Urban Zoning Using Intraday Mobile Phone-Based Commuter Patterns in the City of Brno

Yuri Bogomolov[1(✉)], Alexander Belyi[1], Ondřej Mikeš[2], and Stanislav Sobolevsky[1,3,4]

[1] Department of Mathematics and Statistics, Faculty of Science, Masaryk University, Kotlarska 2, 611 37 Brno, Czech Republic
bogomoloviura@gmail.com , ss9872@nyu.edu

[2] RECETOX, Masaryk University, Kamenice 753/5, 625 00 Brno, Czech Republic

[3] Center for Urban Science and Progress, New York University, 370 Jay Street, Brooklyn, NY 11201, USA

[4] Institute of Law and Technology, Faculty of Law, Masaryk University, Veveri 70, 611 80 Brno, Czech Republic

Abstract. The emergence of information and mobile technology has had a profound impact on modern life, altering the manner in which individuals communicate, access information, entertain themselves, and conduct business. This development has created new opportunities for researchers to access datasets that were not accessible to prior generations of scholars. Historically, studies of urban commuting have relied on census data, which portrays commute patterns as a static number that are updated every few years. However, over the last two decades, the advent of mobile phone datasets has facilitated new research avenues. This study employs mobile phone mobility data to define a signature of urban districts within the Czech city of Brno, utilizing it for the purpose of urban zoning. The proposed signatures provide a tangible classification of neighborhoods with potential applications in urban and transportation planning. This approach is demonstrated using mobile data of 13 thousand inhabitants of Brno and can be applied to other cities wherever similar data is available.

Keywords: Intraday commute patterns · Urban commute · City delineation · Urban zoning · Mobile phone mobility data

1 Introduction

Understanding commute patterns is crucial for city planning, as it can inform decisions on transportation infrastructure, public transport services, and urban development. Existing models are used to predict commute flows based on the spatial structure of cities [1].

The four-step transportation model [7] is a widely used approach for analyzing and forecasting transportation patterns in urban areas. The model consists of four

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023
O. Gervasi et al. (Eds.): ICCSA 2023, LNCS 13957, pp. 482–490, 2023.
https://doi.org/10.1007/978-3-031-36808-0_35

sequential steps: trip generation, trip distribution, mode choice, and trip assignment. After trips are generated based on the population and household data the model matches trip origins with destinations based on commute patterns.

The gravity model [21] is often used to model commute patterns because it provides a simple yet effective way to estimate the volume of traffic between two locations based on distance and attractiveness. In the context of commuting, the gravity model assumes that the volume of traffic (T_{ij}) between two locations (i and j) is proportional to the product of their attractiveness (P_i, P_j) and inversely proportional to some increasing function $f(d)$ of the trip distance or travel time:

$$T_{ij} \propto \frac{P_i P_j}{f(d)} \tag{1}$$

In 2012, Simini et al. proposed the radiation model [11], which accounts attractiveness of neighboring nodes to model commute flows between locations i and j. But the commute is still considered a static flow by the model. In 2021, Schlapfer et al. proposed the universal visitation law of human mobility [10] to describe visit frequencies. However, none of these models cover the time component of the commute, which is crucial to address traffic congestion during peak hours. While existing models focus on the commute flow and frequency, understanding intraday commute patterns is important for city planning for several reasons:

Infrastructure Planning: Intraday commute patterns help city planners to identify the areas where traffic congestion is most common and determine the necessary infrastructure improvements to alleviate traffic flow. For instance, if there is heavy traffic congestion during the morning rush hour, city planners can determine if it is caused by too few lanes, too few exits, or a bottleneck that requires additional infrastructure.

Public Transport Planning: Knowing intraday commute patterns is crucial for planning and improving public transport services. City planners can identify areas with high passenger demand and optimize routes and schedules accordingly. This can help reduce waiting times and improve passenger comfort and convenience, encouraging more people to use public transportation.

Parking Management: Understanding intraday commute patterns helps city planners to optimize parking facilities in areas with high traffic volumes. By analyzing parking utilization patterns, city planners can determine the optimal number of parking spaces required and adjust the pricing structure to encourage carpooling, cycling, or using public transportation.

Environmental Impact: Intraday commute patterns can provide insights into the environmental impact of transportation on a city. By identifying areas with high traffic congestion, city planners can implement measures such as low-emission zones, carpooling or bike-sharing programs, and promote sustainable transport options. For example, the negative health impact of congested traffics flows could be reduced if the peak traffic hours do not overlap with the drop-off and pick-up time in neighboring schools.

The lack of hourly datasets limits the research of intraday commute patterns. Previous attempts covered only narrow mobility flows like bicycle commuting in Melbourne [12] or taxi trips in New York City [4]. The mobile phone data was successfully used to study the city structure based on the population density [15], mobility patterns [19,20] and their relation to socioeconomic status [17,18], urban park catchment areas [5], and changes in the average distance between individuals [6].

Defining city and district boundaries is known to be a challenge [2,3]. Previous attempts to define district signatures and use them for district clustering were based on the municipal request datasets [14], financial activities [13], and mobile call records [8,9,16]. However, all these events are generated only in case of particular activities and therefore have limited coverage. Also, the resulting clusters did not reflect the mobility and transportation needs of inhabitants.

In this paper, we propose to use the mobile origin-destination commute data to build district mobility signatures based on the hourly commute patterns and then use these signatures to group districts into areas with similar temporal patterns. Mobile flow data can be collected for any modern city, which makes the proposed approach generally applicable, while the district signatures provide insights into mobility patterns and transportation needs.

The following sections of the paper describe the input dataset, data processing, and aggregation steps. Then we share our findings and the follow-up experiments. Finally, we discuss the results and provide a conclusion, including recommendations on how our findings may be used for urban and transportation planning.

2 Materials and Methods

2.1 Data

Our research is based on the origin-destination commute dataset provided by a mobile phone company. The dataset provides hourly origin-destination flows for 48 districts of Brno, the second-largest city in the Czech Republic, with a population of about 380,000 inhabitants. We had access to aggregated commute data that covered one full week in October 2019.

There are two types of mobile phone datasets: active and passive. Active datasets are based on records of certain actions, like phone calls or text messages. Every phone communicates with the mobile phone network at least every 30 min, and these communications allow for capturing passive datasets. Active datasets are very useful for studying the interaction graph between people, while passive datasets provide better coverage to study mobility or population density. Our research is based on the passive mobile phone dataset.

In modern cities, people tend to carry their mobile phones for the whole day, and the dataset above provides unique insights into general mobility flows. We aggregated all outgoing trips for all districts for a given hour. As a result, every Brno district was represented by a vector with 168 (7 days $\times$ 24 h) outflow numbers. Due to the differences in population between districts, we normalized

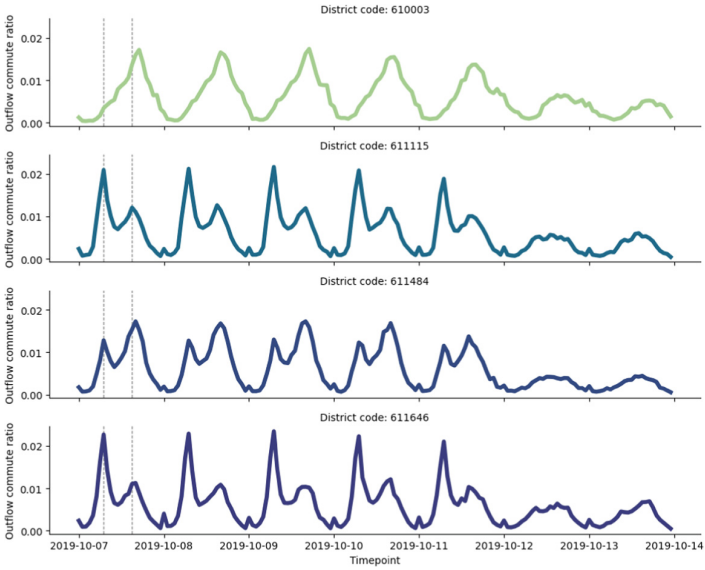


Fig. 1. Hourly commute outflow ratios for four Brno districts. The first five days of data represent weekdays, while the last two days represent weekends. The vertical lines highlight commute spikes, corresponding to 8 am and 4 pm on weekdays.

commute vectors by the total amount of outflow. Therefore every vector component represented what ratio of total outflow was observed in a given hour.

The normalized vectors had clear weekday and weekend patterns (see Fig. 1). Some timelines had a surprising outflow spike around midnight, but after further investigation, this behavior was attributed as an artifact of data collection: the input dataset was provided in daily files, and any trip that spanned across midnight was recorded as a trip that ended at midnight.

We noticed a few repeating patterns across city districts. Some districts had very similar outflow vectors that followed the same pattern (see the second and the fourth plot in Fig. 1), while overall, patterns had significant differences. The initial investigation showed some correlation between commute pattern similarity and district location: e.g., the districts in the city center followed a similar pattern. To study the common patterns holistically, we decided to reduce dimensionality and cluster the outflow vectors.

2.2 Principal Component Analysis and Clustering

We used PCA (Principal Component Analysis) to reduce the dimensionality of vectors. Table 1 gives a summary of PCA performance based on the number of components. Using just two components explains more than 73% of cumulative variance, and we observe smaller incremental improvements for each additional component.

Putting two-dimensional vectors for all districts on a single plot helped to identify patterns of city districts. The similarity of commute vectors for related districts gave us the idea to use commute vectors as a district signature.

Table 1. Cumulative variance of 168-component outflow commute vectors based on the number of PCA components.

Number of components	Cumulative variance explained by the components
1	59.34%
2	73.87%
3	78.67%
4	82.81%
5	85.26%

We considered clustering two types of vectors:

1. The original 168-component outflow vectors.
2. 2-component PCA vectors.

Using the original vectors helps to preserve all the information from the vectors. At the same time, the vector components are highly correlated due to the repeating patterns within day boundaries (e.g., mobility tends to reduce during nighttime) and over the week due to the repeating weekday patterns and weekend patterns.

To mitigate the correlation and focus on the orthogonal components, we clustered 2-component PCA feature vectors. Using the elbow method, we decided to move forward with 3 clusters. The clustering output is presented in Fig. 2.

3 Results

While the district signature does not contain any information about the spatial structure, we received a spatially cohesive separation of the city center, residential area, and the mixed zone in between. The outflow profile of the residential cluster has a clear spike in the morning when people go to work. The city center outflow spikes in the evening. And mixed areas in between have traces of both patterns. Also we discovered that the sum of all outflow vectors is consistent with the intraday bicycle volumes in Melbourne [12].

The overall picture matches our intuition: remote city districts followed one pattern, while the city center districts followed a different pattern. However, the resulting clusters have some asymmetry: the northern part of the city has commute patterns similar to the city center. At the same time, the southern and northwestern borders of the city do not belong to the residential cluster.

After additional research, we found that both irregularities match the industrial zones of the city (see Fig. 3). A large number of industry workers commuting

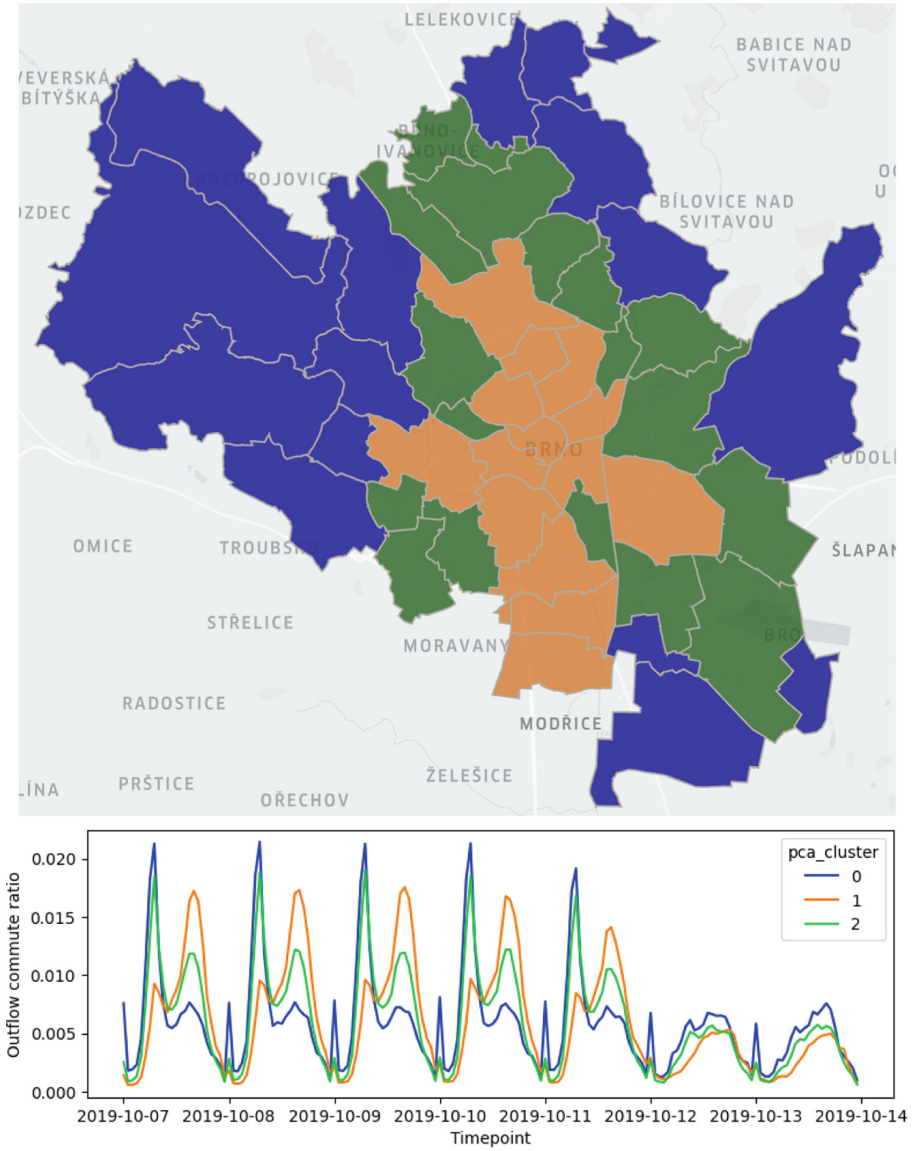


Fig. 2. The top figure demonstrates three city clusters, while the bottom figure shows aggregated hourly outflow commute profile for each cluster. The blue cluster corresponds to residential areas, the orange one covers the city center and business districts, while the green cluster has a combination of both. (Color figure online)

from work increases the afternoon spike, while a limited number of residential properties reduces the commute outflow in the morning.

In addition to the visual inspection of urban zones, we performed a quantitative comparison of clusters based on Brno Census data. We observed a 200%–300%

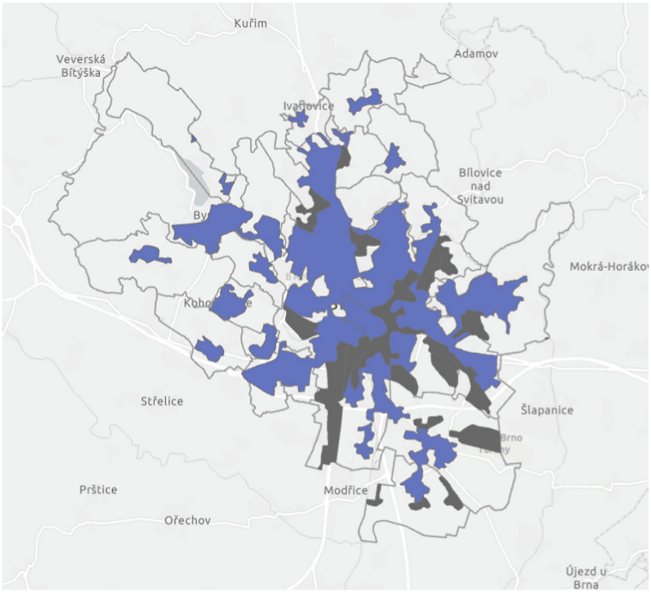
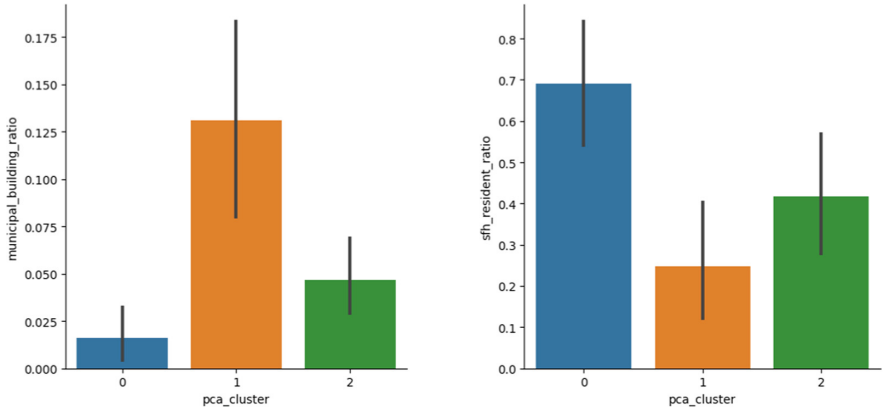


Fig. 3. Brno city zones: urban zones are highlighted in blue, while industrial and commercial zones are highlighted in dark grey. (Color figure online)



(a) Municipal building ratio (b) Single family home resident ratio

Fig. 4. Comparison of the building structure between the identified zones based on commute patterns. The ratio of a) municipal building and b) single-family homes residents among all the buildings/residents within each zone, and 95% confidence intervals showing the variation of those percentages across the locations within the zone

difference in the municipal home ratio (the number of municipal buildings in a cluster divided by the total number of buildings) and the single-family home resident ratio (the proportion of cluster residents living in single-family homes) between the clusters (see Fig. 4). Furthermore, we performed statistical t-tests to compare mean municipal home ratios for all pairs of clusters. The comparison of the residential and mixed clusters yielded a p-value of 3%, while the other two pairs had a p-value below 1%. These results show a statistically significant difference between the clusters.

4 Conclusion and Discussion

To summarize, we identified strong patterns in the urban mobility outflow vectors and proposed using them as a district signature. Characterizing urban locations with their commute timeline allows the delineation of city zones with similar mobility behavior. We found a relationship between the resulting clusters and the ratio of municipal buildings (including libraries, schools, and city administration) and single-family home residents. While this fact is insufficient to establish the causal relation, we plan to conduct further research that will help us better understand the connection between the socioeconomic features of districts and their commute clusters.

Using the mobility signatures to delineate city districts provides great insights into mobility needs and can be directly used for transportation planning and land use classification. At the same time, the relationship between the municipal building and commute patterns contributes new information for city planners.

We proposed an approach leveraging mobility data from mobile phones to define unique signatures of urban locations and apply them to urban zoning. The proposed approach can be applied to any city where a similar input dataset can be collected. Relying only on the mobility data makes the signatures uniform for any country (in contrast to financial or municipal records), while comprehensive dataset coverage enables capturing general commute patterns (counter to taxi or biking datasets).

References

1. Barthélemy, M.: Spatial networks. Phys. Rep. **499**(1–3), 1–101 (2011)
2. Bettencourt, L.M.: The origins of scaling in cities. Science **340**(6139), 1438–1441 (2013)
3. Bettencourt, L.M., Lobo, J., Strumsky, D., West, G.B.: Urban scaling and its deviations: revealing the structure of wealth, innovation and crime across cities. PLoS ONE **5**(11), e13541 (2010)
4. Buchholz, N.: Spatial equilibrium, search frictions and efficient regulation in the taxi industry. Working paper, Technical report (2015)
5. Guan, C., Song, J., Keith, M., Akiyama, Y., Shibasaki, R., Sato, T.: Delineating urban park catchment areas using mobile phone data: a case study of Tokyo. Comput. Environ. Urban Syst. **81**, 101474 (2020)

6. Louail, T., et al.: From mobile phone data to the spatial structure of cities. Sci. Rep. **4**(1), 1–12 (2014)
7. McNally, M.G.: The four-step model. In: Handbook of transport modelling. Emerald Group Publishing Limited (2007)
8. Pei, T., Sobolevsky, S., Ratti, C., Shaw, S.L., Li, T., Zhou, C.: A new insight into land use classification based on aggregated mobile phone data. Int. J. Geogr. Inf. Sci. **28**(9), 1988–2007 (2014)
9. Ratti, C., et al.: Redrawing the map of great Britain from a network of human interactions. PLoS ONE **5**(12), e14248 (2010)
10. Schläpfer, M., et al.: The universal visitation law of human mobility. Nature **593**(7860), 522–527 (2021)
11. Simini, F., González, M.C., Maritan, A., Barabási, A.L.: A universal model for mobility and migration patterns. Nature **484**(7392), 96–100 (2012)
12. Smith, M.S., Kauermann, G.: Bicycle commuting in Melbourne during the 2000s energy crisis: a semiparametric analysis of intraday volumes. Transp. Res. Part B: Methodol. **45**(10), 1846–1862 (2011)
13. Sobolevsky, S., Sitko, I., Des Combes, R.T., Hawelka, B., Arias, J.M., Ratti, C.: Money on the move: big data of bank card transactions as the new proxy for human mobility patterns and regional delineation. the case of residents and foreign visitors in Spain. In: 2014 IEEE International Congress on Big Data, pp. 136–143. IEEE (2014)
14. Wang, L., Qian, C., Kats, P., Kontokosta, C., Sobolevsky, S.: Structure of 311 service requests as a signature of urban location. PLoS ONE **12**(10), e0186314 (2017)
15. Xinyi, N., Liang, D., Xiaodong, S.: Understanding urban spatial structure of shanghai central city based on mobile phone data. China City Plann. Rev. **24**(3) (2015)
16. Xu, Y., Belyi, A., Bojic, I., Ratti, C.: How friends share urban space: an exploratory spatiotemporal analysis using mobile phone data. Trans. GIS **21**(3), 468–487 (2017). https://doi.org/10.1111/tgis.12285
17. Xu, Y., Belyi, A., Bojic, I., Ratti, C.: Human mobility and socioeconomic status: analysis of Singapore and Boston. Comput. Environ. Urban Syst. **72**, 51–67 (2018). https://doi.org/10.1016/j.compenvurbsys.2018.04.001
18. Xu, Y., Belyi, A., Santi, P., Ratti, C.: Quantifying segregation in an integrated urban physical-social space. J. R. Soc. Interface **16**(160), 20190536 (2019). https://doi.org/10.1098/rsif.2019.0536
19. Yu, Q., Li, W., Yang, D., Zhang, H.: Mobile phone data in urban commuting: a network community detection-based framework to unveil the spatial structure of commuting demand. J. Adv. Transp. **2020**, 1–15 (2020)
20. Zhang, B., Zhong, C., Gao, Q., Shabrina, Z., Tu, W.: Delineating urban functional zones using mobile phone data: a case study of cross-boundary integration in Shenzhen-Dongguan-Huizhou area. Comput. Environ. Urban Syst. **98**, 101872 (2022)
21. Zipf, G.K.: The P 1 P 2/D hypothesis: on the intercity movement of persons. Am. Sociol. Rev. **11**(6), 677–686 (1946)