Bagheri Majdabadi, Mehrdad; Mladenovic, Milos; Kosonen, Iisakki; Nurminen, Jukka K.; Roncoli, Claudio; Ylä-Jääski, Antti

A computational framework for revealing competitive travel times with low carbon modes based on smartphone data collection

Published in:
Journal of Advanced Transportation

DOI:
10.1155/2020/4693750

Published: 01/01/2020

Document Version
Publisher's PDF, also known as Version of record

Hindawi
Journal of Advanced Transportation
Volume 2020, Article ID 4693750, 20 pages
https://doi.org/10.1155/2020/4693750

Research Article

A Computational Framework for Revealing Competitive Travel Times with Low-Carbon Modes Based on Smartphone Data Collection

Mehrdad Bagheri,[1,2] **Milos N. Mladenovic**,[2] **Iisakki Kosonen**,[2] **Jukka K. Nurminen**,[1,3] **Claudio Roncoli**,[2] **and Antti Ylä-Jääski**[1]

[1]*Aalto University, Department of Computer Science, Espoo, Finland*
[2]*Aalto University, Department of Built Environment, Espoo, Finland*
[3]*University of Helsinki, Department of Computer Science, Helsinki, Finland*

Correspondence should be addressed to Mehrdad Bagheri; mehrdad.bagheri@aalto.fi

Received 25 June 2019; Revised 27 December 2019; Accepted 11 February 2020; Published 21 March 2020

Academic Editor: Francesco Galante

Evaluating potential of shifting to low-carbon transport modes requires considering limited travel-time budget of travelers. Despite previous studies focusing on time-relevant modal shift, there is a lack of integrated and transferable computational frameworks, which would use emerging smartphone-based high-resolution longitudinal travel datasets. This research explains and illustrates a computational framework for this purpose. The proposed framework compares observed trips with computed alternative trips and estimates the extent to which alternatives could reduce carbon emission without a significant increase in travel time.. The framework estimates potential of substituting observed car and public-transport trips with lower-carbon modes, evaluating parameters per individual traveler as well as for the whole city, from a set of temporal and spatial viewpoints. The illustrated parameters include the size and distribution of modal shifts, emission savings, and increased active-travel growth, as clustered by target mode, departure time, trip distance, and spatial coverage throughout the city. Parameters are also evaluated based on the frequently repeated trips. We evaluate usefulness of the method by analyzing door-to-door trips of a few hundred travelers, collected from smartphone traces in the Helsinki metropolitan area, Finland, during several months. The experiment's preliminary results show that, for instance, on average, 20% of frequent car trips of each traveler have a low-carbon alternative, and if the preferred alternatives are chosen, about 8% of the carbon emissions could be saved. In addition, it is seen that the spatial potential of bike as an alternative is much more sporadic throughout the city compared to that of bus, which has relatively more trips from/to city center. With few changes, the method would be applicable to other cities, bringing possibly different quantitative results. In particular, having more thorough data from large number of participants could provide implications for transportation researchers and planners to identify groups or areas for promoting mode shift. Finally, we discuss the limitations and lessons learned, highlighting future research directions.

1. Introduction

The increased trip duration with public transportation (PT) and bike compared to private car is often a barrier against modal shift [1–4]. Travelers mostly prefer transportation modes that can provide them with travel time that fits into their limited daily travel-time budget and daily activity space [5]. Therefore, in addition to such incentives as reducing carbon emissions and increasing physical activity [6, 7], understanding potential of low-carbon transportation requires considering travel time limitations with low-carbon modes for individual travelers. Although often counterintuitive, previous studies have shown that a portion of current urban car trips could be potentially made with low-carbon (e.g., PT) and active (e.g., bike and walk) modes, without compromising much of travel time [8, 9]. Travelers can have

an underestimated perception of the travel time with car [10], and conversely, an overestimated perception of travel time with PT and bike [11]. Studies show that car drivers who correct their inaccurate knowledge of travel time with PT and bike are more likely to use it in the future or at least consider PT in their choice set [2, 11, 12]. Therefore, it is important to understand that using low-carbon modes can sometimes have comparable travel time to driving car, i.e., there is a potential for time-relevant low-carbon transportation. Above discussion implies two main requirements of computational framework summarized as follows: first, such understanding requires longitudinal data collection of revealed travel behavior of individuals. The alternative low-carbon door-to-door trips should be computed based on current travel behavior of urban travelers, to reflect realistic situations, and it is the collected data that represent current travel behavior [13]. Second, such understanding requires analysis of the collected data to explore opportunities for low-carbon travel alternatives accounting for travel time limitations.

Previously, revealed preference travel surveys have traditionally been used to collect travel behavior datasets, either in person or online. However, this approach has several limitations such as being resource-intensive for longitudinal data collection, having low data quality due to human error in responding [14, 15], and difficulty in capturing complete multimodal door-to-door travel [16]. In response to these challenges, researchers adopted new methods to automate and facilitate data collection. New methods first utilized GPS devices [14, 17, 18], and later on, cellular network [19, 20] and smartphones [21, 22], as many people already carry smartphones during their daily activities. Cellular-network positioning data from call and activity information can provide data consistent with travel surveys, although having issues with geolocation accuracy [23]. Smartphone sensing as the most recently emerging approach uses GPS, accelerometer, and other built-in phone sensors, thus recording all the steps of multimodal door-to-door trips with high spatiotemporal resolution [24–26]. It also has the potential of long-term fine grained temporal coverage, i.e., frequent data points during the whole day, over several months or even years, as well as full spatial coverage in city if diverse and large number of travelers participate [25]. In addition, as this approach uses mobile apps, it is possible to interact with the traveler in case additional socioeconomic or attitudinal information is needed.

Previous research in this domain has largely focused on validation of smartphone-based sensing technology. On the contrary, exploring time-relevant low-carbon transportation opportunities still faces an automated processing challenge of a large amount of smartphone-based high-resolution data [24]. In this context, only limited research has focused on exploring the potential of modal shift to low-carbon modes while considering travel time limitations based on realistic travel data [8, 9, 27, 28]. The authors of [8, 9] used a travel survey collected in Madrid to compute alternative trips with PT, bike, and walk. They concluded that a fraction of current car trips could potentially be changed to PT and cycling without any increase in travel time. However, further analysis could be done especially with longitudinal focus on individual traveler, accounting for emerging transportation modes such as electric bike (e-bike) [29, 30]. In addition, all these previous research studies use surveys for data collection, while none of them utilizes smartphone-based mobility datasets.

The objective of this research is formulation and evaluation of computational framework for analysis of smartphone-based travel data, for understanding the potential of low-carbon travel alternatives while accounting for travel time duration. The proposed framework estimates the extent to which changes in travel mode could reduce carbon emission without a significant increase in travel time. The paper is organized as follows. Section 2 explains our methodology and computational framework. Section 3 explains arrangement of long-term data collection experiment in the Helsinki region, Finland, as well as summary of the collected travel data. Section 4 applies the developed computational framework to the Helsinki region mobility dataset and presents results and insights. Section 5 presents discussion and evaluation of the computational framework and provides suggestions for further research. Finally, Section 6 concludes the paper.

2. Computational Framework

The computational framework includes six components, summarized in Table 1 together with evaluation parameters. We implement the framework as new modules on top of the software system initially implemented for the TrafficSense (TS) project [31, 32]. TS open-source software is explained in [33], and its source code, documentation, and setup instructions are available on github [34]. The first component of the framework is collection and filtering of movement data and is mostly addressed by TrafficSense smartphone app. For data collection, volunteer participants use the TS app to automatically record their daily travel trace for long periods. The app automatically collects anonymous real-time movement data, referred to as point data, from GPS, accelerometer, and other phone sensors, and sends it over the Internet to the TS web server to be stored in a centralized database. The point data are collected at specific time intervals and include timestamped geolocation (longitude and latitude) of traveler at each sampling interval along the trip route, the estimated locational accuracy of the sampled point, and an initial estimate of transportation mode. Sampling interval is 10 seconds when moving to achieve accurate enough description of the trip route, for example, 333 m distance interval at 120 km/h speed. TS backend server refines and transforms the stored data points to retrieve individual legs of trips together with their transportation modes. To tackle noise, the server discards sampled points with a locational inaccuracy of worse than 50 meters. TS app also provides a menu for revision and confirmation of the automatically detected modes. Further details of the computational requirements and software design are presented in the supplementary material (available here).

2.1. Identifying Door-to-Door Trips. This paper implements a trip-extraction module on top of the data collection and

Table 1: Components of the computational framework together with parameters evaluated in experimental results.

Component	Parameters
1. Experiment setup, filtering the sampled data, and extracting door-to-door trips	Temporal coverage and distribution of observed door-to-door trips Spatial coverage and density of trips and origin/destinations throughout the city Distribution and share of recorded trips among all participants
2. Potential of substituting observed car trips with time-relevant low-carbon alternatives	Size and fraction of car trips with possibility of shift to low-carbon time-relevant modes Distribution of modal shift possibilities by trip distance, departure time, and transportation mode Spatial coverage of substitutable car trips throughout the city, depending on the alternative modes grouped as PT and bike
3. Emission savings and increased nonmotorized traveled distance	Size of emission savings as a result of modal shift from car trips Distribution and frequencies of emission savings among trips Distribution of emission savings among different alternative modes Size of increased traveled distance with nonmotorized modes, implying increased physical activity by more walking and cycling Comparison and clustering of trips based on potential emission saving versus increased bike and walk traveled distance
4. Person-based analysis	Identifying active participants for the purpose of a valid person-based analysis Distribution and share of recorded trips among active participants Distribution of modal shift possibilities among active participants Distribution and frequencies of emission savings among participants Comparison of trips based on their emission saving versus increased bike and walk traveled distance
5. Frequently observed unique trips	Distribution of unique trips among active participants Size and fraction of unique car trips with possible shift to low-carbon time-relevant modes Consistency of modal shift possibility among unique car trips Departure time, trip distance, and travel time of unique car trips with consistent modal shift possibility Size of emission savings Comparison of unique trips based on their emission saving versus increased bike and walk traveled distance Distribution of modal shift possibilities among active participants
6. Potential of substituting noncar trips with low-carbon alternatives	Similar variables as listed above, in addition to share of transportation modes among low-carbon alternatives

filtering component. We need to connect the related consecutive trip legs to identify each multimodal door-to-door trip. In addition, each trip should be attributed to the individual traveler that made the trip. As the data collected in TS database, first, only contain isolated trip legs, we implement additional postprocessing in the server to extract whole multimodal door-to-door trips from the recorded individual legs. This module detects the interrelated individual trip legs and combines them into single multimodal door-to-door trip records. For this purpose, the module traverses through the time-sorted legs in order to connect the ones belonging to the same multimodal trip. Both nonmotorized (e.g., walk and bike) and motorized (e.g., PT and private car) legs are considered. There is usually a rather longer pause or idle time between legs that do not belong to the same trip. Such pause could be a sign of a "stay location" or in other words "activity location." For example, pauses more than 10 minutes are considered to be an indication of a stay location and thus start/end of a trip [23]. Our module employs the same approach. A door-to-door trip starts at an activity location and ends at an activity location. Following leg sequences are examples of individual door-to-door trips extracted out of leg records:

(Stay Home) ⟶ WALK ⟶ BUS ⟶ TRAM ⟶ WALK ⟶ (Stay at Work)

(Stay at Work) ⟶ WALK ⟶ BUS ⟶ WALK ⟶ (Shopping in a Shopping Center)

(Shopping in a Shopping Center) ⟶ WALK ⟶ BUS ⟶ WALK ⟶ (Stay at Home)

Our system also detects and discards walking and running exercises as well as round-trip leg sequences where the traveler starts and ends the trip at the same geolocation. Furthermore, filtering process also tries to recognize and discard trips that were detected incorrectly due to erroneous or missed data. As a result, we have identified multimodal door-to-door trips that will be given as input to next computational steps. These data include trip start/end timestamps, origin/destination geolocations, trip legs, transportation mode of each leg, idle times between legs, and trajectory points making up the whole trip.

2.2. Computing Potential of Time-Relevant Low-Carbon Alternatives. For the components two to six, we implement data analysis and visualization as postprocessing modules on top of the original TrafficSense system. The framework

computes *low-carbon alternatives* including walk, bike, and PT as well as their combination in a multimodal route. Among the alternatives, a *time-relevant alternative* has either lower or negligibly higher travel time compared to private car. Our method considers emission and travel time as key criteria for the alternative modes. Among the alternatives of each observed trip, choice priority is given to the mode with lowest emission if it competes with car in terms of travel time. Unlike [8, 9], our model accepts a small compromise in travel time when considering a potential shift from car to low-carbon alternative. Some previous works [27, 28] apply a maximum travel-distance threshold while selecting walk and bike alternatives, according to the statistics on the usual walking and cycling distances in the city. Unlike these works, we neither prioritize nor filter out the computed bike and walk alternatives based on their travel distance. As an example, if walking or cycling for a particular trip and route is not among the fastest choices, they are not chosen as potential alternatives no matter how convenient or short the trip appears to be. For PT trips, we consider maximum access-egress walking distance of total 1 km which should be feasible for most of travelers. In contrast to previous works such as [28], one criterion in our model is that the alternative trips should be feasible according to the original departure time and OD geolocations of the recorded trip. If the desired alternative is bike, a cycling route between O and D should be available according to up-to-date cycling path information. Similarly for PT, trip should be possible according to city PT routes and line schedules during the date and time of original trip. We utilize well-established open-source routing software together with city open-data to make sure that the potential alternative trip is actually possible. Figure 1 presents the algorithm, and Table 2 describes the variables used.

```
Algorithm:
FOR EACH r ∈ R LOOP:
    FOR EACH m_i ∈ M LOOP:
        A_r := {}
        a_i := ComputeRoute(O_r, D_r, t_s, dt_r, m_i, PTS, PTR, NET)
        IF a_i was computed successfully (such route with mode m is possible) THEN:
            T_i := ComputeTravelTime()
            e_i := ComputeEmission()
            Add a_i to A_r
        EndIF
    EndLOOP

    A_r := Sort(A_r) ascending by e
    T_min := min(T)
    tra_r := {}
    FOR EACH a_i ∈ A_r LOOP:
        IF m_i = 'car' THEN:
            BreakLOOP
        ELSE:
            ΔT_i := T_i - T_min
            IF ΔT_i ≤ C THEN:
                tra_r := a_i
                Add {r, tra_r} to TRA
                BreakLOOP
            EndIF
        EndIF
    EndLOOP
    IF m_r = 'car' AND tra_r <> {} THEN:
        Add {r, tra_r} to TRS
    EndIF
EndLOOP
```

Figure 1: Identifying time-relevant low-carbon alternatives.

First, for each observed door-to-door trip ($r \in R$) between origin (O_r) and destination (D_r) geolocations, we compute using the ComputeRoute function the set of alternative multimodal trips (A_r) so that each alternative trip ($a_i \in A_r$) matches O_r, D_r and departure time (t_s) and date (dt_r) of the original trip (r), and each alternative trip is made with a different mode of transportation ($m_i \in M$). Each computed a_i trip might pass through a different or the same route as the original trip. The ComputeRoute function considers dt_r because date of trip can make a difference for PT routes as schedules may change by weekday or month. Date and time may have influence also on computing car routes when road traffic information is available. For some trips, we conclude that certain alternative modes are not feasible. In our model, maximum of $|M|$ alternatives are possible for each r. When computing PT routes, if there are multiple choices with one PT mode, for example, multiple bus lines, we choose the fastest option with least transfers. We also quantify travel time (T_i) and emission (e_i) of each $a_i \in A_r$.

Second, to select potential time-relevant alternatives per trip, priority is first given to low-carbon emission and secondly to travel time while comparing the computed alternatives. We sort the computed alternatives by emission and then compare travel time (T_i) of each alternative $a_i \in A_r$ to the fastest alternative with $T_{\min} = \min T_i$. It is assumed that traveler might switch from the fastest choice (e.g., car with $T = T_{\min}$) to the low-carbon alternative choice (a_i) when added travel time ($\Delta T_i = T_i - T_{\min}$) does not exceed a small constant value of C. Among the alternatives that satisfy this time condition, the one with the lowest emission is saved as the *low-carbon time-relevant alternative* (tra_r) for the trip r. The resulting time-relevant choices are sustainable but still fast enough to compete with car.

Third, the system compares the computed time-relevant low-carbon alternatives to current mobility behavior. If the trip was originally made with car (i.e., $m_r =$ *"car"*) while the trip has a time-relevant low-carbon alternative, we conclude that traveler could have made a better choice in terms of both travel time and emission for this particular trip. In other words, there is a potential for *time-relevant modal shift*. Finally, we perform analysis, including computation of differences between attributes of the time-relevant alternative trips and the observed recorded trips to evaluate parameters such as frequency and size of emission savings as presented in detail later in Section 4. Description so far was to address the requirement of understanding all recorded trips made by all participants. As mentioned before, our methodology also requires understanding individual travelers and their travel patterns. However, for a valid person-based analysis, we first need to select only those participants with enough quantity of recorded trips. For this purpose, an "active participant" is defined as a traveler who has a minimum of 30 "active days," in which at least one trip per day has been recorded from that traveler's smartphone. Next, the computed values are grouped and summarized per active participant.

Table 2: Variables used in the framework.

Parameter	Definition	Value
R	List of recorded (observed) door-to-door trips	
r	Each recorded (observed) door-to-door trip	$r = \{O_r, D_r, t_s, dt_r, m_r\}$
O_r	Coordinates (latitude, longitude) of the origin of a recorded trip (r)	
D_r	Coordinates (latitude, longitude) of the destination of a recorded trip (r)	
t_s	Start time (i.e., departure time) of the recorded trip (r)	
dt_r	Date of the recorded trip	
m_r	Main transportation mode ($m \in M$) in each multimodal trip so that the longest distance of the trip is traveled by mode m.	
M	Modes of transportation	{Car, bus, train, metro, tram, ferry, bike, E-bike, walk}
A_r	List of computed alternative trips for each $r \in R$	
a	An alternative multimodal route for a trip	$a = \{O, D, t_s, d, m, T, e\}$
T	Door-to-door travel-time of each trip	
d	Door-to-door distance traveled along the trip route	
e	Emission per traveler of the whole multimodal door-to-door trip a	Calculated based on equation (1)
$T_{\min}$	Travel time with the fastest choice of computed alternatives (A)	
C	Maximum travel time increase from $T_{\min}$ tolerated by traveler	Assumed 3 minutes
ΔT	Increased travel time with an alternative compared to fastest choice	$T - T_{\min}$
tra_r	Chosen low-carbon time-relevant alternative for a recorded trip r	
TRA	List of low-carbon time-relevant alternatives (TRAs) for all recorded trips regardless of the original mode choice m_r.	Each list element is $\{r, tra_r\} \in$ TRA, and $\lvert\text{TRA}\rvert = \lvert R\rvert$
TRS	Potential time-relevant modal shifts (TRS): list of car trips ($m_r =$ "car") where a potential time-relevant low-carbon alternative is available	Each list element is $\{r, tra_r\} \in$ TRS, and $\lvert\text{TRS}\rvert \leq \lvert R\rvert$
PTS	Public transportation schedule for the city	Usually available in general transit feed specification (GTFS) files created and updated by cities or transportation agencies
PTR	Public transportation route and line data for the city	In GTFS files as explained above
NET	Road network comprising digitalized street network data of the city that may also include biking paths	Provided by OpenStreetMap (OSM)

Table 3 shows an example execution of the algorithm for a 5 km trip from central to south part of Helsinki. Trip origin is "Sturenkatu 9" departing at 16:30 and destination is "Laivurinkatu 39." Computed choices are sorted by emission. The fastest choice is with e-bike, that is, $T_{\min} = T_i$ where $m_i =$ "e-bike." Therefore, the time-relevant choices are bike, e-bike, and car, where T with bike and car is maximum 3 minutes more than $T_{\min}$. However, computed e_i values show that only bike and e-bike routes can be considered as low-carbon time-relevant alternatives. Table 4 shows an example of a recorded trip in our travel dataset, for which the computed PT alternative is almost as fast as car. Table 4 shows details of the computed car and PT trips. The PT trip is multimodal and comprises walk, bus, and train.

Computation of alternative trips and routes is explained below. To implement the function ComputeRoute(), we leverage open-source route planning and mapping APIs together with open data of Helsinki region. Our system uses HTTP REST calls to query car, PT, bike, and walking routes from OpenTripPlanner (OTP) server of Helsinki region [35]. We pass O_r, D_r, t_s, and d_r of the original observed trip as query parameters to OTP. OTP computes and returns door-to-door routes for each requested mode of transportation. Returned routes include all trip legs with timestamps and detailed geolocation steps along the route from origin to destination. We process the returned trip plan data and store it in our database. When planning PT alternatives, we may do minor adjustments in t_s parameter to compensate for inaccuracies of GPS geolocation in the observed trip data. More details are explained in [33]. Functions ComputeTravelTime() and ComputeEmission() quantify travel time (T_i) and carbon emission (e_i) of each alternative route plan a_i. We retrieve value of T_i directly from the JSON response of our query to OTP server. However, OTP at the moment does not compute e-bike. Therefore, we calculate e-bike travel time based on the ordinary bike's travel time returned by OTP. We set e-bike's speed ~16% faster than ordinary bike because according to [36, 37], average cycling speed with ordinary bike is considered 15.5 km/h and average cycling speed with e-bike is considered 18 km/h. Furthermore, to get e_i of each a_i, we have

$$e_i = \sum e_{\text{leg}}, \tag{1}$$

where e_{leg} is the emission per traveler caused by each trip leg of the multimodal multileg a_i trip. e_{leg} is measured by grams of CO_2 (g-CO_2) and calculated as follows:

$$e_{\text{leg}} = d_{\text{leg}} \times e_{\text{pkt}}, \tag{2}$$

where d_{leg} is the distance traveled along each leg of the alternative trip route and e_{pkt} is the average emission per each passenger-km traveled (pkt), measured by grams of

Table 3: An execution of the algorithm for a trip whose origin is "Sturenkatu 9" and destination is "Laivurinkatu 39."

i	Main mode	Travel-time T_i (minutes)	Emission e_i (CO_2 g)	ΔT_i	Is time-relevant compared to $T_{\min}$
1	Walk	44	0	28	False
2	Bike	19	0	3	True
3	E-bike	16 ($T_{\min}$)	0	0	True
4	Tram	30	0	14	False
5	Bus	30	260	14	False
6	Car	18	445	2	True

Table 4: An observed trip in our dataset for which computed PT alternative is almost as fast as car. Trip origin is "Sörnäisten rantatie 7b, Helsinki," and destination is "Kulovalkeantie 6, Espoo," and departure time and date are 21:00 and 9.8.2017.

Computed alternative trip	Total duration	Total distance (km)	Duration traveled by each mode	Distance traveled by each mode (km)
Walk ⟶ Bus ⟶ Walk ⟶ Train ⟶ Walk	00:45:28	18.5	Bus: 00:11:00 Train: 00:19:00 Walk: 00:15:28	Bus: 2.9 Train: 14.9 Walk: 0.7
Car	00:43:43	19.9	Car: 00:43:43	Car: 19.9

CO2 per passenger-km (g-CO_2/pkm) [38]. e_{pkt} is mode-dependent and calculated as follows:

$$e_{\text{pkt}} = \frac{e_{\text{vkt}}}{o_m}, \tag{3}$$

where e_{vkt} is tailpipe emission per vehicle-km traveled (vkt) depending on the transportation type m and o_m is passenger occupancy of m, i.e., average passengers per vehicle depending on the mode. Values of e_{vkt} and o_m can be different in each city depending on its transportation vehicles and passenger volume. We get values of e_{vkt} and o_m for PT modes from statistics of Finland and Helsinki metropolitan area [39, 40]. For example, on average there are $o_m = 1.7$ passengers in each private car and $o_m = 18$ passengers on each city bus vehicle based on 2016 statistics. Vehicle emission is $e_{\text{vkt}} = 151$ g-CO_2/km for private car and $e_{\text{vkt}} = 939$ g-CO_2/km for city bus. Therefore, based on (3), we get $e_{\text{pkt}} = 89$ g-CO_2/pkm for car and $e_{\text{pkt}} = 52$ g-CO_2/pkm for bus.

3. Experiment Setup and Collected Data in the Case Study Region

Mobility dataset for our analysis is collected by TrafficSense app during data collection pilots in Helsinki region. To attract volunteers, first, the research and TS mobile app were advertised by poster banners and online ads in Aalto University's Otaniemi campus and through the university mailing list. Otaniemi is an academic and innovation area hosting the Aalto University campus, located in city of Espoo and a part of Helsinki metropolitan area. Later, a more public advertisement was made using social media. Prize draws were also performed twice to encourage the participants. Data have been collected since 2016 for more than three years. More details about the TrafficSense research and travel data collection are provided on its web portal [31]. Out of the total 135 study participants, 69 have completed an optional questionnaire using a web link provided in the app, reporting their socioeconomic information. Figure 2 shows the income and age distributions of participants as well as that of the whole Helsinki region [41].

The framework filters the collected trip data to discard nonstop round-trips such as walking or running exercises where origin and destination are usually almost at the same geolocation. In addition, only trips with travel distance of more than 500 m and less than 30 km are selected. Trips longer than 30 km are assumed not to be urban trips. This filtering results in $|R| = 25{,}328$ door-to-door trips for the whole region. Furthermore, as explained before in the methodology, for a valid person-based analysis, we need to select only the trip records of "active participants" who have a minimum of 30 "active days." As a result, 68 active participants out of the total 135 registered travelers are identified. These 68 travelers have recorded $|R| = 24{,}377$ trips, that is, 96% of all trips in the city. Active participants are analyzed further in Section 4.3.

Figure 3 shows the timeline overview of the filtered data of all participants and active participants. The peak seen in 2017 is the result of a three-month promotion pilot. For this research, we have used the data collected until the end of March 2019. Figure 4 illustrates distribution of the recorded trips depending on departure time during the day. The overall distribution of observed trips throughout the day reflects the usual daily mobility peaks and lows in cities such as Helsinki and Espoo. For example, the number of trips after 22:00 and before 06:00 is expected to be much smaller than other times of the day, which is well reflected in Figure 4. The morning and afternoon peaks are also seen in the figure. An exception is the relatively low number of observed trips from 06:00 to 08:00 that is usually expected to be part of the morning peak hour. Figure 5 illustrates distribution of the observed trips depending on travel distance. It is seen that a high number of shorter trips (e.g., 0.5 to 2 km) have also been recorded. Figures 6 and 7 illustrate the spatial distribution of the observed trips. In terms of spatial coverage, although

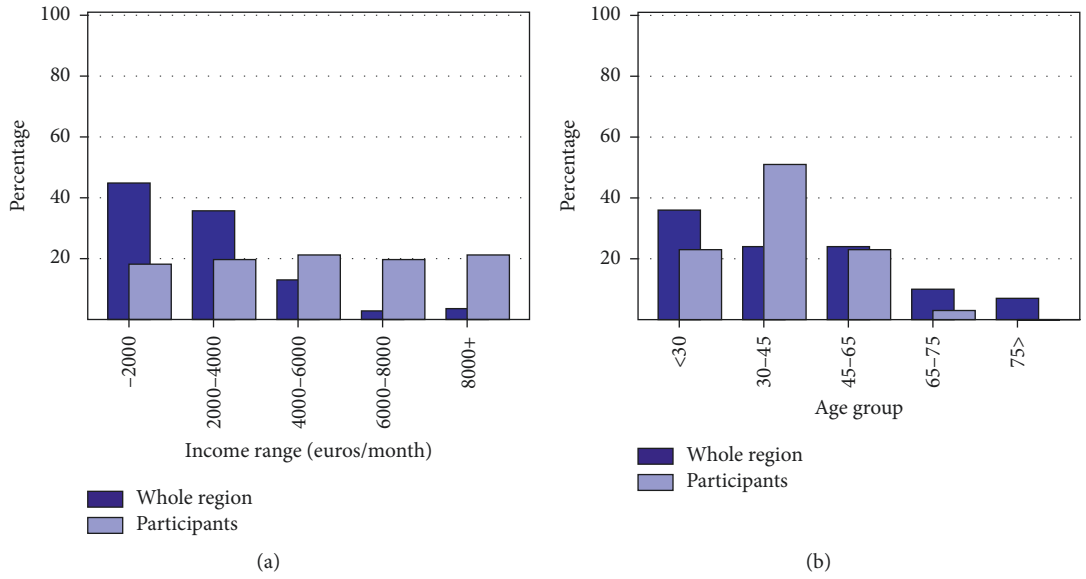


Figure 2: Income and age distribution of participants as well as that of the whole Helsinki region.

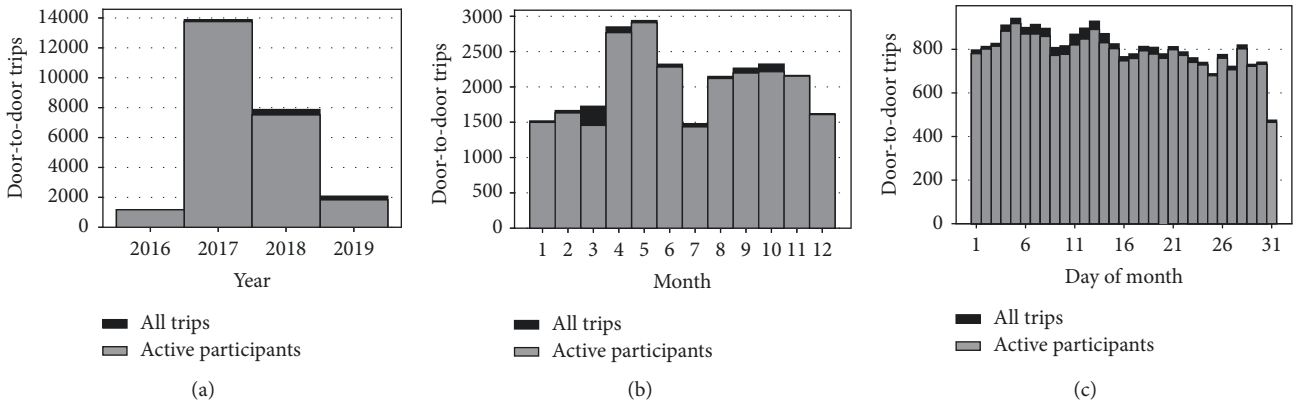


Figure 3: Door-to-door trips observed during the data collection period, by (a) year, (b) month, and (c) day of month.

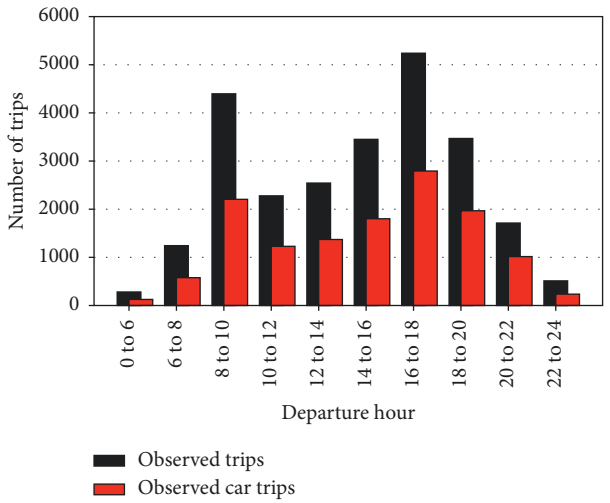


Figure 4: All observed trips vs. observed car trips, grouped by departure time during 24 hours.

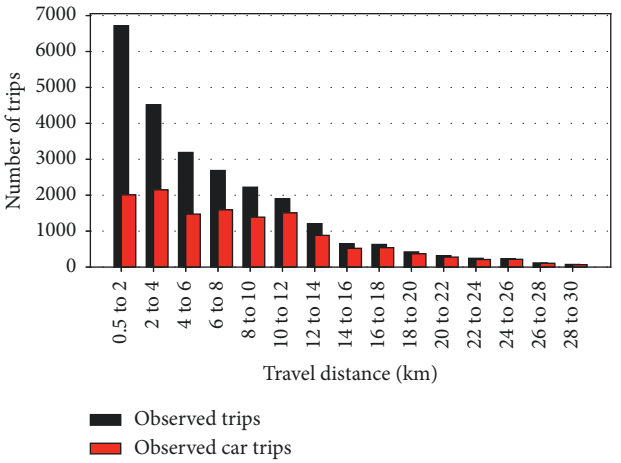


Figure 5: All observed trips vs. observed car trips, grouped by travel distance.

Otaniemi shows a relatively high density of trips, it is seen on the map that the participants have traveled all over the metropolitan area. Therefore, although the participants so far may not be representative of the whole Helsinki region, the collected data have a good spatial coverage over the city. Figure 8 shows the distribution of trips recorded per traveler. On average, around 190 trips have been observed per person. The mean share of each participant from all recorded trips is 0.75% with standard deviation of 1.35. The cumulative distribution of trips per traveler is exponential and, for example, denotes that 20% of participants have recorded 80% of all trips in our dataset.

4. Experimental Results

We apply our framework to the travel dataset of Helsinki region that was explained in previous section. Evaluations and visualizations in this section present examples of quantitative results and what can be achieved using our method and how this framework can be used for analyzing traveler behavior data to understand potential for time-relevant low-carbon mobility.

4.1. Potential of Substituting Car Trips with Time-Relevant Low-Carbon Alternatives. As explained before in Section 2, for each observed car trip, the choice priority is given to the alternative mode with lowest emission if it competes with car in terms of travel time. Out of the total $|R| = 25{,}328$ observed door-to-door trips in Helsinki region, there are 13,324 detected car trips, for which $|\mathrm{TRS}| = 2{,}730$ trips have a time-relevant low-carbon alternative (i.e., PT, bike, or walk). Therefore, 20% of the observed car trips, that is, 11% of all observed trips, have the possibility of being substituted with low-carbon choices without compromising travel time, if not hindered by other choice factors such as weather, physical effort, car ownership, and personal preferences. The potential of PT as alternative is 3% (425 trips), the potential of cycling as alternative is 17% (2298 trips), and the potential of walking as alternative is almost zero. Among the PT alternatives, 31% involve traveling with metro and train, indicating the good potential of rail transportation as a substitution for car. These cases are seen across a wide range of trips from 1 to 22 km.

Figure 9 illustrates spatial aspect of the results by showing distribution of the potential time-relevant mode substitutions depending on travel distance. Total column height in each distance range denotes the fraction of observed car trips in that range with a lower carbon alternative. As expected, the overall potential of cycling in each distance range decreases by travel distance. On the other hand, potential of PT as an alternative to car does not change much by trip distance. Figure 10 illustrates modal share of low-carbon modes, with all PT modes counted in one group. PT and bike compete with each other, with intersection point at the 8 to 10 km range. For trips longer than 10 km, most of the potential alternatives are PT trips. Figure 11 illustrates the spatial distribution of car trips with a low-carbon time-relevant alternative throughout the Helsinki region. This illustration indicates that spatial potential of bike as an alternative is more sporadic throughout the city, while the potential of bus is more focused towards ODs in city center. Bike can be a substitute of private car in several areas of the city and for relatively shorter trips.

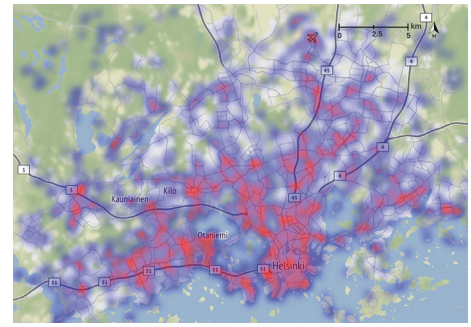


Figure 6: Spatial distribution of the observed trips illustrated by density map of origins/destinations. Blue color represents lower density and red color represents higher density. The map shows the Helsinki metropolitan region including cities of Helsinki, Espoo, and Vantaa (map copyright of OpenStreetMap, under ODbL).

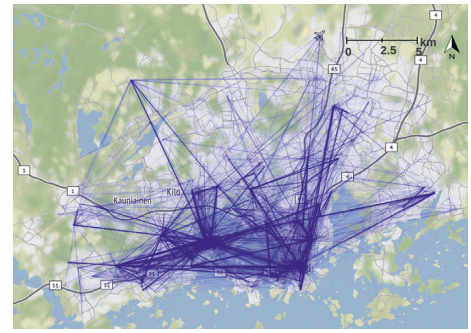


Figure 7: Desire line distribution of the observed trips, based on trip OD pairs. The lines are weighted by the repetition of unique trips. The map shows the Helsinki metropolitan region including cities of Helsinki, Espoo, and Vantaa (map tiles by Stamen Design, under CC BY 3.0; data by © OpenStreetMap, under CC BY SA).

Figure 12 illustrates a temporal aspect of the results by showing distribution of the potential alternatives depending on departure time during the day. Overall column height in each time range denotes the total modal shift percentage among the car trips observed in that range. Figure 13 shows share of each group of low-carbon modes among the potential time-relevant alternatives. The share of walking is almost zero, while share of cycling is always larger than PT during the whole day.

In addition to conventional modes of transportation, we can also test the case if travelers have access to electric bikes. With this assumption, the potential of cycling as an alternative increases from 17% to 24% of all observed car trips. The total potential of time-relevant low-carbon modes increases from 20% to 27%.

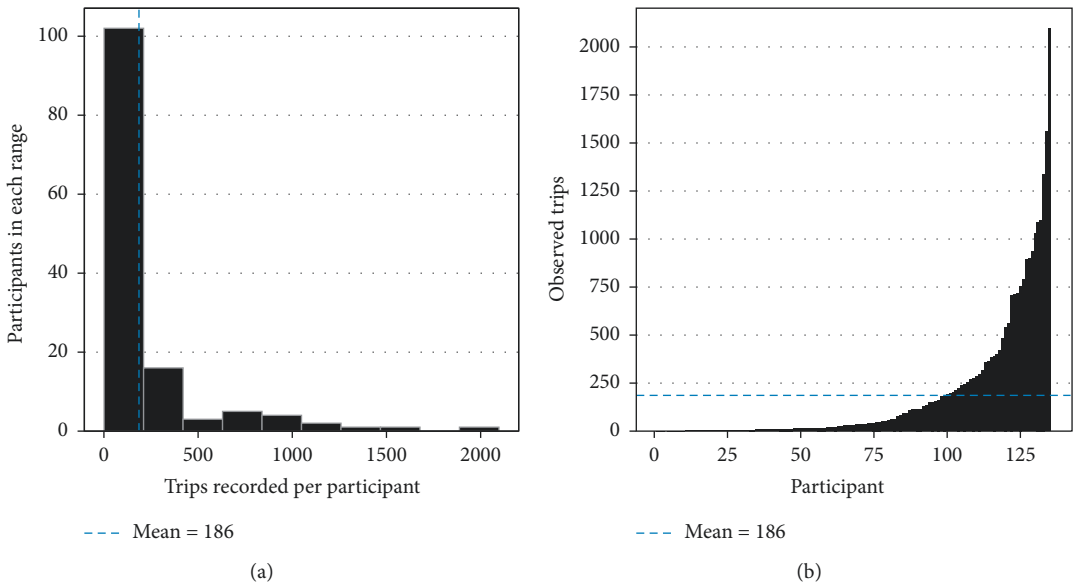


Figure 8: Distribution and share of trips recorded among all participants.

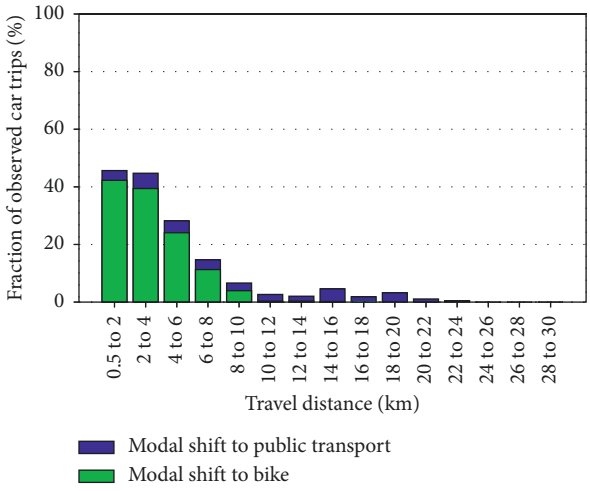


Figure 9: Observed car trips with a time-relevant low-carbon alternative, depending on travel distance.

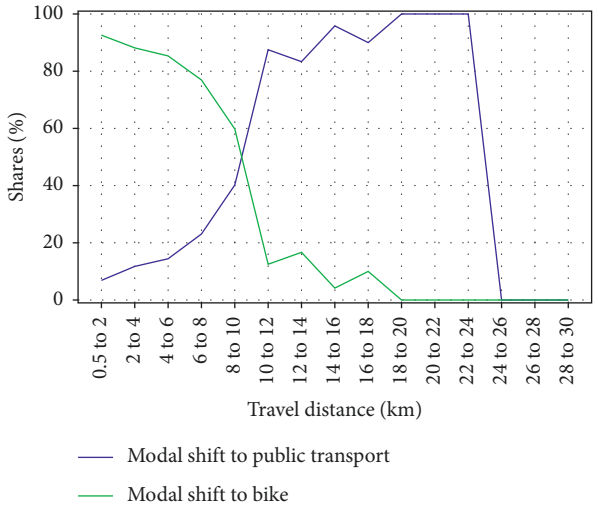
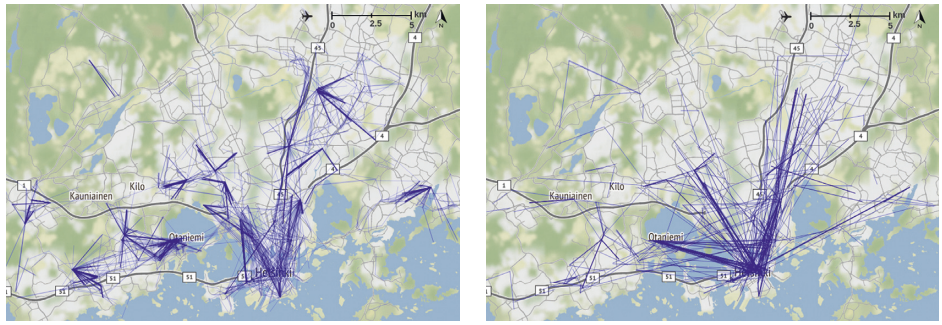


Figure 10: Share of each group of low-carbon modes in potential time-relevant modal shift from car. (a) To bike. (b) To public transport.

4.2. Emission Savings and Increase in Active Travel. This section compares the computed alternative trips with the observed car trips in order to quantify potential carbon emission savings as well as increased nonmotorized distance (i.e., with active travel modes of bike and walk) if travelers shift from car to the low-carbon alternatives. Total size of potential emission savings is 1,645 CO_2 kg, that is, 8.4% emission reduction for all Helsinki trips ($|R| = 25{,}328$) with mean 0.60 CO_2 kg per car trip. It should be noted that 1 kg of CO_2 equals to the amount of emission from energy usage of an average household for 40 minutes. Figure 14 illustrates the size and distribution of per-trip emission saving. Figures 15 and 16 also illustrate emission saving as well as increased distance traveled by cycling and walking as a result of mode change. The figures show the range of changes grouped by the suggested alternative mode. For example, when bus is the low-carbon alternative, per-trip emission savings range from 0.08 to 1.5 CO_2 kg, where 75 percentile is between 0.08 and 0.77 CO_2 kg. Among the motorized low-carbon potentials, train involves a wider range of emission saving as well as a wider range of increased active travel. Figure 17 illustrates the correlation between active travel and emission dimensions, where each circle represents one door-to-door car trip. At least two clusters of trips can be seen in Figure 17(a), one with higher amount of emissions reduced and another one with higher amount of active travel increased. As illustrated in Figure 17(b), the former cluster includes 62% bus and 38% rail-based modes as alternatives that both naturally include some access/egress walking too. The latter cluster comprises almost 100% bike alternatives.



(a) (b)

Figure 11: Desire lines of ODs of car trips with potential low-carbon time-relevant alternatives, when the alternative is (a) bike and (b) PT (map tiles by Stamen Design, under CC BY 3.0; data by © OpenStreetMap, under CC BY SA).

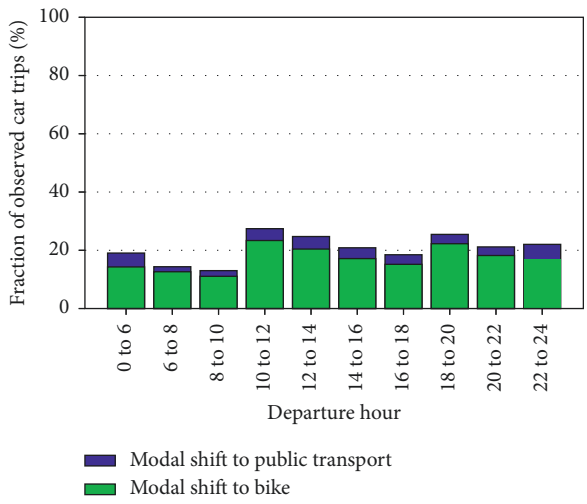


Figure 12: Observed car trips with a time-relevant low-carbon alternative, depending on departure time during the day.

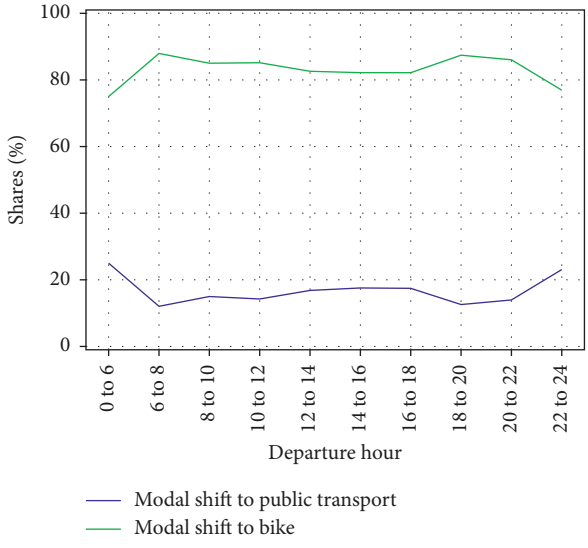


Figure 13: Share of each group of low-carbon modes in potential time-relevant modal shift from car during the day.

4.3. Person-Based Analysis. Section 4 so far presented the viewpoint of the whole trips recorded by all participants in the study. This section in turn focuses on the individual travelers. As explained in the methodology, a valid person-based analysis is made only based on the "active participants" who have a minimum of 30 "active days." This filtering results in 68 active participants out of the total 135 registered participants. Figure 18 shows the distribution of recorded trips among active participants. These 68 travelers have recorded $|R| = 24{,}377$ trips, that is, on average around 360 trips per person. Therefore, 96% of all trips in the dataset belong to the active participants. The mean share of each active participant from all recorded trips is 1.47% with standard deviation of 1.68. The cumulative distribution of shares is exponential and, for example, denotes that 20% of active participants have recorded 60% of trips. Figure 19 illustrates the active participants by the date span of recorded trips as well as their number of active days.

For these active participants, $|\mathrm{TRS}| = 2{,}693$ out of the 13,056 car trips, that is 21%, have the potential of modal shift. On average, each active participant has 23% of their car trips with a low-carbon time-relevant alternative. Figure 20(a) illustrates frequency and volume of total emission savings per active participant if they chose the possible lower carbon alternatives for all their car trips. By shifting to low-carbon mode, a total of 1,608 CO_2 kg is saved, that is, 8.2% emission reduction for all Helsinki, and on average 24 CO_2 kg per participant. Figure 20(b) illustrates the correlation between active travel and emission dimensions, where the same two clusters of PT and bike alternatives are seen as presented before in Figure 17.

Figure 21(a) illustrates the range and distribution of the modal shifts among the 68 active participants. Figure 21(b) shows the complementary cumulative probability distribution (CCDF) of the modal shift percentage per person, showing, for example, that half the participants can find a low-carbon alternative for at least 20% of their car trips. In addition, based on Figure 21(a), we could focus for instance on the top quartile of active participants (i.e., 13% of all travelers) who have the most substitutable car trips and perform further analysis from there.

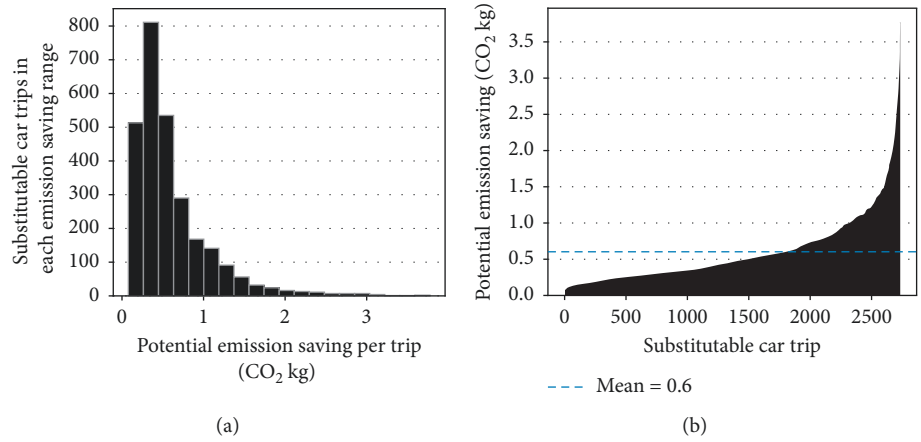


Figure 14: Size and distribution of potential emission savings among door-to-door car trips with alternatives.

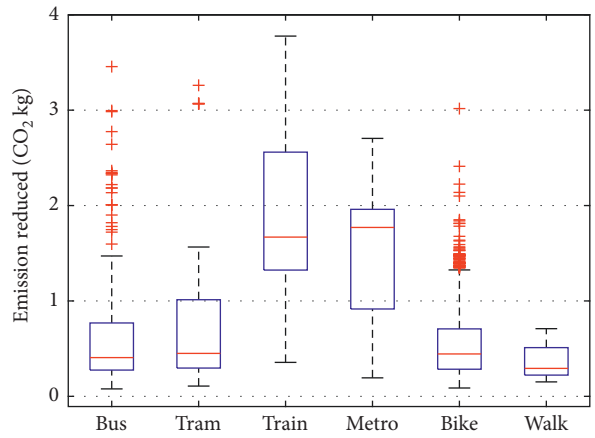


Figure 15: Size range and distribution of potential per-trip emission saving if travelers adopt low-carbon alternative modes. Red marks denote the statistical outliers.

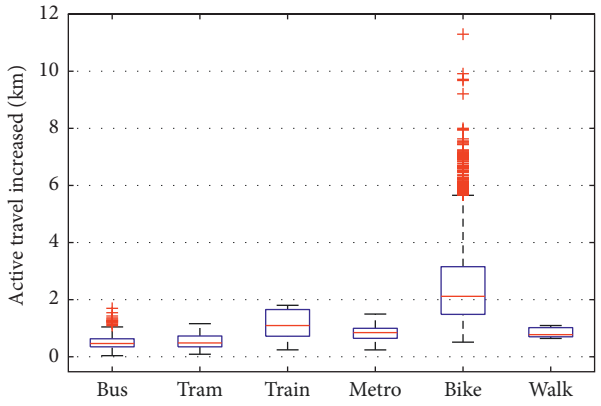


Figure 16: Size range and distribution of potential per-trip changes in active travel if travelers adopt low-carbon alternative modes. Red marks denote the statistical outliers.

4.4. Frequently Observed Unique Trips. So far in this paper, every single observed door-to-door trip together with its computed alternatives was counted against the aggregated numbers and figures presented above. However, clustering of trip ODs by their proximity shows that many trips of the same participant have almost the same origin and destination, thus being unique trips repeated several times. This section focuses on these unique trips that each traveler frequently takes. Figure 22 shows the size and distribution of unique trips among active participants. For the 68 active participants, we have observed a total of $|R| = 9{,}313$ unique trips, that is, on average 137 unique trips per participant, and each unique trip repeated on average 3 times.

There are 5,017 unique observed car trips that are on average 9 km long and took on average 22 minutes. 86% of these unique car trips have less than 3 minutes variability in travel time among the repetitions. Figure 23 evaluates consistency of modal shift possibility among the unique car trips. As seen in the histogram, from these unique car trips, 17% have always had a time-relevant low-carbon alternative, while 3% had an alternative only on average half of the times, and 80% never had an alternative. The conclusion is that the majority of the frequent car trips with any low-carbon alternative consistently have the alternative.

Figure 24 illustrates the 17% of unique car trips that have consistent possibility of modal shift. These repeated car trips are on average 3.5 km long and take 15 minutes on average. In addition, 95% of these car trips have less than 3 minutes variability in car travel time. They can be substituted 81% of the times with bike and 10% of the times with bus and 9% of times with rail-based transport (i.e., tram, subway, and train). Departure time of these trips is distributed during the whole day. Figure 25 shows modal shift and the potential savings. The result of mode substitution is 0.5 CO_2 kg of emission saving on average per single trip.

Focusing on individual travelers, it is seen that, among the 68 active participants, 94% have at least one car trip with a low-carbon alternative. As illustrated in Figure 26, on average, 20% of each participant's unique car trips have a low-carbon time-relevant alternative.

4.5. Potential of Substituting Noncar Trips with Low-Carbon Alternatives. So far in Section 4 we discussed car trips with possibility of shifting to low-carbon alternatives. Car trips

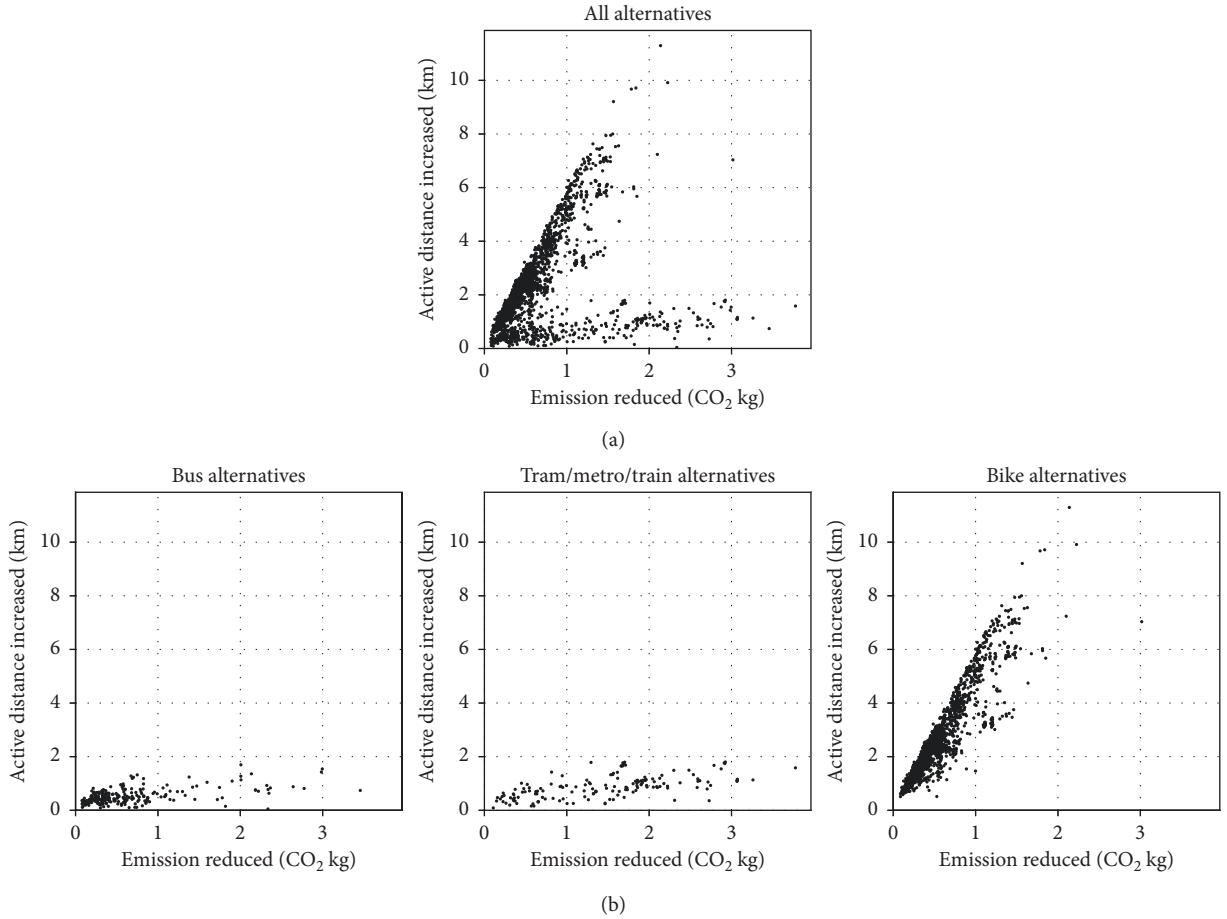


Figure 17: Potential emission savings and active travel increase if travelers shift from car to low-carbon time-relevant alternatives. Each point represents one door-to-door trip. (a) All car trips with any low-carbon alternative. (b) Car trips grouped based on the type of their alternative mode.

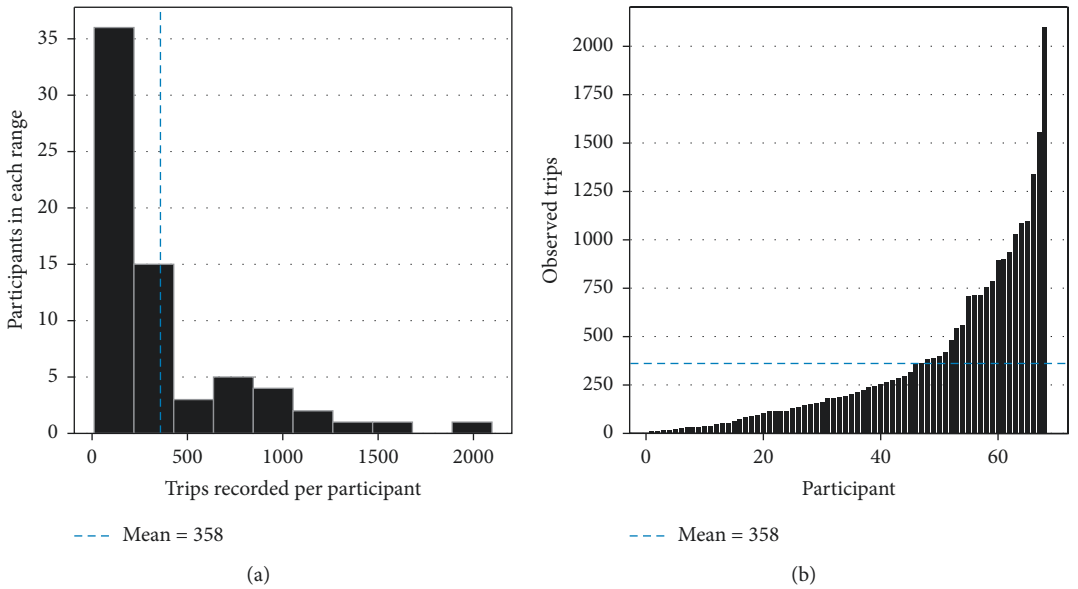


Figure 18: Distribution and share of trips recorded by active participants.

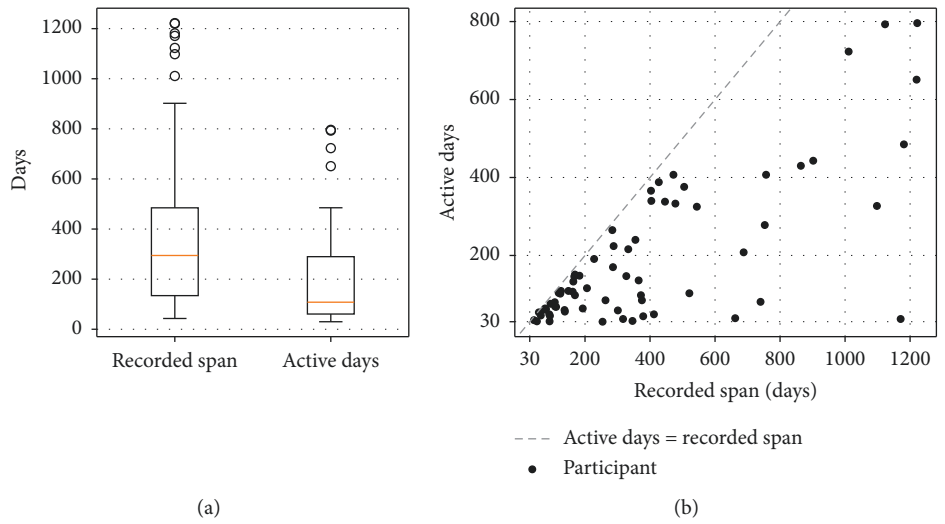


Figure 19: Range and distribution of travelers by their recorded date span and the number of their active days.

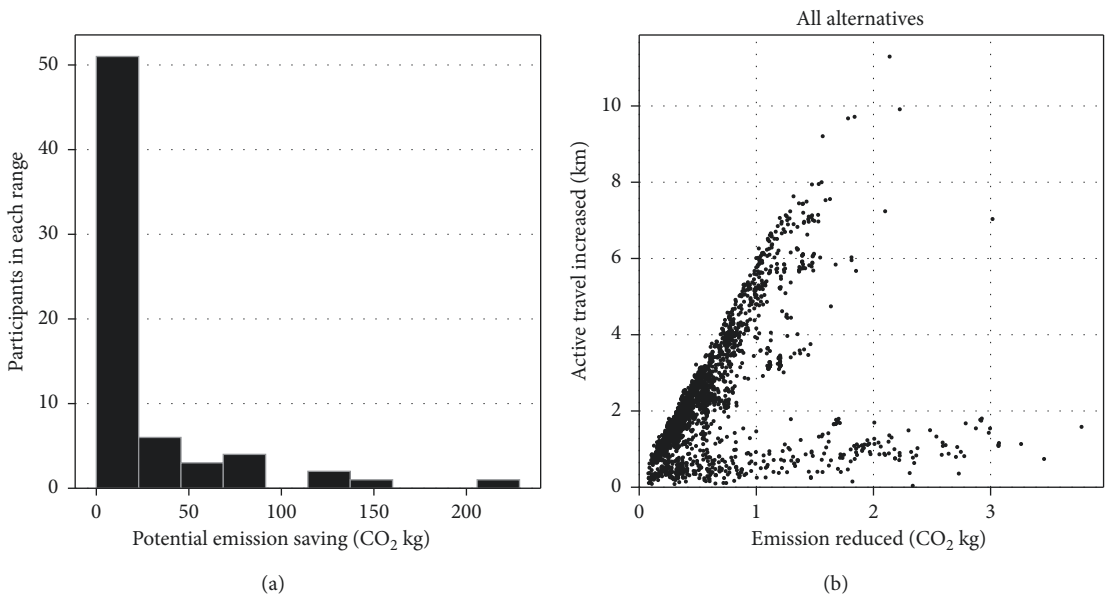


Figure 20: (a) Distribution of total emission savings per active participant. (b) Potential emission savings versus active travel distance increase, where each point represents one door-to-door trip of an active participant.

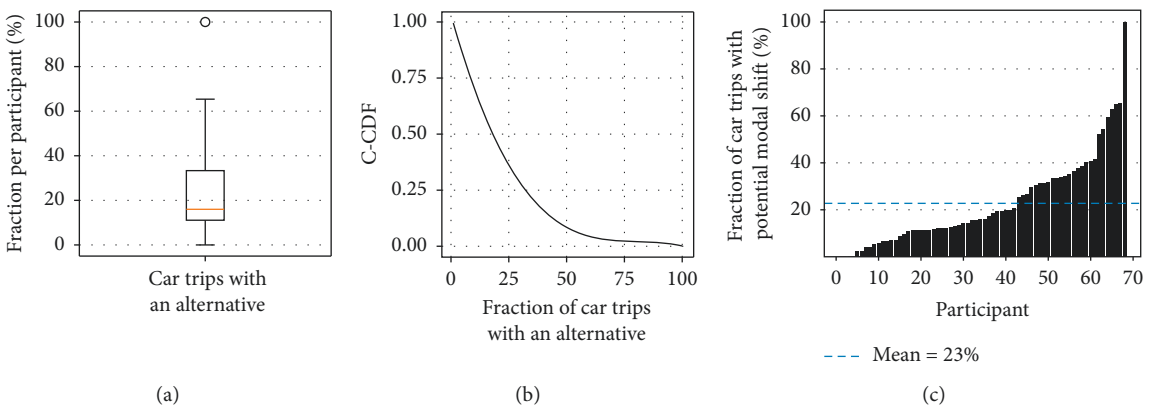


Figure 21: Range and distribution of per-traveler possibilities of time-relevant low-carbon modal shift.

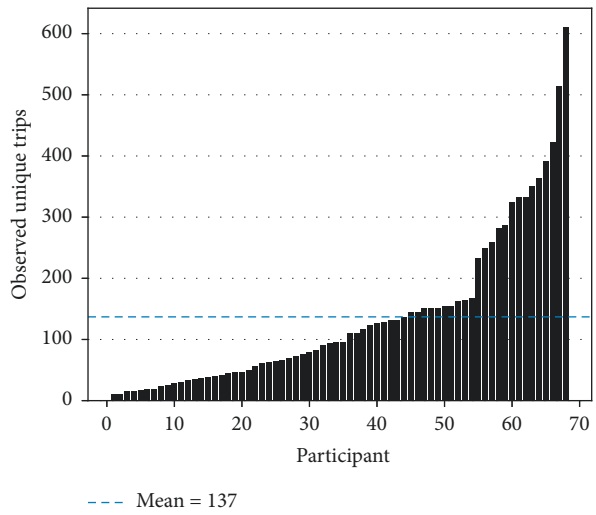


Figure 22: Share of unique trips among active participants.

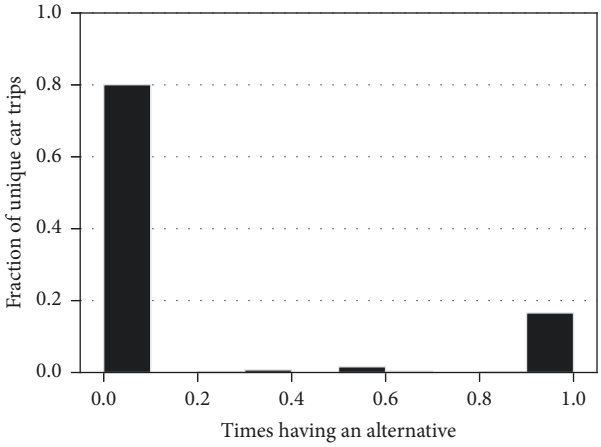


Figure 23: Consistency of having an alternative among unique car trips, illustrated by histogram.

make up around half of the total recorded trips. This section reviews the other half where the observed mode is not car but still has a time-relevant alternative with lower carbon emission. For example, there are bus trips that can be instead made by bike without losing time and thus saving emission. As seen in Figure 27, most of such trips are a shift from bus to bike and some also from the bus to rail-based transportation. There are 11,803 noncar trips from which 1,124 (10%) have a time-relevant lower carbon alternative, resulting in potential emission saving of 0.2 CO_2 kg per trip and 3.6 CO_2 kg per participant. Among the 4,012 observed PT trips, 18% have an alternative of bike. Figure 28 illustrates the emission savings.

5. Discussion

5.1. Discussion of Time-Relevant Low-Carbon Alternatives. Following are observations from analysis of Helsinki region and based on the trip data of participants who took part in the TrafficSense experiment. It should be noted that such implications may differ depending on the sampling design as well as the city where the travel data is collected. The previous studies based on a household mobility survey in Madrid have shown that 18% of reported car trips have a time-relevant alternative, with modal share among alternatives as 75% PT, 15% bike, and 10% walk [8, 9]. In comparison, our results showed that 20% of all observed car trips have an alternative, with modal share among alternatives as 16% PT and 84% bike. The model used for Madrid discards using PT or bike before 06:00 and after 22:00 due to personal safety concerns. Unlike such previous works, our method currently does not consider time of the day as a personal safety limitation while choosing the walk, bike, and PT alternatives. Test case for our method has been Helsinki metropolitan area, where with a few exception areas, traveling by bike or PT is considered safe in all times of the day. However, such limit can be considered if needed when our method is applied to mobility datasets of other cities.

Another set of studies based on a travel survey in Montreal has shown that 27% of reported motorized trips, that is, both PT trips and car trips, have a time-relevant alternative of walk or bike [27, 28]. In comparison, our results showed that 18% of observed PT trips have an alternative of bike. In comparison to our method, these studies do not compute PT trips as an alternative for car and do not consider how much the computed alternatives increase travel time compared to the observed trip. Their criteria for choosing a bike or walk trip as a car or PT substitute include travel distance as well as age of traveler and time of the day. Regarding the potential of cycling, personal willingness and physical ability may limit cycling only to shorter trips and relatively flat terrain [42]. In this regard, the maximum distance people are willing to bike is a subjective value and varies from city to city. For example, Morency et al. [27] concluded that the maximum threshold distance for cycling in Montreal is on average 5.4 km. Although our method does not consider a maximum distance, the cumulative distribution of travel distances shows that 90% of the computed time-relevant bike choices are already shorter than 5.6 km.

5.2. Usefulness. Experimental results from case study of Helsinki region presented examples on how our framework can be used to explore individual and aggregated traveler behavior. One of the method's main advantages is longitudinal evaluation of potential of time-relevant low-carbon mobility. As discussed before, long-term evaluation of traveler behavior is difficult with previous data collection methods. The analysis outcome can be used to identify potentials for improvement towards sustainable transportation systems, especially if data collection with mobile apps becomes more popular among travelers. Policy makers need to have numerical and visual data on measures such as ODs, travel route, travel time, departure/arrival times, and traveler modal choice. Although conventional data collection and analysis methods and tools such as surveys and GIS software provide such information, applying our framework with smartphone-based datasets enables higher resolution in both raw and refined data as well as the possibility of travel statistics on the level of individual traveler as presented in

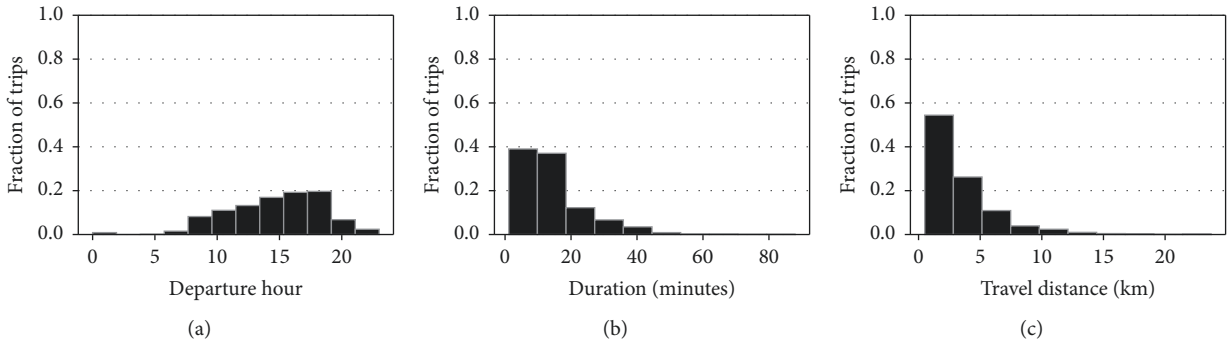


Figure 24: Attributes of the unique car trips that consistently have low-carbon time-relevant alternatives, illustrated by histograms.

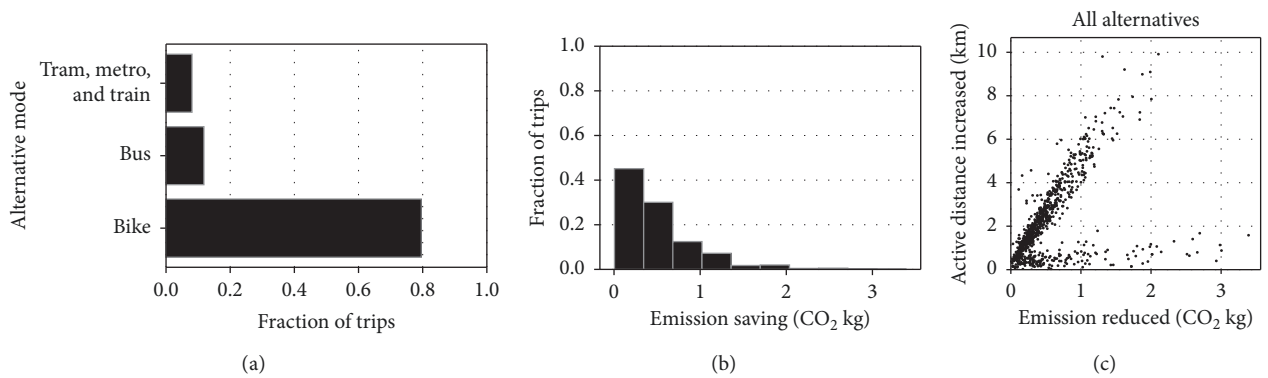


Figure 25: Modal shift from the unique car trips with consistent low-carbon time-relevant alternatives. Emission savings and increased travel distance of bike and walk.

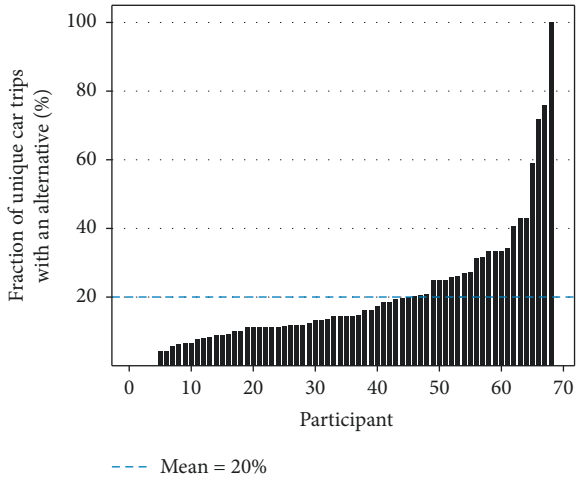


Figure 26: Potential of time-relevant modal shift from unique car trips among active participants.

Section 4.3. In addition, it is possible to focus more specifically on travelers who show higher potential of modal shift, for example, those travelers in the top quartile of potentials distribution shown before in Figure 21(a). In addition, analysis can investigate demographics and socioeconomic situation of travelers in this focus group, as well as whether people in the focus group have limited access to PT and bike, disability, or other reasons that keep them from changing from car usage. After the data collection and analysis phase, there can be a corresponding policy development phase, after which, it might be required to reach the local residents again in order to materialize the modal shift potential through different persuasive methods. With smartphone-based methods such as TS, a permanent contact point to local residents is established at the time when traveler installs the mobile app. Further notifications, feedback, and encouragements can be sent to users from this mobile app channel. Moreover, the contact remains anonymous when only the mobile app channel is used for communication to travelers, whereas in the survey-based approach, it can be challenging to find and reach the respondents again.

5.3. Accuracy and Noise in Sampling and Results. Previous literature has extensively studied accuracy and noise in smartphone-based data collection, showing that locational accuracy of GPS data points depends on factors such as clear sky view, holding the phone in hand, pocket, or attached in the bike or car, and the phone model. For example, experimental results for cycling along a 2.5 km inner-city bicycle track show maximum inaccuracy of 5 meters in most cases (95 percentile) and 20 meters in the worst case scenarios [43, 44]. In addition, in our own experiment, we store the accuracy of each sampled

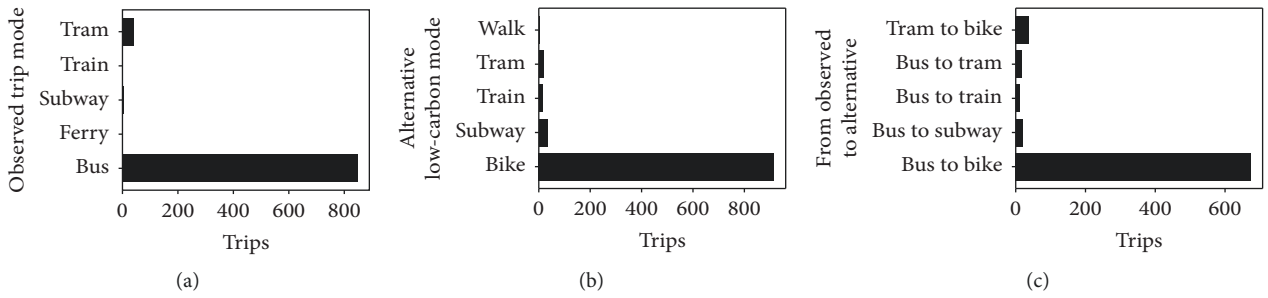


Figure 27: Distribution of the noncar trips that have low-carbon time-relevant alternatives, together with distribution of their low-carbon alternatives.

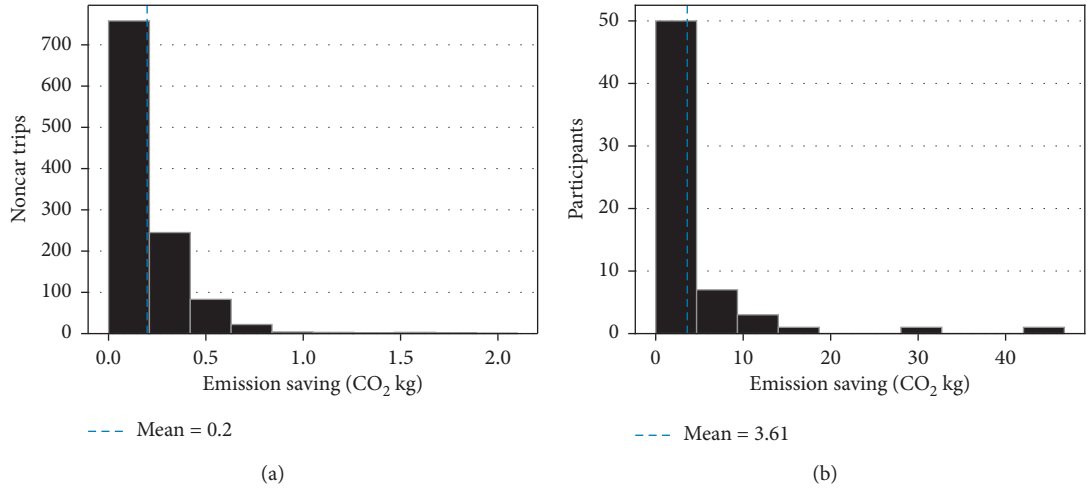


Figure 28: Size of emission savings illustrated per trip as well as per participant.

GPS data point as estimated by the Google API that is used to get phone locations. The mean accuracy of all sampled data points is 2.72 meters, and the mean accuracy per participant's smartphone is 5 meters. To tackle noise, the TrafficSense app discards sampled points that have an accuracy of worse than 50 meters. In future works, methods such as Google's Roads API [45] can be used to match the GPS points onto the actual road networks in order to achieve even a higher locational accuracy of vehicle traces. In addition, further data processing is needed to extract parameters such as transportation modes, traveler's activity locations, and start/end of each trip leg from the collected point data. To address questions of accuracy in this context, TrafficSense considers a 100 meters threshold to match the GPS traces to the expected path of scheduled public transport vehicle during mode detection of motorized public transport such as bus [33]. Similarly, Hemminki et al. and Shin et al. [46, 47] developed automatic transportation mode detection by employing various phone sensors such as accelerometer and GPS, together with statistical analysis and classifier training. In addition, Jiang et al. [48] proposed methods to classify daily mobility networks and extract trip/stay sequences, while Du and Aultman-Hall [49] developed a method for a more accurate identification of trip starts and trip ends.

Furthermore, to compute accurate alternative trip paths, we have used the OpenTripPlanner (OTP) API [50]. OTP is a well-established open-source software platform among the multimodal trip planning models. It is also used by some cities, such as for the online trip planning portal of Helsinki, Reittiopas [51]. OTP suggests routes based on real maps, road network, and up-to-date cycling and walking paths. For trip planning, it uses a single time-dependent graph containing both street and PT networks. Street network data are provided by OpenStreetMap (OSM), and PT schedules and route data are provided by General Transit Feed Specification (GTFS) files that are created and updated by cities or transportation agencies. OTP's routing API computes walk and bike trips generally using the A-star algorithm with a Euclidean heuristic [52]. PT trips including their walk legs and transit-ride legs are computed using multiobjective A-start with Tung–Chew heuristic [53] for queue ordering. OTP uses its internal algorithm together with OSM road network to estimate travel time [50].

5.4. Limitations and Future Work. Here, we discuss technological and data limitations together with future research directions. As expected, there are challenges during the recruitment and engagement of participants. This research has presented a proof of concept, targeting future more uniform and wider travel data. Nevertheless, as mentioned

in Section 3, 69 out of the 135 participants have completed an optional anonymous questionnaire. The questionnaire results include home postal codes, showing that participants come from various residential locations throughout the city. Moreover, the respondents come from all income categories as seen before in Figure 2. Among the respondents, there are 53 males and 16 females, which is (unfortunately) a common trend with first adopters. Future experiment could involve larger numbers of volunteers with more diverse demographic and socioeconomic profile.

Considering the privacy challenges in longitudinal smartphone-based data collection, we took the safest approach to preserve participant's privacy, with the intention to not process the trip information in a way that may give clues about identity of the individual participants. In addition, we have devised the methodology to minimize the user interaction. As a result, TrafficSense data collection is passive and anonymous, and we do not have exact information about residential/work location of participants in our database. Another engagement challenge is that there might be participants who travel several days a week, for a long period, and frequently repeat some of their unique OD trips, but do not necessarily keep the data collection app on during all the time. At the moment, there are no certain means to identify such "frequent travelers" that are not actively recording their trips. Finally, the quantity of the collected travel data affects how well our analysis reflects current travel behavior and potential of time-relevant mobility. Our computational framework has the potential of providing more precise results if we collect trip records from a more varied population of travelers. Such mobility dataset would contain travel data that are spatially more evenly distributed throughout different areas of the city. This can be achieved in future experiments by involving more volunteers with more diverse socioeconomic profile.

Access/egress walking legs of PT alternative trips are already considered in our computations. However, OTP at the moment does not take into account access-egress of walking to/from car and finding the parking space. In future work, by using complementary methods together with OTP, we could also estimate the time spent walking from/to car as well as the time spent searching for a parking spot. This could relatively increase potential of PT and bike as time-relevant options, as PT does not require parking search and cycling usually does not require any access/egress walk. Similarly, total travel time of a bike trip should ideally include the time spent to get dressed according to weather condition, for example, putting on/off the weather shielding clothing (i.e., water and wind proof jacket, trousers, gloves, and helmet). Currently, our method does not consider this time. In addition, it should be noted that cycling is not always a feasible choice for all travelers due to reasons such as not owning a bike, lack of access to bike sharing, or difficulty of biking because of bad weather conditions. Our computational framework is able to retrieve real-time weather information from local open-data services; therefore, in future work we can filter cycling trips based on weather conditions and for example rule out cycling on dates with heavy rain. Another point to discuss is the minimum 10 minutes idle time used in Section 2.1 as a threshold to identify trip start/ends in Helsinki region [23] and its relation to typical waiting times at intermediate PT stops. However, in some other cities, travelers might wait longer than 10 minutes for a bus or train. For this reason, a future work could test with slightly lower or higher threshold values and compare the identified door-to-door trips. As for carbon emission estimation, this paper considered tailpipe and not the total emission values. As the total carbon emission per vehicle-km of e-bike, electric train, tram, and metro is different in each region depending on electricity production and distribution network, further research is needed to obtain those values and use them in the framework.

The computational framework can be used for travel behavior analysis in any other city. Most parts of the source code and configurations would remain unchanged as the trip extraction, and analysis and visualization of alternatives are independent of city road network. Few changes might be needed such as obtaining the PT network and scheduling information of the target region or replacing the link to OTP API server with a local route planning service. In future research, similar methods can be integrated into previously developed persuasive web-based systems and smartphone apps that utilize personalized feedback and gamification [54, 55]. For example, the computed time-relevant alternatives can be recommended to travelers in a way that public transportation and cycling are gradually perceived as better choices not only because of their environmental or health benefits, but also as competitors to car in saving travel time. In future research, the proposed framework could be used to observe and evaluate actual changes in travel behavior of participants, for example, at the time when a new policy is being implemented. Examples are evaluation of new transportation services such as Western extension of Helsinki region metro system, the new city bike sharing, and assessment of new types of transportation ecosystems such as mobility-as-a-service (MaaS) [56–59].

6. Conclusion

This paper presented formulation, development, and evaluation of a computational framework for comparing observed travel behavior with computed low-carbon alternative trips in order to estimate the extent to which modal shifts could reduce carbon emissions without a significant compromise in travel time. The framework makes use of the emerging smartphone-based travel datasets and presents quantitative results and visualizations from a set of temporal, spatial, per-traveler, and whole-city viewpoints. For instance, we illustrated the estimated modal shifts, emission savings, and active-travel growth, as clustered by suggested alternative mode, departure time, trip distance, and spatial coverage throughout the city. The framework also estimated potential changes for trips frequently repeated by the same travelers. Furthermore, we have explained the lessons learned, limitations, and implications for future work.

In this paper, we evaluated the framework with long-term trip data of sampled travelers in Helsinki metropolitan region, Finland. The results showed that, for instance, on

average, 23% of car trips of each traveler had a lower carbon alternative and half the travelers had lower carbon alternatives for at least one-fifth of their car trips. Had the preferred alternative been chosen, about 8% of the carbon emissions could have been saved. Among the frequent unique car trips that had lower carbon alternative, 85% consistently had this possibility regardless of variations in travel time. Frequent car trips could be substituted 10% of the times with bus and 9% of the times with rail-based transport (i.e., tram, subway, and train). Chances of modal shift were higher for trips less than 20 minutes long and shorter than 8 km. In addition, the spatial potential of bike as an alternative was much more sporadic throughout the city compared to that of bus, which has relatively more trips from/to city center. Moreover, among the observed noncar motorized trips, 10% of public transport trips could also be substituted with bike and walk. This experimental evaluation showed usefulness of the method for exploring time-relevant low-carbon potentials. However, a large number of lower carbon alternatives were bike trips that may not be suitable for all participants and in all weather conditions. The size of any realized gain will depend on these factors as well as the regional quality of public transportation and bike paths.

The framework could be used for different cities, with changes such as linking to the local public-transportation information, considering regional cycling conditions, and adjusting the minimum activity-location idle time. Analyzing different cities would provide possibly different quantitative measures. In particular, the framework has the potential to provide more precise evaluations, when used with sizable data from larger population of volunteers with more diverse socioeconomic profile. Having such thorough datasets, the framework could provide implications for transportation researchers and planners to identify groups or areas for promoting mode shift.

Data Availability

A part of the data used to support the findings of this study is available from authors upon request. Shareable data include database schema and data table structures, as well as data records with aggregation level higher than individual traveler. Full individual trip records and traces cannot be shared due to research terms and GDPR regulations.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research was partly funded by the TrafficSense Project as part of the Aalto Energy Efficiency Research Programme and FINEST Twins Center of Excellence.

Supplementary Materials

Additional explanations about computational architecture, requirements, and software design. Figure 1: the main needs of framework. Table 1: needs, system requirements, and its implementation. Figure 2: system requirements. Figure 3: components of the whole system including both the original TrafficSense system as well as our time-relevant analysis method. Implemented contributions of this paper are highlighted in bold. Figure 4: screenshots of Traffic Sense mobile app, where traveler's route and mode of transportation is shown (map data copyright of Google). If needed, traveler can also click on each trip leg to revise the automatically detected modes. Figure 5: a door-to-door trip in Espoo, Finland, extracted by TrafficSense software from the collected travel data and illustrated using Python and Google Maps API. (*Supplementary Materials*)

References

[1] B. Gardner and C. Abraham, "What drives car use? A grounded theory analysis of commuters' reasons for driving," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 10, no. 3, pp. 187–200, 2007.

[2] R. Mitra, R. A. Ziemba, and P. M. Hess, "Mode substitution effect of urban cycle tracks: case study of a downtown street in Toronto, Canada," *International Journal of Sustainable Transportation*, vol. 11, no. 4, pp. 248–256, 2017.

[3] A. De Witte, J. Hollevoet, F. Dobruszkes, M. Hubert, and C. Macharis, "Linking modal choice to motility: a comprehensive review," *Transportation Research Part A: Policy and Practice*, vol. 49, pp. 329–341, 2013.

[4] J.-W. Grotenhuis, B. W. Wiegmans, and P. Rietveld, "The desired quality of integrated multimodal travel information in public transport: customer needs for time and effort savings," *Transport Policy*, vol. 14, no. 1, pp. 27–38, 2007.

[5] P. L. Mokhtarian and C. Chen, "TTB or not TTB, that is the question: a review and analysis of the empirical literature on travel time (and money) budgets," *Transportation Research Part A: Policy and Practice*, vol. 38, no. 9-10, pp. 643–675, 2004.

[6] T. Shannon, B. Giles-Corti, T. Pikora, M. Bulsara, T. Shilton, and F. Bull, "Active commuting in a university setting: assessing commuting habits and potential for modal change," *Transport Policy*, vol. 13, no. 3, pp. 240–253, 2006.

[7] J. Kim, K. Choi, S. Kim, and S. Fujii, "How to promote sustainable public bike system from a psychological perspective?" *International Journal of Sustainable Transportation*, vol. 11, no. 4, pp. 272–281, 2017.

[8] A. Monzon and L. A. Vega, "How to measure the potential of transferring trips from car to public transport and soft modes—the case of Madrid," in *Proceedings of the 2006 European Transport Conference*, Strasbourg, France, 2006.

[9] A. Monzon, L. A. Vega, and M. E. Lopez-Lambas, "Potential to attract drivers out of their cars in dense urban areas," *European Transport Research Review*, vol. 3, no. 3, pp. 129–137, 2011.

[10] G. Beirão and J. A. Sarsfield Cabral, "Understanding attitudes towards public transport and private car: a qualitative study," *Transport Policy*, vol. 14, no. 6, pp. 478–489, 2007.

[11] R. Sharples, "Travel competence: empowering travellers," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 44, pp. 63–75, 2017.

[12] N. J. A. Van Exel and P. Rietveld, "Could you also have made this trip by another mode? An investigation of perceived travel possibilities of car and train travellers on the main travel corridors to the city of Amsterdam, The Netherlands,"

Transportation Research Part A: Policy and Practice, vol. 43, no. 4, pp. 374–385, 2009.

[13] P. R. Stopher and S. P. Greaves, "Household travel surveys: where are we going?" *Transportation Research Part A: Policy and Practice*, vol. 41, no. 5, pp. 367–381, 2007.

[14] E. Murakami and D. P. Wagner, "Can using global positioning system (GPS) improve trip reporting?" *Transportation Research Part C: Emerging Technologies*, vol. 7, no. 2-3, pp. 149–165, 1999.

[15] J. Wolf, M. Oliveira, and M. Thompson, "Impact of underreporting on mileage and travel time estimates: results from global positioning system-enhanced household travel survey," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1854, no. 1, pp. 189–198, 2003.

[16] K. Clifton and C. D. Muhs, "Capturing and representing multimodal trips in travel surveys," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2285, no. 1, pp. 74–83, 2012.

[17] J. Wolf, R. Guensler, and W. Bachman, "Elimination of the travel diary: experiment to derive trip purpose from global positioning system travel data," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1768, no. 1, pp. 125–134, 2001.

[18] L. Zhu, J. Gonder, and L. Lin, "Prediction of individual social-demographic role based on travel behavior variability using long-term GPS data," *Journal of Advanced Transportation*, vol. 2017, Article ID 7290248, 13 pages, 2017.

[19] X. Chen, X. Wan, F. Ding et al., "Data-driven prediction system of dynamic people-flow in large urban network using cellular probe data," *Journal of Advanced Transportation*, vol. 2019, Article ID 9401630, 12 pages, 2019.

[20] Y. Asakura and E. Hato, "Tracking survey for individual travel behaviour using mobile communication instruments," *Transportation Research Part C: Emerging Technologies*, vol. 12, no. 3-4, pp. 273–291, 2004.

[21] C. D. Cottrill, F. C. Pereira, F. Zhao et al., "Future mobility survey: experiences in developing a smartphone-based travel survey in Singapore," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2354, no. 1, pp. 59–67, 2013.

[22] S. Vlassenroot, D. Gillis, R. Bellens, and S. Gautama, "The use of smartphone applications in the collection of travel behaviour data," *International Journal of Intelligent Transportation Systems Research*, vol. 13, no. 1, pp. 17–27, 2015.

[23] S. Jiang, G. A. Fiore, Y. Yang, J. Ferreira Jr., E. Frazzoli, and M. C. González, "A review of urban computing for mobile phone traces: current methods, challenges and opportunities," in *Proceedings of the 2nd ACM SIGKDD International Workshop on Urban Computing*, p. 2, Chicago, IL, USA, August 2013.

[24] J. Gadziński, "Perspectives of the use of smartphones in travel behaviour studies: findings from a literature review and a pilot study," *Transportation Research Part C: Emerging Technologies*, vol. 88, pp. 74–86, 2018.

[25] Z. Wang, S. Y. He, and Y. Leung, "Applying mobile phone data to travel behaviour research: a literature review," *Travel Behaviour and Society*, vol. 11, pp. 141–155, 2017.

[26] V. Raman, J. Shen, A. Kansal, V. Bahl, and R. R. Choudhury, "I am a smartphone and I know my user is driving," in *Proceedings of the 2014 6th International Conference on Communication Systems and Networks (COMSNETS)*, pp. 1–8, Bangalore, India, January 2014.

[27] C. Morency, H. Verreault, and A. Frappier, "Estimating latent cycling and walking trips in Montreal," in *Proceedings of the TRB 2017 Annual Meeting*, Washington, DC, USA, 2017.

[28] F. Godefroy and C. Morency, "Estimating latent cycling trips in Montreal, Canada," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2314, no. 1, pp. 120–128, 2012.

[29] A. Fyhri, E. Heinen, N. Fearnley, and H. B. Sundfør, "A push to cycling—exploring the e-bike's role in overcoming barriers to bicycle use with a survey and an intervention study," *International Journal of Sustainable Transportation*, vol. 11, no. 9, pp. 681–695, 2017.

[30] C. Gorenflo, I. Rios, L. Golab, and S. Keshav, "Usage patterns of electric bicycles: an analysis of the WeBike project," *Journal of Advanced Transportation*, vol. 2017, Article ID 3739505, 14 pages, 2017.

[31] "TrafficSense—research project at Aalto university," 2018, https://aaltotrafficsense.wordpress.com/about/.

[32] M. Heiskala, J.-P. Jokinen, and M. Tinnilä, "Crowdsensing-based transportation services—an analysis from business model and sustainability viewpoints," *Research in Transportation Business & Management*, vol. 18, pp. 38–48, 2016.

[33] M. Rinne, M. Bagheri, T. Tolvanen, and J. Hollmén, "Automatic recognition of public transport trips from mobile device sensor data and transport infrastructure information," in *Proceedings of the International Workshop on Personal Analytics and Privacy (PAP 2017)*, vol. 10708, pp. 76–97, Dublin, Ireland, September 2017.

[34] "TrafficSenses software and mobile app—source code, documentation and setup instructions," 2020, https://github.com/aalto-trafficsense/.

[35] "Digitransit Finland," 2020, https://digitransit.fi/en/.

[36] Cynergy E-bikes, http://www.cynergyebikes.com/about-ebikes-s/117.htm.

[37] Bicycle Performance, https://en.wikipedia.org/wiki/Bicycle_performance.

[38] A. Tiwary and J. Colls, *Air Pollution: Measurement, Modelling and Mitigation*, CRC Press, Boca Raton, FL, USA, 2010.

[39] "LIPASTO—unit emissions (database of transport emissions in Finland)," 2017, http://lipasto.vtt.fi/yksikkopaastot/indexe.htm.

[40] "Jälki—weekday travel carbon calculator (for Helsinki metropolitan area)," 2017, http://www.hsljalki.fi/en/menu/info.

[41] Executive Office and Urban Research and Statistics, *Statistical Yearbook of Helsinki 2018*, Executive Office, Urban Research and Statistics, Helsinki, Finland, 2018.

[42] E. Heinen, B. van Wee, and K. Maat, "Commuting by bicycle: an overview of the literature," *Transport Reviews*, vol. 30, no. 1, pp. 59–96, 2010.

[43] M. Bierlaire, J. Chen, and J. Newman, "A probabilistic map matching method for smartphone GPS data," *Transportation Research Part C: Emerging Technologies*, vol. 26, pp. 78–98, 2013.

[44] M. Liebner, F. Klanner, and C. Stiller, "Active safety for vulnerable road users based on smartphone position data," in *Proceedings of the 2013 IEEE Intelligent Vehicles Symposium (IV)*, pp. 256–261, Gold Coast, Australia, June 2013.

[45] Google roads API. http://developers.google.com/maps/documentation/roads/snap.

[46] S. Hemminki, P. Nurmi, and S. Tarkoma, "Accelerometer-based transportation mode detection on smartphones," in *Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems—SenSys '13*, pp. 1–14, Rome, Italy, November 2013.

[47] D. Shin, D. Aliaga, B. Tunçer et al., "Urban sensing: using smartphones for transportation mode classification," *Computers, Environment and Urban Systems*, vol. 53, pp. 76–86, 2015.

[48] S. Jiang, J. Ferreira, and M. C. González, "Activity-based human mobility patterns inferred from mobile phone data: a case study of Singapore," *IEEE Transactions on Big Data*, vol. 3, no. 2, pp. 208–219, 2017.

[49] J. Du and L. Aultman-Hall, "Increasing the accuracy of trip rate information from passive multi-day GPS travel datasets: automatic trip end identification issues," *Transportation Research Part A: Policy and Practice*, vol. 41, no. 3, pp. 220–232, 2007.

[50] OpenTripPlanner bibliography. http://docs.opentripplanner.org/en/latest/Bibliography/.

[51] "Reittiopas (trip planner for Finland)," 2017, https://www.reittiopas.fi/.

[52] L. Mandow and J. L. P. De La Cruz, "Multiobjective A^* search with consistent heuristics," *Journal of the ACM*, vol. 57, no. 5, pp. 1–25, 2010.

[53] C. Tung Tung and K. Lin Chew, "A multicriteria Pareto-optimal path algorithm," *European Journal of Operational Research*, vol. 62, no. 2, pp. 203–209, 1992.

[54] V. Sunio and J. D. Schmöcker, "Can we promote sustainable travel behavior through mobile apps? Evaluation and review of evidence," *International Journal of Sustainable Transportation*, vol. 11, no. 8, pp. 553–566, 2017.

[55] J. Jariyasunant, M. Abou-Zeid, A. Carrel et al., "Quantified traveler: travel feedback meets the cloud to change behavior," *Journal of Intelligent Transportation Systems*, vol. 19, no. 2, pp. 109–124, 2014.

[56] C. Weckström, R. Kujala, M. N. Mladenović, and J. Saramäki, "Assessment of large-scale transitions in public transport networks using open timetable data: case of Helsinki metro extension," *Journal of Transport Geography*, vol. 79, Article ID 102470, 2019.

[57] N. Haglund, M. N. Mladenovic, R. Kujala, C. Weckstrom, and J. Saramaki, "Where did Kutsuplus drive us? Ex post evaluation of on-demand micro-transit pilot in the Helsinki capital region," *Research in Transportation Business & Management*, Article ID 100390, 2019 In press.

[58] C. Weckström, M. N. Mladenović, W. Ullah, J. D. Nelson, M. Givoni, and S. Bussman, "User perspectives on emerging mobility services: ex post analysis of Kutsuplus pilot," *Research in Transportation Business and Management*, vol. 27, pp. 84–97, 2018.

[59] K. Pangbourne, M. N. Mladenović, D. Stead, and D. Milakis, "Questioning mobility as a service: unanticipated implications for society and governance," *Transportation Research Part A: Policy and Practice*, vol. 131, pp. 35–49, 2019.