heterogeneity of human activity dynamics during the COVID-19 pandemic.

[2] https://data.london.gov.uk/dataset/super-output-area-population-lsoa-msoa-london

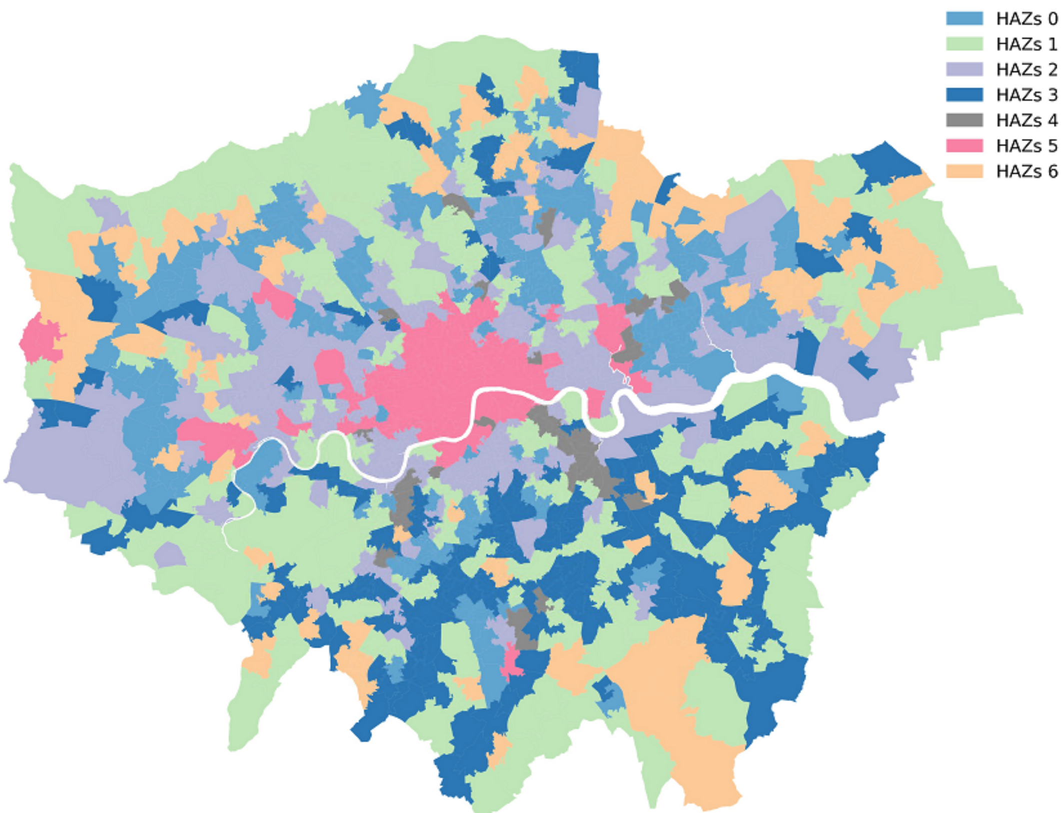

Figure 4: The spatial distribution of HAZs in London.

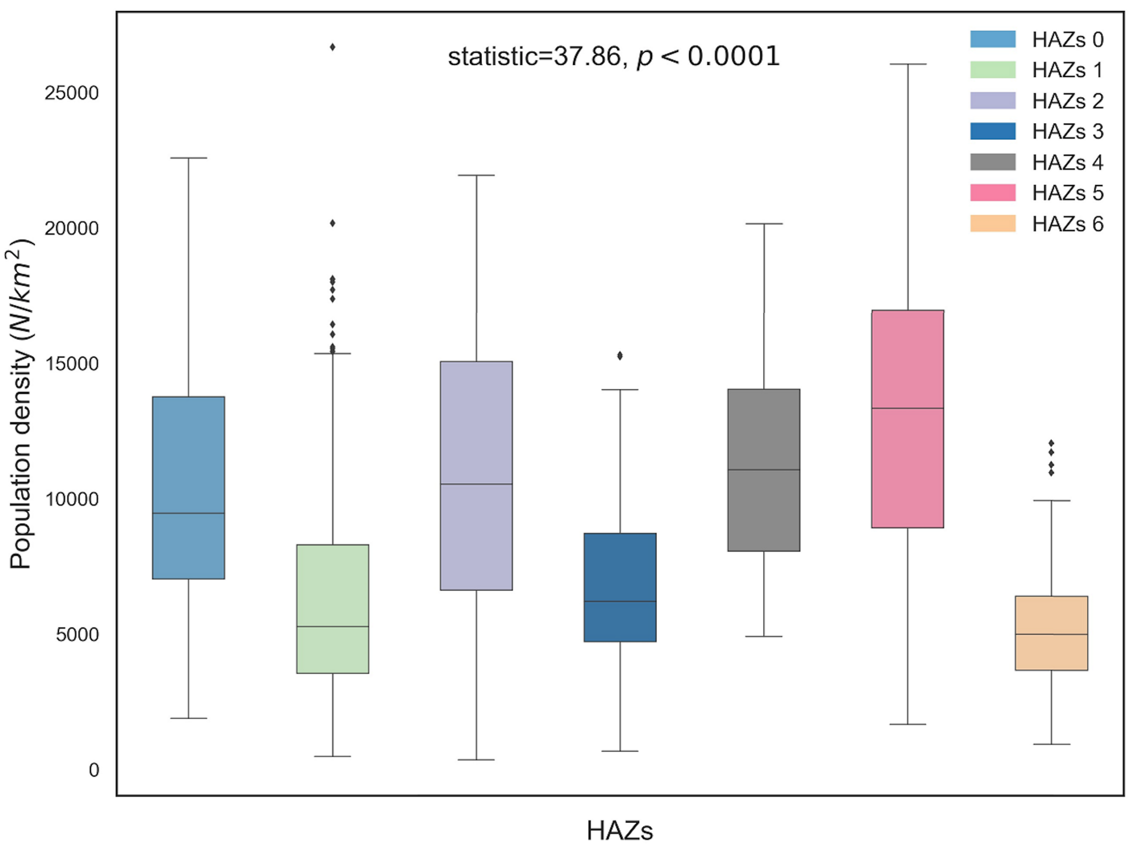

Figure 5: The population density in different types of HAZs in London.

5 Acknowledgements

This work is supported by the UCL Grand Challenges 2020-21. We thank the Location sciences for their providing data.

References

Ahas, R., Aasa, A., Yuan, Y., Raubal, M., Smoreda, Z., Liu, Y., Ziemlicki, C., Tiru, M., and Zook, M. (2015). Everyday space–time geographies: using mobile phone-based sensor data to monitor urban activity in harbin, paris, and tallinn. *International Journal of Geographical Information Science*, 29(11):2017–2039.

Chinazzi, M., Davis, J. T., Ajelli, M., Gioannini, C., Litvinova, M., Merler, S., y Piontti, A. P., Mu, K., Rossi, L., Sun, K., et al. (2020). The effect of travel restrictions on the spread of the 2019 novel coronavirus (covid-19) outbreak. *Science*, 368(6489):395–400.

Hadjidemetriou, G. M., Sasidharan, M., Kouyialis, G., and Parlikad, A. K. (2020). The impact of government measures and human mobility trend on covid-19 related deaths in the uk. *Transportation research interdisciplinary perspectives*, 6:100167.

Hariharan, R. and Toyama, K. (2004). Project lachesis: parsing and modeling location histories. In *International Conference on Geographic Information Science*, pages 106–124. Springer.

Hasan, S., Schneider, C. M., Ukkusuri, S. V., and González, M. C. (2013). Spatiotemporal patterns of urban human mobility. *Journal of Statistical Physics*, 151(1):304–318.

Iacus, S. M., Santamaria, C., Sermi, F., Spyratos, S., Tarchi, D., and Vespe, M. (2020). Human mobility and covid-19 initial dynamics. *Nonlinear Dynamics*, 101(3):1901–1919.

Kraemer, M. U., Yang, C.-H., Gutierrez, B., Wu, C.-H., Klein, B., Pigott, D. M., Du Plessis, L., Faria, N. R., Li, R., Hanage, W. P., et al. (2020). The effect of human mobility and control measures on the covid-19 epidemic in china. *Science*, 368(6490):493–497.

Liu, Y., Liu, X., Gao, S., Gong, L., Kang, C., Zhi, Y., Chi, G., and Shi, L. (2015). Social sensing: A new approach to understanding our socioeconomic environments. *Annals of the Association of American Geographers*, 105(3):512–530.

Pappalardo, L., Simini, F., Barlacchi, G., and Pellungrini, R. (2019). scikit-mobility: a python library for the analysis, generation and risk assessment of mobility data.

Sevtsuk, A. and Ratti, C. (2010). Does urban mobility have a daily routine? learning from the aggregate data of mobile networks. *Journal of Urban Technology*, 17(1):41–60.

Sirkeci, I. and Yucesahin, M. M. (2020). Coronavirus and migration: Analysis of human mobility and the spread of covid-19. *Migration Letters*, 17(2):379–398.

Tu, W., Cao, J., Yue, Y., Shaw, S.-L., Zhou, M., Wang, Z., Chang, X., Xu, Y., and Li, Q. (2017). Coupling mobile phone and social media data: A new approach to understanding urban functions and diurnal patterns. *International Journal of Geographical Information Science*, 31(12):2331–2358.

Tu, W., Hu, Z., Li, L., Cao, J., Jiang, J., Li, Q., and Li, Q. (2018). Portraying urban functional zones by coupling remote sensing imagery and human sensing data. *Remote Sensing*, 10(1):141.

Zheng, Y. (2015). Trajectory data mining: an overview. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 6(3):1–41.

Biographies

Tongxin Chen, is a PhD student at SpaceTimeLab, Department of Civil, Environmental and Geomatic Engineering, University College London. His research interests include urban analytics, spatio-temporal data analysis, crime mapping and applied machine learning.

Tao Cheng, is a professor in Geoinformatics at Department of Civil, Environmental and Geomatic Engineering, University College London. She is the Founder and Director of SpaceTimeLab for Big Data Analytics. Her research interests span network complexity, Geocomputation, space-time analytics and Big data mining (modelling, prediction, clustering, visualisation and simulation) with applications in transport, crime, business, health, social media, and natural hazards.

Di Zhu, is an assistant professor at Department of Geography, Environment and Society, University of Minnesota, Twin Cities. His research interests include exploring theories, algorithms and models to bridge geospatial analytics, artificial intelligence and social sensing. He uses integrated thinking and cross-disciplinary methods to facilitate intelligent spatial understanding and geographic knowledge discovery, focusing on the human-environment complexities within socioeconomics, population, business and public health.