Temporary Migration Flow Inference and Analysis From Perspective of Mobile Phone Network Data

SANTI PHITHAKKITNUKOOK\textsuperscript{1,2}, SORANAN HANKAEW\textsuperscript{1}, MERKEBE GETACHEW DEMISSIE\textsuperscript{3}, ZBIGNIEW SMOREDA\textsuperscript{4}, AND CARLO RATTI\textsuperscript{5}

\textsuperscript{1}Department of Computer Engineering, Chiang Mai University, Chiang Mai 50200, Thailand
\textsuperscript{2}Excellence Center in Infrastructure Technology and Transportation Engineering (ExCITE), Chiang Mai University, Chiang Mai 50200, Thailand
\textsuperscript{3}Department of Civil Engineering, University of Calgary, Calgary, AB T2N 1N4, Canada
\textsuperscript{4}Sociology and Economics of Networks and Services Department, Orange Labs, 92794 Chatillon, France
\textsuperscript{5}SENSEable City Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Corresponding author: Santi Phithakkitnukoon (santi@eng.cmu.ac.th)

This work was supported by the Schulich School of Engineering, University of Calgary Startup Grant.

\section*{ABSTRACT}

Temporary migration trips are circulatory journeys that can be associated with tourism as these trips aren’t made on a regular basis. They’re often left unmeasured due to limited data sources, and thereby missing from regional and national travel demand models. This article presents a methodology for inferring large-scale temporary migration trips from mobile phone network data or Call Detail Record (CDR), and analyzing their spatial determinants based on urban assets derived from Google Places data. As a case study, CDR of mobile phone users in Portugal was used in our study from which insightful statistical patterns of both intra and inter-district flows were observed. Analysis of spatial determinants shows that places for leisure and population density are among the top influential factors that attract inflows into the district, while places for public service such as city hall, local government office, courthouse, and cemetery are negative urban assets for attracting district’s inflows. The presented methodology and insights gained from our analysis can be useful for transport and urban planners to establish better informed travel demand modeling and urban planning strategies.

\section*{INDEX TERMS}

Mobile network data, CDR, temporary migration, spatial determinants, urban asset, opportunistic sensing data.

\section*{I. INTRODUCTION}

Motivated by its countless practical applications, human mobility has been one of the most fascinating research topics over the last decades. Understanding movement of human beings builds a foundation for travel demand forecast as transportation engineers and researchers seek to develop predictive models for how people travel and influential factors that affect their travel choices and decisions.

Transportation infrastructure has been designed and built in relation to the movement of population, which takes place whenever a trip is made. Trips are made regularly and for different purposes. Different trip types vary in their frequency and periodicity [1]. Circulatory trips largely involve regular cyclic schedules, such as commute (daily journey to work or school), weekly travel to more distant work locations, and weekend trip to a second home. Occasionally, circulatory trips can be associated with tourism where trips aren’t made on a regular basis, such as annual holiday travel, business trip, and long-holiday travel, which are potentially long-distance trips that involve temporary change of the place of residence, which can be considered as temporary migration. On the other hand, non-circulatory trips are associated with permanent change of the place of residence i.e., permanent migration, which are mostly long distance trips that are influenced by socio-political, economic, and ecological factors, such as occupation opportunities and family reasons [2]. Due to its characteristics, circulatory trips are generally more predictable than those of non-circulatory trips. Commutes account for the majority of circulatory trips, which are mostly...
short distances and highly predictable [3]. Long-distance trips are largely associated with temporary or permanent migrations.

Long-distance trips are not well represented in most regional and national travel demand models. Although long-distance trips are considered a much smaller portion when compared to all trips generated, they account for a substantial portion of total distance travelled [4], especially on high capacity routes that require considerable budgets to build. So, it is essential that these long-distance trips are also included in the travel demand models. Moreover, with the increasing environmental impact associated with long-distance trips [5], there is a need for appropriate planning and operation of these trips to reduce their impacts. Long-distance trips are normally observed and measured loosely in forms of migration statistics, i.e., the amount of people changing their residence, for which data collected from registries and field inquiries once every a number of years [6]. However, as the official migration statistics only focus on changes in place of usual residence or permanent migration [7], the temporary migration such as touristic trips thereby go unmeasured. As a result, these temporary moves of population are missing from most migration statistics.

Population fluctuations created by these temporary migration trips can significantly affect traffic patterns, housing prices, retail sales, and the use of public transportation, medical services, recreational facilities, as well as other publicly and privately provided goods and services. As functional linkages are often found between temporary migration and touristic flows, the socio-economic and demographic characteristics of the population are also substantially impacted [8]. Unfortunately, there are limited data sources that provide complete and consistent coverage of temporary migration trips. Estimates are typically derived from a variety of administrative records, business statistics, sample surveys, and interviews [1], [9], [10]. Consequently, an extensive and large-scale analysis of temporary migration flows, such as trip generation and attraction, determinants, and consequences, is not possible. This study therefore aims at addressing this shortcoming by introducing a new and feasible method based on an opportunistic sensing approach to gather temporary migration trip information on a large (countrywide) scale.

The opportunistic sensing approach has the edge over its counterpart approach, i.e., active sensing, by which popular sensors such as GPS tracking devices are used to track individuals in travel surveys [11]. Although it can generate a fine-grained human mobility data, the privacy issues and regulations, e.g., EU GDPR (general data protection regulation), have largely limited this type of detailed mobility data to be available for a large-scale analysis. Recently, only data from some specific groups of tracked individuals can be collected, such as university students [12], customers of a company [13], and urban cyclists [14]. Retrospective survey questions can also be used to study the timing of a migratory event. However, retrospective methods are subject to recall bias.

On the other hand, data collected by the opportunistic sensing approach is a data that is originally collected for one purpose, but it also creates an opportunity to be used for another purpose. In our case, a set of mobile phone network data and also known as call detailed records (CDR) that are originally collected by a cellular network service provider for billing purposes, but the data can also be exploited to understand human mobility [15].

CDR contains communication logs and associated user locations. Each time the mobile phone user connects to a cellular network by making or receiving a phone call or using internet, the locations of connected (nearest) cellular towers of both ends are recorded along with communication information such as timestamp, call duration, and user identification. Collectively, these location records of individual users can be explored systematically to investigate various aspects of human mobility and transportation such as social network influenced mobility [16], public transit demand [17], traffic volume [18], [19], trip generation [20], trip distribution [21], [22], route choice [23], transport mode [24], and migration [25], [26]. However, an effort to apply CDR data to infer and analyze temporary migration trips has not yet been made. So, this study builds on the state of the art by utilizing CDR data to advance our understanding of temporary migration flows and their spatial determinants, which therefore makes the following distinct contributions: (i) a heuristic based approach to infer temporary migration flows using CDR data; and (ii) a feasible framework for analyzing spatial determinants of temporary migration. The proposed methodology and insights gained from our analysis can be useful for transport and urban planners to establish better informed travel demand modeling and planning strategies.

Our methodology is described in the next section, which includes CDR data preprocessing, subject selection, temporary migration flow inference, and a framework for identifying spatial determinants of temporary migration flows. Results of our analysis are then presented and discussed in the succeeding section. The paper finally concludes with a summary, limitations of the study, and an outlook on future work.

II. METHODOLOGY

While circulatory migration refers to repeated movements, temporary migration involves a one-time only temporary stay and a return to the primary residence which eventually closes the migration cycle [27]. Many types of mobility can be classified as temporary migration, which are all important for specific purposes. Our study focuses only on extended stays at a second residence [9], [16] that are not one-day trips or short-term overnight visits. These extended stays can principally link to tourism that also represents one form of temporary migration. To study these temporary migration flows, we make use of a CDR data from which a set of suitable mobile phone users are extracted as subjects and their residence as well as temporary migration flows are inferred based on statistical characteristics. As there are
factors that can influence the temporary migration, in this study we explore particularly spatial determinants induced by area’s local places that potentially attract temporary migration flows. The overview of our methodological approach for the analysis is shown in Fig. 1.

A. MOBILE PHONE NETWORK DATA
An anonymized CDR data used in this study was collected from 1,891,928 mobile phone users who were subscribers of one of the largest telecom operators in Portugal, which account for over 18% of the country population. Data includes communication logs over the course of 14 months from April 2, 2006 – June 30, 2007 where half of September and whole October records are missing. Each call detailed record includes identifications (IDs) of both calling and called users, IDs of connected cellular towers of both ends, call duration, and timestamp. Geolocation information of cellular towers is given in a separate lookup table with their corresponding IDs. There is a total of 6,358 cellular towers, each serves an area of 14 km$^2$ on average and its coverage decreases to 0.13 km$^2$ in metropolitan areas. Only the cellular towers that are located within the Continental Portugal are considered in this study, hence excluding the autonomous regions of Portugal, i.e., Azores and Madeira islands.

Individual phone numbers were hashed with unique IDs to safeguard personal privacy by the telecom operator before leaving their data storage facility. Information about text messages and data usage (Internet) are not included in this dataset.

Over the course of the 14 months, the data includes over 500 million records of cellular connectivity. To give an overview of how connectivity is distributed over time, Fig. 2 shows the average number of cellular network connections separated by hours of the day, days of the week, dates of the month, and across the 14-month period. Hourly, the connectivity is low during the late-night hours, and it rises from 7:00 to reach its first peak around 13:00 which is a lunch hour, then it reaches the second peak around 18:00, during the after-hours. Weekly, the connectivity rises from Monday to Friday, and then drops significantly on the weekends where Sunday has the lowest connectivity. Monthly, the connectivity spreads consistently throughout the month. Over the 14-month period, a significantly low connectivity is observed in September which is due to the missing records. Intuitively, high connectivity is observed during the summertime (May - August) and holiday season (November – January).

B. SUBJECT SELECTION
As the level of connectivity implies the amount of mobile phone user’s location records, selecting users with a higher level of connectivity can potentially provide us with a finer grained mobility information. However, selecting users with a very high connectivity may result in an undesired small subject size. After trying various options, we decided that it was reasonable to select the users with at least five connections in each of the 14 months as our subjects for this study, which yields 538,394 users.

Since migration was our focus of the study, so we needed to be able to eventually detect the change of residence of our subjects. To do so, firstly each subject’s residence must be identified. We applied the same approach with [16] to identify the subject’s residence by assigning the most frequently connected cellular tower location during the night hours (22:00 – 7:00) as an approximate location of residence. Hence, only subjects with the night-hour connectivity were considered further in our analysis, which subsequently left us with 148,215 subjects. To show the accuracy of this residence detection approach, for each of the 14 months we identified a residential location of each subject and only considered subjects whose residence remains the same across all months – i.e., 27,004 subjects. Subjects were grouped together according to belonging districts of detected residence, and then compared against the actual census population density distribution. Statistically, the CDR-based
population is highly comparable with the actual census data with a relatively high correlation coefficient ($R$-value) of 0.9401, which is in line with [16]. The comparing census-based and CDR-based population values across all districts of Portugal are listed in Table 1, where their locations and boundaries are illustrated on the Continental Portugal map shown in Fig. 3.

| District          | Census population | CDR-based population |
|-------------------|-------------------|----------------------|
| Lisbon            | 2,250,533         | 5,185                |
| Porto             | 1,817,117         | 6,274                |
| Setubal           | 851,258           | 2,640                |
| Braga             | 848,185           | 2,377                |
| Aveiro            | 714,200           | 1,301                |
| Leiria            | 470,930           | 567                  |
| Santarem          | 453,638           | 1,451                |
| Faro              | 451,006           | 393                  |
| Coimbra           | 430,104           | 781                  |
| Viseu             | 377,653           | 1,005                |
| Viana do Castelo  | 244,836           | 562                  |
| Vila Real         | 206,661           | 1,122                |
| Castelo Branco    | 196,264           | 391                  |
| Evora             | 166,706           | 727                  |
| Guarda            | 160,939           | 170                  |

The number of temporary migration trips varies from the minimum of one trip to the maximum of three trips made by each subject. There is a total of 56,899 temporary migration trips detected. A bar chart displaying the distribution of temporary migration trips across these 54 patterns is shown in Fig. 5. The number of migrations detected in each pattern varies from 1 to 3,897 trips. Most trips were detected intuitively in the summertime when many people take a vacation. Most trips were seen in the pattern #3, which represents single trips made in July, followed by 2,205 single trips made in August represented by the pattern #4. Among the two-trip patterns, the pattern #24 has the highest number of 153 detected trips made in August and December, which are in the summer and Christmas holiday periods. Pattern #52 has 19 detected trips, which is the highest among the three-trip patterns, which represents trips made in August, December, and April. There are certainly other potential temporary migration patterns that could take place, but in this study, we only considered these patterns that were observed from our dataset and temporary migration definition.

There are 40,521 temporary migration trips that flow within their own districts or intra-district flows detected from 34,891 trip makers. Porto has the highest flows of 15,758 trips followed by 13,210 trips made by Lisbon’s residents. On the other hand, Guarda and Beja are the among the lowest intra-flow districts with 285 and 462 flows, respectively. The amount of these intra-district flows ranked by districts is shown in Fig. 6. In addition, Fig. 7 shows the distribution of these intra-district flows geographically. Districts that are nearby Lisbon and Porto tend to also have a relative higher level of intra-district flows compared to other districts that are more distant from these two top tourist destination cities. When considering distances traveled by these 40,521 intra-district flows, the trip length varies from 1.20 km to 178.23 km with the average of 12.32 km. Trip lengths were calculated based on the road network distance using the Google Directions API [28]. The distribution of all intra-district trip lengths is shown in Fig. 8, which appears to follow a power-law distribution i.e., $y = ax^b$ where $a = 3.517 \times 10^3$ and $b = -6.548 \times 10^{-1}$. Fig. 8 also shows the distance decay effect where the demand for intra-district temporary migration trips peak at around 12.5km and then decline as trip length increases.

C. TEMPORARY MIGRATION INFERENCE

According to [27], temporary migration refers to a one-time only temporary stay and a return to the primary residence which eventually closes the migration cycle. Based on this definition, we selected subjects who reside at a primary residence and move to a secondary residence, then later return to the primary residence. Each person may make this temporary migration trip multiple times over the 14 months. Methodically, from the pool of 148,215 subjects who had nighttime connectivity, individual residential location was identified for each of the 14 months, from which we selected the subjects who made at least one temporary trip over this period of observation. We also took into consideration a 2-month gap between each migration trip made over the 14 months. As a result, we retained 47,673 subjects who were temporary migrants that exhibit one of 54 observed temporary migration patterns shown in Fig. 4, where A and B indicate a primary and secondary residences.
The temporary migration trips that flow across districts or inter-district flows were classified into incoming and outgoing flows (or inflows and outflows) for each of the 18 districts. The total of 16,378 inter-district flows were detected from 15,027 travelers. The amounts of inflows and outflows per each district are shown in Figs. 9 and 10, respectively. On average, there are 910.70 temporary migration flows per district. Faro has the highest inflows, attracting
S. Phithakkitnukoon et al.: Temporary Migration Flow Inference and Analysis From Perspective

First group includes trip lengths in the range of 3 – 200 km, which are mostly short-distance trips between nearby districts. Trips that are in the range of 200 – 400 km are considered long-distance trips where most trips are made between big cities, i.e., Lisbon and Porto. Very long-distance trips are those that are longer than 400 km, which mostly are trips made between northern and southern districts.

Across the country, districts attract different levels of inflows, which can be due to a number of reasons. From the tourism’s point of view, area attractiveness refers to the elements of a destination that draw visitors away from their usual environment [30]. The term attractiveness is thus often used to describe destination attributes or characteristics that attract...
visitors or lead them to choose that destination. The literature on influential factors in tourism attractiveness affirms that destination attractiveness is a central determinant of competitiveness and success [31], [32]. Thus, studies of city attractiveness for tourism have mainly sought to understand the determinants and their links to destination competitiveness, and to identify multiple attributes and assets that influence urban tourism performance [33], [34].

As the urban assets (i.e., economic, institutional, physical, and social environments) are strong predictors of a city’s competitive performance [35], we further analyzed spatial characteristics of each district to examine the determinants of district attractiveness by collecting and analyzing urban asset information based on the type of places situated in the area. Since local businesses, government offices, and public services are among those of urban environments that comprise the urban assets, we gathered information about places that are located in each district by using the Google Places API requests [36]. Information about the place is returned for each request made, such as place’s name, address, opening hours, geolocation (latitude and longitude), type, reviews, etc., from which we used only the geolocation and place type information for our analysis. There are 90 different place types specified by the Google Places API, as shown in Table 2.

By providing a geolocation and its radius in an API request, a set of places located within the specified circular area is returned as a result. The maximum value allowed for the radius is 50 kilometers. Technically, in order to gather all places within a district, we needed to locate center points with 50-km radius that can cover the whole district’s area with minimal overlaps. To achieve this, for each district’s area we drew a Voronoi diagram [37] with our desired radius distance as the maximum spacing between centers. As a result, a set of Voronoi polygons were obtained and their corresponding centers were then used as a set of center points for API requests. An example of the resulting set of center points obtained from our Voronoi diagram-based approach is shown in Fig. 14.

With this approach, there are a couple of data cleansing steps that need to be carried out. Due to some overlapping circle areas of Voronoi diagram, some duplicate places were obtained. So, for each district, we searched through our collected places and removed all duplicates. Moreover, the center points that are near district boundary can potentially bring in some places that belong to neighboring districts. Thus, for each district, we filtered out collected places that do not belong to the district according to their addresses.

The districts with the highest numbers of places are Lisbon (30,703), followed by Porto (30,692) and Braga (20,281). On the other hand, districts with the lowest numbers of places are Portalegre (4,858) followed by Braganca (5,200) and Beja (5,281). In terms of the place type, countrywide there are...
TABLE 2. A list of 90 different place types.

| Place Type                | Place Type                | Place Type                |
|---------------------------|---------------------------|---------------------------|
| Accounting                | Electrician               | Night club                |
| Airport                    | Electronics store         | Painter                   |
| Amusement park            | Embassy                   | Park                      |
| Aquarium                  | Fire station              | Parking                   |
| Art gallery               | Florist                   | Pet store                 |
| 6th                        | Funeral residence         | Pharmacy                  |
| Bakery                    | Furniture store           | Physiotherapist           |
| Bank                      | Gas station               | Plumber                   |
| Bar                       | Gym                       | Police                    |
| Beauty salon              | Hair care                 | Post office               |
| Bicycle store             | Hardware store            | Real estate agency        |
| Book store                | Hindu temple              | Restaurant                |
| Bowling alley             | Residence goods store     | Roofing contractor        |
| Bus station               | Hospital                  | RV park                   |
| Cafe                      | Insurance agency          | School                    |
| Campground                | Jewelry store             | Shoe store                |
| Car dealer                | Laundry                  | Shopping mall             |
| Car rental                | Lawyer                    | Spa                       |
| Car repair                | Library                   | Stadium                   |
| Car wash                  | Liquor store              | Storage                   |
| Casino                    | Local government office   | Store                     |
| Cemetery                  | Locksmith                 | Subway station            |
| Church                    | Lodging                   | Supermarket               |
| City hall                 | Meal delivery             | Synagogue                 |
| Clothing store            | Meal takeaway             | Taxi stand                |
| Convenience store         | Mosque                    | Train station             |
| Courthouse                | Movie rental              | Transit station           |
| Dentist                   | Movie theater             | Travel agency             |
| Department store          | Moving company            | Veterinary care           |
| Doctor                    | Museum                    | Zoo                       |

22,583 stores followed by 14,572 restaurants, and 11,870 lodgings. The bottom of the type list includes three Hindu temples, six synagogues, and 12 casinos (five in Faro, two in Coimbra, and one in Setubal, Viana do Castelo, Vila Real, Porto, and Lisbon). Figure 15 shows the ranked districts based on the number of places versus place types.

Based on their characteristics, the places were grouped into six categories; *eatery* (including place types 7, 9, 15, 54, 55, 61, 72 from Table 2), *service* (1, 6, 8, 10, 17-20, 22-24, 27, 28, 30-34, 36, 38, 40, 42, 44, 45, 47, 48, 51, 52, 56, 59, 62, 64-71, 73, 80, 84, 89), *leisure* (3, 4, 5, 13, 16, 21, 39, 53, 57, 58, 60, 63, 74, 78, 79, 88, 90), *education* (49, 75), *shopping* (11, 12, 25, 26, 29, 35, 37, 41, 43, 46, 50, 76, 77, 81, 83), and *transport* (2, 14, 82, 85-87).

To examine the determinants of district attractiveness, we used the ordinary least squares (OLS) regression [38], which is a regression model based on linear least squares method for estimating unknown parameters. Our observed independent variables include the number of places of each place type category in the district, district population density, district area (km²), river length within the district (km), and distance to the nearest coastline from the district centroid (km). For the dependent variable, the amount of inflows drawn by each district is used. However, the inflows inferred from the CDR are only partial. Flows must be expanded to its population through an expansion factor, which is based on a sampling ratio [26]. An expansion factor was derived for each district as a ratio between the total number of population of the district (we used the population between the age of 20 to 59) and the number of sampled users identified as residents of that district from the CDR. The expanded inflow value of each district was then used as the dependent variable for our regression analysis.

III. RESULTS

With the aforementioned dependent and independent variables, an OLS regression was performed. Results from the regression are shown in Table 3, where the R-value is 0.804, which signifies that the inflow value (dependent variable) is relatively explained by the changes in our independent variables. According to the coefficient — a measure of how the change in variable affects the expanded inflow, *leisure* is the most effective variable with a positive coefficient of 59.4785 followed by river length (11.1437), shopping (3.7896), distance to the nearest coastline (2.9821), district area (0.1493), and population density (0.0722). On the other hand, variables that negatively affect the expanded inflow (from highest to lowest coefficients) are education (−32.7243), service (−10.4135), eatery (−1.8918), and transport (−0.7926). Note that a positive coefficient indicates that as the value of the independent variable increases, the dependent variable’s mean also tends to increase, whereas a negative coefficient suggests an inverse relationship.

The result implies that a district with more places for leisure tends to draw a larger amount of inflows from other districts. Statistically, the significance of this implication is underpinned by a low standard error (6.755), i.e., the amount of variation in its coefficient, and a high t statistic value (8.805), i.e., the precision with which the coefficient is measured, as well as a very low p-value \(6.99 \times 10^{-22}\), which...
is a measurement of how likely its coefficient is measured through our regression model by chance.

The total length of rivers within the district (river length) is listed as the second most effective variable according to the coefficient, however its p-value is relatively high (0.754) while typically a p-value of less than 0.05 is acceptable, i.e., indicating strong evidence against the null hypothesis. It is likewise for other variables with a positive coefficient except for the population density, which has a low p-value of 8.97e-9 suggesting that the district population density is another determinant for its district’s inflow though not as strong as leisure.

When considering negative factors for the district’s inflow, service with its coefficient of -10.4135 appears to be the only variable that is statistically significant (p-value = 0.03). Interestingly, this seems to suggest that places for public services such as bank, cemetery, city hall, courthouse, local government office, and so on are not urban asset that helps attract its inflow.

Overall, our regression model produced by the given variables is statistically significant. Its F-statistic is much higher than the required level of 3.95 to reject the null hypothesis, as also reflected in a very low Prob (F-statistic) that indicates the accuracy of the null hypothesis.

IV. CONCLUSION

It is important for temporary migration trips to be considered in regional and national travel demand models as they largely represent long-distance trips that account for a substantial portion of total distance travelled. However, due to limited data sources, estimates are typically derived from a variety of administrative records, which can be inconsistent and incomplete coverage. This study therefore introduces an opportunistic sensing approach and methodology to infer and analyze large-scale temporary migration using mobile phone network data or CDR, which contains logs of communication and associated locations, originally collected by telecom operator for billing purposes.
Mobile phone users in Portugal were analyzed in this study. Subjects were selected based on their cellular network connectivity from which residence and temporary migration trips were inferred according to probable patterns. This consequently enabled us to observe insightful statistical patterns of both intra-district and inter-district trips.

Lisbon and Porto as well as their nearby districts tend to have a relative higher level of intra-district flows compared other districts. Inter-district trips were classified into inflows and outflows from which we observed that Faro, which is a coastal and tourist district, draws most inflows but does not produce a large amount of trips to other districts unlike Lisbon, Porto, and Setubal that highly attract inflows as well as produce outflows. Lisbon has the most overall flows whose top outflow destinations are Faro, Setubal, and Santarem. Furthermore, the flow intensity rises in the summertime, especially in August, a midsummer when many people typically take vacations. Most trips were headed into the southern coastline areas. Overall, the average trip length was 196 km.

Moreover, this study presents a methodology to carry out an analysis of spatial determinants of district attractiveness. As there is a strong linkage between area attractiveness and its urban assets that influence tourism performance, so we examined spatial characteristics of each district to assess the determinants of district attractiveness by collecting and analyzing urban asset information based on the type of places located within the district for which we utilized the Google Places data. We’ve developed a Voronoi-based approach for analyzing urban asset information based on the type of places such as bank, cemetery, city hall, courthouse, and service such as bank, cemetery, city hall, courthouse, and local government office are not urban asset that helps attract its district’s inflow.

Nonetheless, there are a number of limitations of this study that can be further investigated in the future. Firstly, there are other possible spatial determinants such as road network, built environments, and so on, which wasn’t considered in this study and worth exploring in future investigation by taking into account those related geographic information by which potential challenges may include map matching and impact measurement for instance. Secondly, there are also non-spatial determinants e.g., state of mind, social events, family, reasons, and holidays, which could be difficult to measure without an interview or survey. Lastly, future study may investigate other types of temporary migration beyond the 58 patterns as considered in this study e.g., weekend getaway, one-day trips, more than one month stays, and so on.

REFERENCES

[1] M. Bell and G. Ward, “Comparing temporary mobility with permanent migration,” Tourism Geographies, vol. 2, no. 1, pp. 87–107, Jan. 2000, doi: 10.1080/1461676003063466.
[2] J. L. Yang, W. H. Dong, D. F. Kong, G. Liao, J. Wang, W. Li, W. L. Mei, and H. F. Dai, “Characterization and analysis of 2-(2-phenylethyl)-chromone derivatives from agarwood (Aquilaria crassna) by artificial hedging for different times,” Molecules, vol. 21, no. 7, p. 911, Jul. 2016, doi: 10.3390/molecules21070911.
[3] C. Song, Z. Qu, N. Blumm, and A.-L. Barabási, “Limits of predictability in human mobility,” Science, vol. 327, no. 5968, pp. 1018–1021, 2010, doi: 10.1126/science.1177170.
[4] F. T. Charlene Rohr, J. Fox, A. Daly, B. Patruni, and S. Patil, “Modeling long-distance travel in the U.K.,” RAND Eur., Cambridge, U.K., Tech. Rep. TR-899-DFT, 2010.
[5] L. Christensen, “Environmental impact of long distance travel,” Transp. Res. Proc., vol. 14, pp. 850–859, Jan. 2016, doi: 10.1016/j.trpro.2016.05.033.
[6] Measuring Disability in a Census, U.S. Bureau of the Census, U.S. Dept. Commerce, Washington, DC, USA, 2017. [Online]. Available: https://www.census.gov/content/dam/Census/library/working-papers/2017/demo/measuring-disability-in-a-census.pdf.
[7] American Community Survey: Operations Plan, U.S. Bureau of the Census, Washington, DC, USA, 2003. [Online]. Available: http://www.census.gov/acs/www/data_documentation/geography/
[8] S. Kozhukhov, X. Chen, D. Yang, G. Issanov, K. Samarkanov, and S. Aliyeva, “Assessment of tourism impact on the socio-economic spheres of the Issyk-Kul Region (Kyrgyzstan),” Sustain., vol. 11, no. 14, pp. 1–18, 2019, doi: 10.3390/su11143886.
[9] S. K. Smith, “Toward a methodology for estimating temporary residents,” J. Amer. Stat. Assoc., vol. 84, no. 406, pp. 430–436, Jun. 1989, doi: 10.1080/01621459.1989.10478787.
[10] S. K. Smith and M. House, “Temporary migration: A case study of Florida,” Population Res. Policy Rev., vol. 26, no. 4, pp. 437–454, Aug. 2007, doi: 10.1007/s11111-007-9037-6.
[11] L. Shen and P. R. Stopher, “Review of GPS travel survey and GPS data processing methods,” Transp. Res. Proc., vol. 34, no. 3, pp. 316–334, May 2014, doi: 10.1016/j.trpro.2014.04.002.
[12] A. Cuttone, S. Lehmann, and M. C. González. “Understanding predictability and exploration in human mobility,” EPJ Data Sci., vol. 7, no. 1, p. 17, Dec. 2018, doi: 10.1100/epjds.s13688-017-0219-1.
[13] S. Phithakkittunkoon, T. Horanont, A. Witayangkurn, R. Siri, Y. Sekimoto, and R. Shibasaki, “Understanding tourist behavior using large-scale mobile sensing approach: A case study of mobile phone users in Japan,” Pervas. Mobile Comput., vol. 18, pp. 18–39, Apr. 2015, doi: 10.1016/j.pmjc.2014.07.003.
[14] F. Rupi, C. Poliziani, and J. Schweizer, “Data-driven bicycle network analysis based on traditional counting methods and GPS traces from smartphone,” ISPRS Int. J. Geo-Information, vol. 8, no. 8, p. 322, Jul. 2019, doi: 10.3390/ijgi8080322.
[15] F. Calabrese, L. Ferrari, and V. D. Blondel, “Urban sensing using mobile phone network data: A survey of research,” ACM Comput. Surv., vol. 47, no. 2, p. 25, Jan. 2015, doi: 10.1145/2655691.
[16] S. Phithakkittunkoon, Z. Smoreda, and P. Olivier, “Socio-geography of human mobility: A study using longitudinal mobile phone data,” PLoS ONE, vol. 7, no. 6, Jun. 2012, Art. no. e39253, doi: 10.1371/journal.pone.0039253.
[17] M. G. Demissie, S. Phithakkittunkoon, T. Sukhvibul, F. Antunes, R. Gomes, and C. Bento, “Inferring passenger travel demand to improve urban mobility in developing countries using cell phone data: A case study of Senegal,” IEEE Trans. Intell. Transp. Syst., vol. 17, no. 9, pp. 2466–2478, Sep. 2016.
[18] K. Lwin, Y. Sekimoto, and W. Takeuchi, “Estimation of hourly link population and flow directions from mobile CDR,” Information, vol. 17, no. 9, pp. 2935–2950, 2016.
[19] M. G. Demissie, G. H. D. A. Correia, and C. Bento, “Exploring cellular network handover information for urban mobility analysis,” J. Transp. Geography, vol. 31, pp. 164–170, Jul. 2013, doi: 10.1016/j.jtrangeo.2013.06.016.
[20] A. Bwambale, C. F. Choudhury, and S. Hess, “Modelling trip generation using mobile phone data: A latent demographics approach,” J. Transp. Geography, vol. 76, pp. 276–286, Apr. 2019, doi: 10.1016/j.jtrangeo.2017.08.020.
[21] M. G. Demissie, F. Antunes, C. Bento, S. Phithakkittunkoon, and T. Sukhvibul, “Inferring origin-destination flows using mobile phone data: A case study of mobile phone users in Portugal,” Proc. 13th Int. Conf. Elect. Eng/Electron., Comput., Telecommun. Inf. Technol., Jun. 2016, pp. 1–6, doi: 10.1109/ECTICon.2016.7561328.
SANTIPHITHAKKITNUKON received the B.S. and M.S. degrees in electrical engineering from Southern Methodist University, Dallas, USA, in 2003 and 2005, respectively, and the Ph.D. degree in computer science and engineering from the University of North Texas, USA. He is currently an Associate Professor with the Department of Computer Engineering, Faculty of Engineering, Chiang Mai University, Thailand. Before joining Chiang Mai University, he was a Lecturer in computing with The Open University, U.K., a Research Associate with Newcastle University, U.K., and a Postdoctoral Fellow with the SENSEable City Laboratory, Massachusetts Institute of Technology, USA. His research interest includes the area of urban informatics.

SORANAN HANKAEW received the B.Sc. degree in statistics and the M.Eng. degree in computer engineering from Chiang Mai University, Thailand. He is currently a Data Science Researcher at Chiang Mai University. His research interests include urban data analytics and intelligent transportation systems.

MERKEBE GETACHEW DEMISSIE received the M.Sc. degree in transport systems from the Royal Institute of Technology (KTH), Sweden, in 2009, and the Ph.D. degree in transportation systems from MIT Portugal Program, in 2014. He is currently an Assistant Professor with the Department of Civil Engineering, University of Calgary, Canada. Before joining the University of Calgary, he was a Postdoctoral Fellow with the University of Coimbra and the Instituto Pedro Nunes, Portugal. His research interests include transport demand modeling, intelligent transportation systems, data mining, and machine learning.

ZBIGNIEW SMOREDA received the Ph.D. degree in sociology from Paris-Est University. He is currently a Senior Researcher at Orange Labs. Before integrating the (future) Orange Labs, in 1995, he worked as an Assistant Professor with Warsaw University, a Researcher and a Lecturer with GRIFS (Université de Paris 8), a Researcher with GAST (France Télécom) and the Observatoire Mondial des Systèmes de Communication. His work at SENSE/Orange Labs is related to sociology of communication and in particular to social uses of ICT and social network forms and transformations associated with technologies.

CARLO RATTI received the Ph.D. degree in architecture from the University of Cambridge. He is currently an Architect and an Engineer who practices architecture in Turin and teaches at MIT, where he is also a Professor and directs the SENSEable City Laboratory. His research interests include urban design, human–computer interfaces, electronic media, and the design of public spaces. He is a member of the Ordine degli Ingegneri della Provincia di Torino, the Architects Registration Board (U.K.), and the Association des Anciens Élèves de l’École Nationale des Ponts et Chaussées.