A Context- and Trajectory-Based Destination Prediction of Public Transportation Users

Malte Bieler
Is with the Department of Technology, School of Economics, Innovation and Technology, Kristiania University College, Oslo, 0152, Norway. Email: malte.bieler@kristiania.no.

Raghava Rao Mukkamala
Is with the Centre for Business Data Analytics, Department of Digitalization, Copenhagen Business School, Frederiksberg, 2000, Denmark and the Department of Technology, Kristiania University College, Oslo, 0152, Norway. Email: rrm.digi@cbs.dk.

Tor-Morten Grønli*
Is with the Department of Technology, School of Economics, Innovation and Technology, Kristiania University College, Oslo, 0152, Norway. Email: tor-morten.gronli@kristiania.no.

xxxxxxx

Digital Object Identifier 10.1109/MITS.2021.3132772
Date of current version: 8 January 2022

*Corresponding author

1939-1390/22©2022IEEE

***Abstract*—**Location prediction based on contextual information is the core of a plethora of personalized location-based services (LBSs). Several applications require the use of techniques for predicting travel destinations based on human movement. Network analyses of human behavioral data show how the spatial-temporal regularity of human movement can be harnessed for inferring human mobility patterns. However, techniques are often based on a limited number of contextual features, which may limit prediction accuracy, especially if only historical location data are used. Using movement data obtained from public transportation users, we investigate the utility of contextual features derived via the installation of Bluetooth beacons in transportation vehicles and software tools in end-users' travel applications. Using a multiclass random forest classifier, we show that contextual information of a user's past travel history and at journey onset goes beyond spatial information and boosts destination prediction accuracy. The likely destination and travel-path length obtained at journey onset can then serve as the input for a stochastic-based model to predict a destination based on acquired trajectory information. Here we show that previously predicted destinations boost the performance of a Markov chain network. Thus, various contextual information at the start of a journey provides information beyond the location information acquired during a journey's progression that can be employed for destination prediction. These findings have strong implications for LBSs as they require accurate destination prediction at early stages of a journey while at the same time mitigating the privacy concerns associated with collection of location data.

The use of sensors, advanced communication technologies, and high-speed networks are collectively setting new standards in intelligent transportation systems (ITSs). These technologies impact how people move among places in urban environments, and thus shape human mobility profoundly. In addition, while not only shaping human movement, these technologies allow for observation and examination of human movement in R&D. The advantages that these smart technologies entail for transportation are numerous. For example, Internet of Things data can be processed and exploited using techniques such as machine learning to reveal information on customer behavior to support traffic management.

Concurrent to the emergence of the aforementioned technologies, advances in mobile computing and positioning technology have paved the way for the distribution of real-time, personalized services in smart cities as well as the collection and management of a vast amount of human movement data [1]–[3]. Such data can be utilized, among others, in various domains of smart cities, including urban planning, power grid utilization, traffic control, and public health [4]–[6]. In particular, new possibilities are being extended through the general availability of mobile computing solutions and personal smartphones especially. This opens the possibility for a reliable stream of live data from citizens.

Sensor-rich mobile devices incorporate much of a user's external surroundings. Together with the sensor network of a city, this creates a digital footprint of the user's experienced context and is possibly a manifestation of his or hers intrinsically motivated behavior. Here, we refer to the *concept of context* as entailing everything that surrounds a user and gives meaning to or influences behavior. Among the contextual data used for personalized applications, location information has received the greatest interest by service providers as well as academic research [7]–[10], leading to rapid advances in the realm of location-based services (LBSs), ranging from navigation and social networking to personalized recommendation services. These, in turn, have become catalysts in the growth of areas such as telecommunications and transportation.

Location information is critical to understand and anticipate human behavior [11]. Pervasive mobile devices and public sensor infrastructure simplify the tracking of a person's movement, represented as a sequence of time-stamped locations [12], [13]. Research has shown that despite individual randomness, human movement is highly regular [14], [15]. This regularity in human movement can be harnessed to investigate the causes of, and make predictions about, human movement in urban environments [16], [17], which can be used in areas such as in city development or route planning [4]–[6], [18]. However, although spatial information is often used for destination predictions, fewer approaches utilize temporal information and fewer still apply other types of contextual information [19].

The specific purpose of transportation is to fulfill a demand for mobility as transportation can exist only if it moves people. Individual spatial-temporal information can potentially be gathered and shared continuously when people make use of public transportation. This would lead to valuable insights on mobility patterns and user preferences, which can be used for location prediction and recommendations, respectively [20].

Rodrigue et al. [21] propose that the mobility of people can be expressed by accessibility, i.e., the capacity of a location to be reached by or to reach different locations. Accessibility is determined by the level of development of a transportation system to support mobility. Thus, highly accessible places are more frequented by people than less-accessible locations. Overall, a well-developed system is linked with an array of opportunities at economic and social levels.

Location and distance are core concepts of accessibility. Location enables the estimation of the relativity of places in relation to the transportation infrastructure, which supports human movement. Similarly, distance, the connectivity between locations, exists if two locations are linked by means of the transportation infrastructure. Distance expresses the friction of space and locations, with the least friction being the most accessible. Overall, tracking individuals and obtaining contextual information related to accessibility enhance understanding of human mobility.

Location-based applications provide essential intelligence to business and governments; however, certain developments urge a diminishing reliance of personalized applications on users' location information. Combining ubiquitous positioning, motion recognition, and human behavior modeling, mobile devices are increasingly considered a cognitive mobile device for intelligent transportation. These technological advancements have led to greater consciousness on data privacy with respect to a user's willingness to share personal information for mobile applications, including those in intelligent transportation. Thus, despite the enormous potential of utilizing insights on human movement, the continuous tracking of people's location by means of pervasive technology raises concerns, both at the individual and national security level [22], [23]. In the long run, it is desirable to reduce nonessential tracking of individuals and tailor personalized applications to intermittent usage in cases when needed.

Here we investigate the importance and value of contextual information in destination prediction in public transportation. The prediction of a user's destination based on detailed knowledge of its environment, both at the onset of travel as well as during travel progression, constitutes the possibility for relevant personalized applications. To enhance privacy and reduce the unnecessary long-lasting tracking of users' sensor data, we focus our efforts on the possibility of predicting a passenger's destination and possible route in a medium-size Scandinavian city solely on the available contextual information at departure and coarse GPS location information of the public transportation vehicle. The historical data used for destination prediction contains a unique device identifier and is stored for only up to six months. The data set entails only the contextual information available during boarding and disembarking. Privacy concerns may arise if information beyond and during the trip is obtained by means of a smartphone. However, no GPS data are obtained via a smartphone at any instance of the trip. Only coarse, real-time GPS data publicly available for all transportation lines is used, but collection of GPS data at the vehicle level consequently occurs only when on board.

We investigate the importance of contextual information of destination prediction accuracy by means of several approaches. First, previous research has shown that destinations can be derived from the spatial and temporal data at the beginning of the journey [3], [24], [25]. We go beyond this and investigate whether the correctness of destination prediction is enhanced by augmenting location and temporal information with further contextual data. Second, smartphone GPS data are a prime source for the tracking of people and predicting their future location. Here we restrict our destination prediction to location data at a coarse level that are solely available from public registries once a user boarded a public transportation vehicle. Finally, although research has often regarded destination and route predictions independently [26]–[29], we merge these two approaches by examining how trajectory prediction can be boosted by considering only the likely paths and destinations obtained from context-based destination prediction.

Related Work

Although controversy exists on the homogeneity of human movement (e.g., in [30]), much research indicates that the movement of people is defined by regular spatial-temporal patterns where each individual can be characterized by a set of defined routes and visits to a few locations [14], [15], [31]. In other words, a spatial probability distribution underlies the reproducible patterns of human movement, which can be harnessed for the design of predictive systems for destination prediction in public transportation [28], [32], [33].

Pervasive location-acquisition technologies often rely on GSM, Wi-Fi, or GPS records [34]. The latter has often been used as the main source to reconstruct and predict human movement trajectories. For instance, the authors in [35] make use of a person's history of visited locations and apply a Markov model to obtain probabilities on the likely path and destination of a user's travel. Similarly, in [26], Alvarez-Garcia et al. apply a hidden Markov model (HMM) based on historical data and predict a user's next location based on the start of a journey.

Providing data on spatial, temporal, and social features, location-based social networks enable the prediction of user mobility. Araújo et al. [20] introduced the a two-layer ensemble learner approach, the ensemble random forest Markov mobility prediction model, based on the random forest algorithm and Markovian property. The proposed model achieves a higher accuracy and F1-score when compared to other trajectory prediction models.

Generally, historical trajectories are used to mine human movement and develop methods for next-location prediction. Such methods can be based on trajectories from individual users or on all the available trajectories of larger groups examining collective movement. Chen et al. [16] addressed this dissociation by presenting three models that take the individual as well as the collective movement patterns into consideration, and acknowledge that movement patterns are time dependent. Hence, predictive models must account for temporal variations.

In addition to differences at the individual-versus-collective level, the mining of historical trajectories for destination prediction is generally associated with the data sparsity problem. The *data sparsity problem* refers to the fact that available historical data sets do not cover all the possible trajectories that a user may take. To address this issue, Xue et al. [25] developed a method named *subtrajectory synthesis (SubSyn)*. SubSyn decomposes historical trajectories into subtrajectories comprising two neighboring locations and connects these two subtrajectories into a new, merged "synthesized" trajectory. Thereby, the number of trajectories that can have predicted trajectories increases, and thus, the model predicts destinations for up to 10-times more query trajectories than a baseline algorithm. Krumm and Horvitz [29] developed a method called *predestination*, which uses a driver's history of destinations as well as the data collected during travel progression for destination prediction. In particular, although much work assumes a closed-world network of users not visiting new locations, the latter work leverages an open-world methodology, also taking into account the likelihood of users visiting previously unobserved locations based on trends in the data.

Generally, additional external information may improve the accuracy of predicted destinations. It is assumed that the external features of certain traveling patterns are associated with greater possibilities of reaching a certain destination than others. Although historical movement patterns may be useful to understand the movement of users, they do lack contextual information to better understand the intended activities. Such contextual information can be related to, for example, road conditions and driving habits [36], but temporal aspects such as time of the day and day of the week are also valuable contextual features that can complement historical trajectory data [3], [24]. Wu and Li [3] performed a detailed investigation of the problem of annotating contextual semantics to mobility patterns, i.e., associating the mobility records in a person's trajectory with relevant surrounding context. The authors used spatial context and annotated the venue a user was visiting at a defined location.

In addition to the use of probabilistic models, machine learning methods have been used to predict a user's destination. Recently, Zhao et al. [28] presented a bidirectional, long short-term memory network to model the sequence of visited locations. By implementing an attentional mechanism, meaningful locations of a route that have a strong relationship to the possible destination were defined and given more impact on the eventual prediction of the destination. Using various machine learning methods to predict the destination of a user based on smartphone GPS data using information such as location information 5 min prior to departure, time of the day, and the day of the week, the authors in [27] reveal that a decision tree-based algorithm proved to be the most accurate resulting in 96% accuracy in destination prediction.

Overall, the lack of data on identical trips, visits to different destinations under the same context, or a trajectory history of too-short time spans may limit prediction accuracy. Furthermore, prediction techniques will not be adequate if the starting location of a trip is a new, unique location, i.e., no historic data on movement history are available. To circumvent these difficulties, Huang et al. [37] investigated the relationship between activity and location changes. The authors demonstrate that activity transitions are more regular than location transitions and developed an HMM-based predicting approach, which takes users' activity transitions into account. Based on the next activity, the authors could derive the next location of a path in a much smaller candidate destination set, reaching 87% accuracy.

Similarly, Xia et al. [38] addressed the problem of location imprecision by adding semantic information added through stay-point detection and semantic-place recognition. The aim was to predict next-personally semantic places with historical visiting patterns derived from mobile device logs. Using a decision-tree-based algorithm and a Markov model, the authors examined differences in prediction accuracy and found that the decision-tree-based algorithm is superior to a Markov model.

Although previous studies often aimed at separating the destination- and route-prediction problem, Imai et al. [24] sought to converge these two techniques. The authors approached the tradeoff between timing and accuracy by combining path predictions using contextual information at the start of a journey with trajectory prediction. Thus, an initial prediction of the destination was made and updated as the trip progressed. This led to superior prediction accuracy when compared to applying techniques independently.

Thus, relying on the regular movements of people, various analytical tools have been used to derive predictions on future locations of public transportation users. From related work, we extract the insight that merging various approaches using contextual information is not only feasible, but also required, for optimal destination and route prediction.

Methodology

Scenario

In this section, we provide details on the scenario for the aforementioned research as well as our chosen approach for data collection and analyses. A middle-sized Scandinavian city (with an estimated 150,000 inhabitants) provides several modes of public transportation, including bus, train, tram, and city ferry. These public transportation vehicles are equipped with hardware transmitters that enable the assignment of users to a certain public transportation vehicle. This enables the collection of data regarding the user's transportation usage and can be supplemented with publicly available real-time data of the public transportation vehicle. The collected contextual data are used here to derive the probability of arriving at a certain destination based on historical travel patterns. Thus, at the early stages of travel, contextual information can be utilized to predict the destination of a public transportation user.

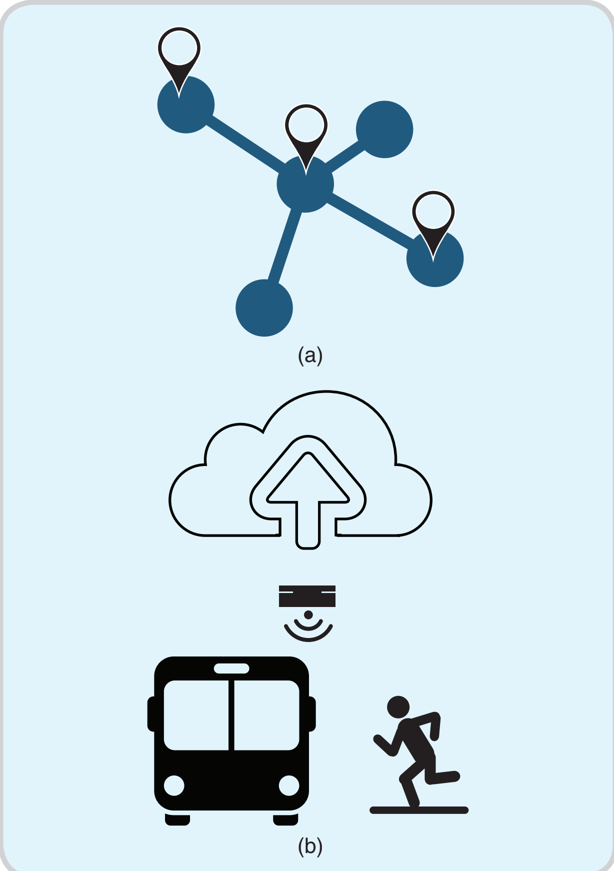

FIG 1 A scenario. The information available at the start of the journey can be obtained by linking a passenger to a public transportation vehicle by means of a software development kit embedded in a travel application and Bluetooth beacons, whose identified signal assigns a passenger to a transportation vehicle (b). The assignment of a passenger to a vehicle allows a more precise trajectory tracking by matching passenger information with real-time public transportation location data obtained from a government-owned transportation company in the cloud (a).

Data Collection

Public transportation vehicles are equipped with Bluetooth beacons that broadcast their identifiers to the mobile phone of a passenger in proximity. The passengers make use of a ticketing application with an installed software development kit that allows for the collection and sending of coarse contextual information. The origin and destination of a journey are given by the identification of users through the Travel Pass mobile app. The installed beacons on public transportation vehicles pick up signals from devices on board and ping them at regular intervals. The beacons are installed in the ceiling of the vehicles, and depending on the model at strategic places, for instance, in the front, middle, or back. The location and time of a journey's start can be derived from the beacon's signal strength. Thus, for the duration of the entire trip (given the strength of the beacon signal), a person is associated with a certain transportation vehicle and contextual information, such as the time and location of boarding and disembarking (see Figure 1).

Bluetooth is given access once customers have given consent in the travel application. As long as the user keeps Bluetooth turned on, the consent is valid and tracking is available. The user can at any time withdraw the consent, turn Bluetooth off, or request to have profile data deleted. Given the architecture of the travel application solutions, connectivity issues are generally not an interfering issue because information is collected locally on the phone and stripped of any personal-related information before being submitted to the cloud's back end for analysis. Within the cloud, passenger data are matched with more accurate location data (GPS) of public transportation stations recorded in a national registry for public transportation. From the largest to the smallest level, the country is divided into counties, municipalities, and basic statistical units (BSUs). The country is divided into approximately 14,000 BSUs, which are geographically coherent, homogeneous (with respect to the nature and basis for economic activities, conditions for communications, and structure of buildings), and temporally stable units, which are useful to work with and represent regional statistics. Overall, BSUs aim at a more efficient statistical basis for analysis on regional and municipal levels for management and planning purposes. Because the aforementioned conditions must be fulfilled, BSUs are measured in squared kilometers may vary in size. Notably, location data are represented by the presence of the passenger within a broader area of the municipality (i.e., BSU) rather than precise GPS coordinates.

The final data set thus contains location information of the journey's start and end location, the time stamp of the start and end of the journey, the transportation vehicle's ID, and station location data. Overall, we base our analyses on data collected from 26,088 customers during a period of approximately 175 days (Table 1). When compared to a publicly available data set comprising the number of travel searches served by an open national journey planner application programming interface acting as a back-end system in several major journey planning services (developer .entur.org/mobility-trends-covid-19), we find that daily travel counts of our data set and travel searches (as a proxy of national mobility) across the country are highly similar. Here we regard the quantity of travel route searches as a proxy for human mobility within the last months.

Data Preprocessing and Feature Engineering

As a part of the extract, load, and transform (ETL) process, raw passenger data are uploaded to a cloud database, where they undergo further prepossessing. A *journey* is defined as a set of trips with an intent. A journey can consist of multiple trips, which in total define the entire journey. Because we often cannot find the intent of a journey from location events alone, we use the time spent in a BSU to define a journey. When a user stays for more than 15 min in one place, we consider this the origin for the next journey. Naturally, we use the same time measure for the destination. Once a user arrives at a location and stays for more than the 15-min parameter, we consider this location to be the destination of the journey. Following the ETL process, the data are organized and cleaned using Python in a Microsoft Azure Databricks environment. Based on obtained passenger information and real-time vehicle data, we acquired several contextual features that are either available or can be achieved by processing the data set (see Table 2).

Data Analysis

We base our analyses on users out of the 26,088 who had a sufficient amount of movement within the period of data collection. This condition is met if two requirements are fulfilled. First, the user must have a journey count larger than the median journey count across passengers (i.e., 101 journeys in 175 days). Second, the selected journeys' destinations must have been visited several times (i.e., lay above the first quartile/0.05% of all a user's journeys that a specific destination was the final stage of the trip). This led to a data set of 3,002 users that use public transportation frequently.

Supervised Learning

We applied a supervised learning technique to predict the destination of a user based on contextual features of past trips. We used a tree-based learning algorithm: the multiclass random forest from Python's scikit-learn library. The random forest classifier is built using several decision trees in an ensemble method (a bootstrap aggregation, i.e., additional data in the training stage were generated to

Table 1. The descriptive statistics.

Statistics	Value
Time frame of analyzed data	23 February 2021–18 August 2021
Size of data set	497 MB
Number of customers	26,088 (20% of city population> age 10)
Average number of journeys per customer	38
Average distance of journey	4,918 m
Most-frequented day	Tuesday
Least-frequented day	Sunday

Table 2. A description of contextual features. Note that in the "Human Mobility Patterns for Destination Prediction" section, only "Hour of the Day," "Day of the Week," and "Journey Origin" are used for destination prediction.

Feature	Description
Hour of the day	Full hour of the day at the start of the journey
Day of the week	Day of the week on which the journey took place
Journey origin	The place, defined by the BSU, where the journey started
Distance	Linear distance as measured by the percentage of traversed diameter of the urban area
Transport mode	Type of vehicle used in the first trip of the journey: bus, train, or city ferry
Traveler type	Categorization of the passenger based on past transportation mode use
	If the passenger has only used a bus in the past, then the passenger is a bus-only traveler
	Several categories exist: bus only, train only, city ferry only, bus/train, bus/ferry, train/ferry, and bus/train/ferry
Outward/inward flow	Measurement that defines a departure and destination location/station as a station where more users depart from rather than arrive at or vice versa
Home-work location	Several conditions define the rough estimate of work and home location
	Home location: The start of the journey must take place between 5:30 and 9:30 a.m. The end location of a journey is the last visited location between 3:30 p.m. and 12 a.m. The end location of a journey must be the final destination in at least 33% of all journeys during the collected time period.
	Work location: The end location of a journey must be visited between 5 and 10 a.m. The previous end location must be a start location between 2:30 and 5:30 p.m. The end location of the journey must be a destination in at least 15% of the journeys in the collected time frame.
Hotspot	All stations whose total departure/arrival journey count density is larger than the average across all stations
Distance	Linear distance of the journey (in meters)

decrease the variance of the prediction model). A Gini index was used as an attribute selection indicator to generate individual decision trees. The most-voted class of all trees was chosen as the final solution. The data were split into an 80:20 training-versus-test set. The independent categorical variables (e.g., day of the week) were recoded into numerical representation without an arbitrary ordering by means of one-hot encoding. The dependent variable of possible destination locations was label encoded. We used the F1-score, a harmonic mean of precision and recall, as a statistical prediction of accuracy where

$$F-\text{measure} = 2 * \frac{\text{precision} * \text{recall}}{\text{precision} + \text{recall}}$$

and where

$$\text{precision} = \frac{\text{true positive}}{\text{true positive} + \text{false positive}}$$

and

$$\text{recall} = \frac{\text{true positive}}{\text{true positive} + \text{false negative}}.$$

The final F1-score is obtained by microaveraging (biased by class frequency), given the nonnormal distribution of output classes. To choose the optimal set of hyperparameters for the random forest, we used scikit-learn's RandomizedSearchCV function. For this, we first created a parameter grid from which to sample during fitting. On each iteration, the algorithm chose a different combination of hyperparameter features. We choose 100 different combinations and set the number of folds to use for cross validation to three.

The feature importance was calculated with Python's scikit-learn's built-in function, which provides the impurity-based feature importances (Gini importances). As a baseline, we selected the most common destination of the passenger and evaluated the overlap of actual predicted destination with setting the predicted destination to the most common destination found with a user's historical data.

Statistical Modeling

We used a stochastic-based Markov chain model [39] to obtain a probabilistic description of various possible next-location visits of a user's trajectory, which is composed of dependent random events (Figure 2). Specifically, a state in the Markov model corresponds to a current location of a station as defined by latitude and longitude coordinates of the transportation vehicle, and state transition corresponds to moving from one station to the next. Thus, the probability of the next likely location of a trajectory can be obtained by

$$\text{Pr}_{\text{next\_location}} = \text{Pr}_{\text{current\_location}} \times \text{transition probability}.$$

Our Markov model is characterized by a state space, a transition matrix describing the probabilities of particular transitions, and an initial state across the state space. Based on a user's past movement, we have a set with several states of

$$S = S_1, S_2, S_3 \ldots, S_r,$$

where each state described one location in a movement grid of the user. The journey starts at one of these processes and as the user travels, the state moves from one state to another. The movement progression from one state to the next thus depends on the probability that a user is moving from that state to the other and can be described by a stochastic probability matrix:

$$P = [P_1, P_2, P_3, \ldots, P_r], \text{ where } \sum_{i=1}^{r} P_i = 1.$$

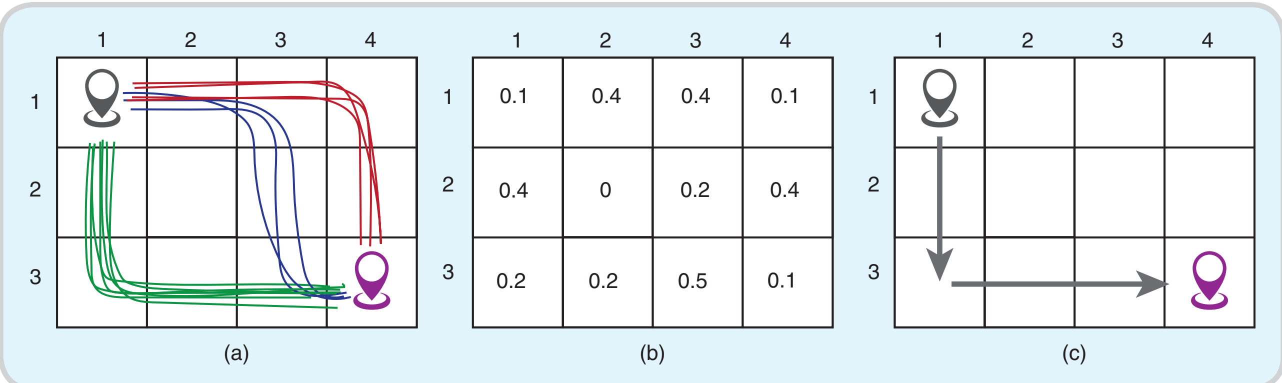

FIG 2 The schematic of a grid world. (a) A 3×4 grid world, where each grid represents the latitude-longitude location of a public transportation station. The Markov chain has 12 possible states. The colored lines indicate the historical mobility of a user traveling from the origin (gray) to the destination (purple). Here *state* refers to the latitude-longitude location of a public transportation station. (b) The transition matrix where each state in the state space is included once as a row and once as a column. Each cell in the matrix displays the probability of transitioning from the row's state to the column's state. (c) An illustrative trip a user is undertaking to derive the probability of starting at [1,1] and arriving at location [3,4] after a trip with two stops: $(1 \to 2 \to 4 = 1 \to 2 \times 2 \to 4 = 0.4 \times 0.4) + (1 \to 3 \to 4 = 1 \to 3 \times 3 \to 4 = 0.4 \times 0.1) = 0.56$. Thus there is a 56% chance that the user is will travel to location [3,4] after two stops, if the user started out in location [1,1].

The probability of transitioning to any of the locations can be represented by transition probability matrix P, as indicated in the following equation. The probability of traveling from i to j is given by p_{ij} (ith row and jth column) in the transition probability matrix P:

$$P = \begin{bmatrix} p_{11} & p_{12} & p_{13} & \dots & p_{1r} \\ p_{21} & p_{22} & p_{23} & \dots & p_{2r} \\ & & & & \\ p_{r1} & p_{r2} & p_{r3} & \dots & p_{rr} \end{bmatrix}.$$

Within the transition probability matrix, the probability of moving to the next state depends on only the current location and not where the user has been before and can be described as

$$\begin{aligned} \Pr(X_{n+1} = x \mid X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) \\ = \Pr(X_{n+1} = x \mid X_n = x_n), \end{aligned}$$

where the probability of X_{n+1} depends only on the probability of X_n, which precedes it.

The size of the Markov chain depends on each user's unique visited locations, thus each transition probability matrix per user has N possible states, where (I, J) of that matrix is the probability of transitioning from state I to state J, i.e., transitioning from location I to J. The product of subsequent matrices describes the probability of the likely destination.

We applied our Markov model on two scenarios for optimal comparison. First, we used the location of where the journey began as initial state and used the initial state/location and transition probability matrix to derive the probability of arriving at the most visited location. Here the transition steps (visits of different locations) are defined by the average number of station visits of a journey, given the starting state/location. Second, we used as the initial state the location of where the journey began, and used the starting state/location and transition probability matrix of arriving at the predicted destinations previously revealed by supervised learning. The steps (visits of different locations) are defined by how many steps were taken/stations were visited to reach that destination. Following the law of large numbers, we estimate the probability of reaching a defined destination by iterating (10,000 times) through the transition probability matrix. Finally, statistical testing was done using Python's library used for scientific and technical computing (SciPy). The values are displayed as mean ± standard error of the mean.

Results

Regular Patterns Underlie Human Mobility at Individual Levels

Despite large differences in individuals' mobility behavior, research has shown that individual trajectories in urban environments are predictable with high accuracy [14]. Human movement is goal directed and often restricted to a select number of origins and destinations. This makes it possible to determine humans' forthcoming travel destinations. To derive a prediction method of users' destinations, we first set out to gain a better understanding of the users' mobility behavior, both at the level of the individual as well as at the group level. In particular, we wondered how human mobility in public transportation is shaped by spatial-temporal features such as location or day of the week. For this, we analyzed the location (as defined by the center of the BSU, i.e., the subdivision of the municipality a user is located at) and time stamp distributions of users' journeys. We found that the origin and destination of a total 26,088 public transportation users were dispersed across the metropolitan area within the time period of 175 days [see Figure 3(a)]. We can observe at least two clusters of condensed journey origins and destinations. Next, we randomly selected four of the

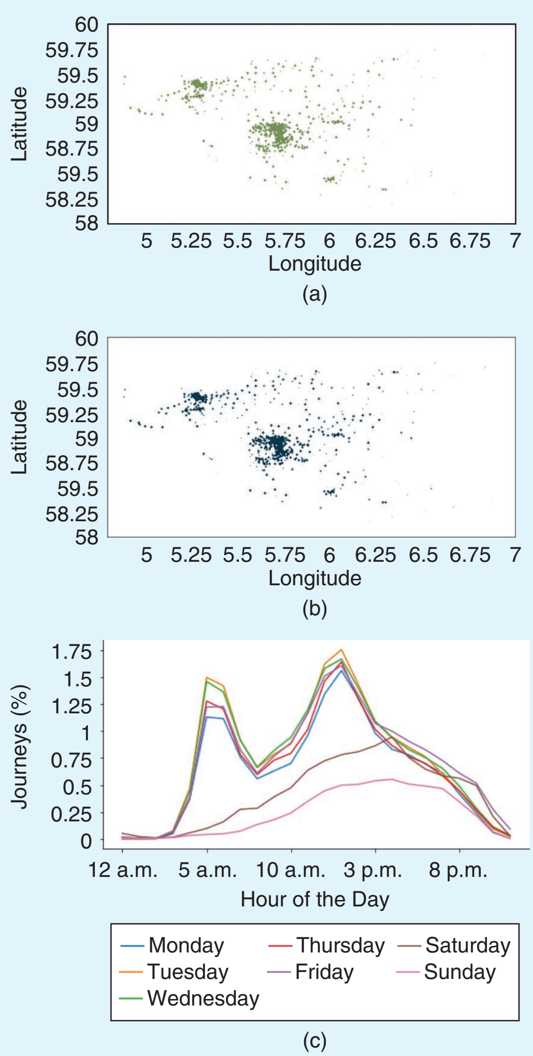

FIG 3 Spatial-temporal mobility at the individual and population levels. (a) A scatterplot showing the spatial distribution of journey origins of journeys undertaken by all users. (b) The same as (a) but for journey destinations. (c) A line plot showing the distribution of journeys across users displayed by hour of the day and day of the week.

26,088 users and observed that travel patterns are restricted to a few selected locations at the individual level [see Figure S1(a) in "User's Journey Counts Are Dependent on Day of the Week and Month of the Year"], thus highlighting the individual's relevant location to which they restrict their travel. In addition to the spatial level, at the temporal level we observed that human mobility follows highly regular patterns both at shorter time scales (hours and week days) [see Figure 3(b)] as well as at longer temporal scales (weeks and months) [see Figure S1(b) and (c)]. Notably, the spatial mobility of individuals is influenced by the temporal context, i.e., a person's travel behavior is different depending on the day of the week (see Figure S2 in "Users Display Different Origin-Destination Patterns"). Thus, individual users display distinct travel patterns at a spatial-temporal scale.

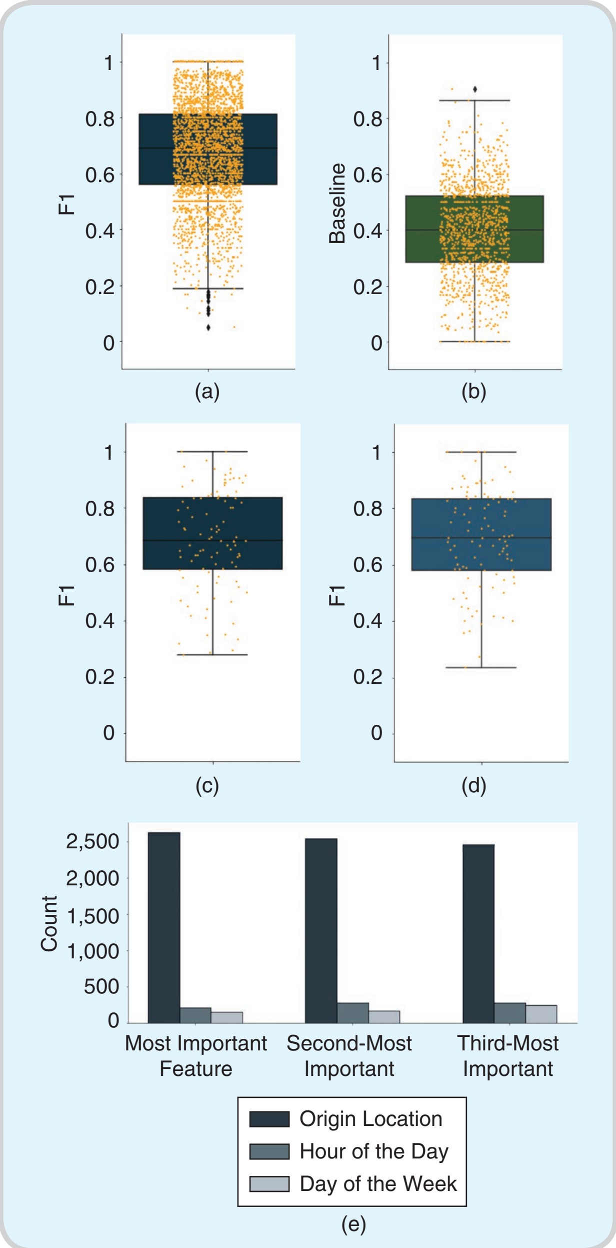

FIG 4 Using spatial-temporal information for next-place prediction. (a) A boxplot showing the F1-score for each customer. The orange dots represent individual users. (b) The same as (a) but for baseline prediction accuracy. (c) The same as (a) but for only 200 randomly selected customers. (d) The same as (a) but for a hyperparameter-tuned model on 100 randomly selected customers. (e) A histogram displaying the distribution of the top-three most important features for all customers.

Human Mobility Patterns for Destination Prediction

The use of spatial-temporal information as outlined in the previous section has been shown to be suitable for predictive analyses of human mobility patterns [24]. Therefore, as the next step, we set out to investigate to what extent we can infer the destination of public transportation users based on gathered spatial-temporal information. As can be seen from Figure S3(a) in "The Quantity of Journeys a User Is Undertaking Is not Related to Prediction Accuracy," public transportation users selectively visit certain locations as the majority of users have a maximum of 20 unique destinations they visited within the last 175 days (mean = 11 unique destination visits). To determine which of these destinations a user is going to visit, we used available spatial information of the journey's origin (i.e., the BSU in which the departure station is located) as well as temporal information (hour of the day and day of the week). Using a random forest classifier for destination prediction, we obtained an F1-score across all customers ($n = 3{,}002$) of 0.68 ($\pm$ 0.003) [see Figure 4(a)]. The F-score was significantly higher when compared to baseline (0.4 $\pm$ 0.002) ($p < 0.05$, $t = 8{,}784.30$) [see Figure 4(b)]. Notably, the accuracy of our ensemble method did not correlate with the overall number of journeys taken by the users [see Figure S3(b)]. Next, we tested an optimized random forest classifier on 100 randomly selected customers [see Figure 4(c)] to maximize prediction accuracy. We found that our optimized random forest classifier reached an average F1-measure of 0.69 ($\pm$ 0.002) [see Figure 4(d)]. Finally, we evaluated the importance of the features, i.e., location of the BSU's origin, hour of the day, and day of the week. We found that for most cases, the feature of origin location scored highly. It was the highest-ranked feature in $n = 2{,}662$ users [see Figure 4(e)]. Hour of the day was the second-highest ranked feature, while day of the week was not significant. In summary, using an ensemble learning method, the final destination of a passenger's journey can be predicted based on basic spatial-temporal information with fairly good accuracy.

Contextual Features From a City's Station Network and People's Past Behavior

There is a consensus that incorporating more contextual data allows for better learning, however, the use of additional contextual features for destination prediction is not predominant [19]. We hypothesized that the prediction of human mobility behavior is improved when including additional relevant contextual beyond the location of origin, hour of the day, and day of the week. Thus, as a next step, we set out to derive more contextual data based on spatial-temporal information available in the data. Broadly, context can be divided into that which is specific to the individual user (e.g., the mode of transportation a user is using) as well as that which is defined by the city's population where an individual contributes to the overall contextual feature (e.g., hotspots of human movement). With respect to the latter aspect, we can obtain information on which transportation mode the user is using at the onset of the trip (based on matching a user's ID and received signal from an installed beacon with the vehicle's ID), or by taking past information into consideration, which transportation mode the user is generally using [see Figure S4(a) in "Additional Contextual Information Suitable as Input for Destination Prediction Models"]. Furthermore, we can get information on the user's average traveled distance to receive a better approximation of a destination [see Figure S4(b)]. In particular, we can analyze whether a passenger is commonly traveling within the metropolitan area (i.e., not traveling farther than the diameter of the metropolitan area) or commutes beyond, which is often associated with a certain type of transportation vehicle (e.g., short- versus long-distance buses).

Additionally, based on additional departure and arrival information, the likely home and workplace of the passenger can be inferred (see Table 2). In addition to contextual features related to the individual user, we can derive contextual features that are defined by the collective behavior of people in an urban environment [40]. For instance, whether an origin or destination station is a hotspot can be determined, which may assist in determining the probability of a user arriving at a certain destination. We observe in our data set that within the city, several clusters of common origins [see Figure S4(c)] as well as destinations [see Figure S4(d)] exist. The origin and destination clusters largely overlap. Thus, the majority of people tend to travel the city by starting and ending a journey in a hotspot. In contrast, at the level of individual stations, we found that some stations have more inward than outward passenger flow or vice versa, i.e., the efflux and influx of passengers is not in balance [see Figure S4(b)]. Thus, several contextual features exist at individual and group levels. These features can be utilized to create prediction models that conform to the urban environment at hand.

Additional Contextual Information for Next-Place Prediction

As a next step, we aimed to supplement spatial-temporal features (see Figure S4) with additional contextual information (see Figure S4) using our multiclass random forest classifier. When incorporating contextual features, the average F1-score of all customers is 0.81 ± 0.002 [Figure 5(a)], reaching significance when compared to the baseline [$p < 0.05$, Figure 5(b)]. We found that the most important features for destination prediction based on all contextual information available is the transportation mode used at the onset of the journey ($n = 708$) and the hour of departure ($n = 299$) [see Figure 5(c)]. At lower rankings, the origin location feature gains in importance, while transportation mode and hour of departure lose significance. The traveled distance of past journeys remains relatively constant [see Figure 5(c)]. As the next step, we performed a random search to optimize our predictions on 100 randomly selected customers. The F1-measure in the absence of hyperparameter tuning reached, on average, 0.82 ± 0.013 [see Figure S3(c)], while parameter optimization led to a similar average F1-score of 0.81 ± 0.013 [see Figure S3(d)]. In summary, incorporating additional contextual information boosts destination prediction accuracy.

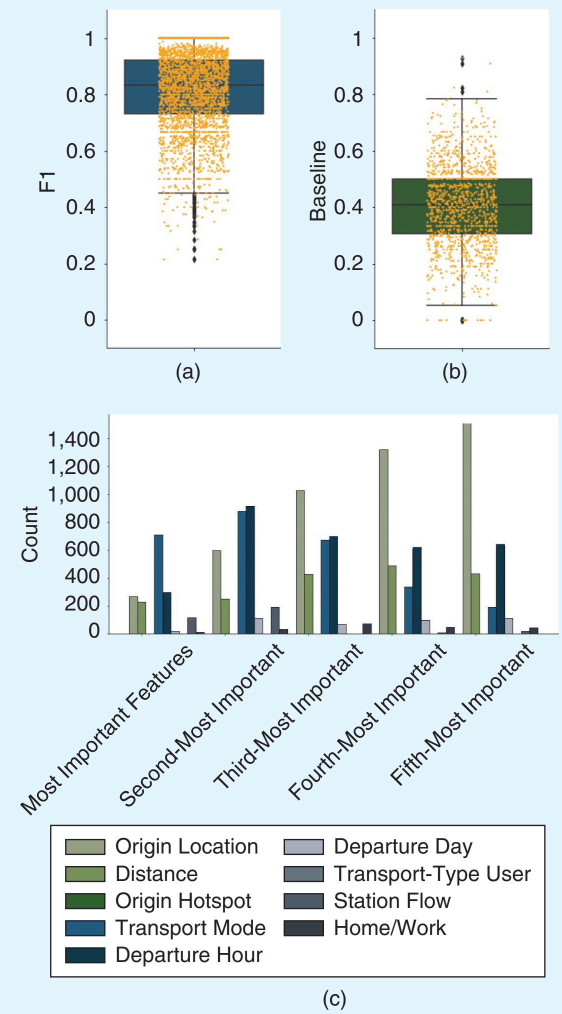

FIG 5 An implementation of contextual information for destination prediction. (a) A boxplot showing next-place prediction accuracy with a random forest classifier for individual journeys. (b) The same as (a) for baseline prediction accuracy. (c) A histogram displaying the distribution of the top-10 features for all customers.

Route Prediction Using Destination Contextual Information

Destination and trajectory predictions are often being investigated and applied independently. Therefore, we examined to what extent contextual predictions of destination prediction can enhance trajectory predictions. A trajectory prediction often makes use of stochastic-based models such as Markov chains, which help analyze dependent random events, i.e., future events depend only on the present event, not the past event. To examine the possibility of merging destination and trajectory prediction models, we used a Markov chain model on station location data obtained during the morning rush hours during weekdays. We used GPS data from a public transportation database to circumvent the use of GPS data from users' smartphones.

In the first step, we used the most common destination and its rank location of a user's history of travel trajectories to determine the step size needed to derive the overall probability of a travel path and its destination, as shown in Figure 2(a). Using this information, we calculated the probability of reaching that destination [see Figure 2(b) and (c)], which was low [see Figure 6(a)]. In comparison, the likely destination probabilities of the same users using our random forest classifier trained on contextual information at the beginning of the journey was significantly higher, as depicted in Figure 6(b). However, a trajectory prediction based a station's location data could be enhanced by using the obtained likely destinations from contextual destination prediction as the likely destination input to the Markov chain model due to better knowledge of the step size needed to arrive at the destination, as presented in Figure 6(c). Although the increment in accuracy is low, the predicted destination and its rank of how many public transportation stops are required to reach it enhance our estimate of the probability obtained with our Markov chain model. Overall, the contextual information at the start of the journey is a stronger predictor of likely destination, rather than using location data in an incremental manner to create a likely estimate of travel destination. However, predicted destinations based on contextual information available at the onset of the journey can enhance trajectory prediction

Discussion

Here we provide a method for destination prediction relying on contextual information obtained in an ITS by means of users' mobile devices, Bluetooth radio transmitters installed on public transportation vehicles, and public registries of public transportation vehicles' location. We have shown that contextual information at the onset of a trip is a crucial predictor of destination location. In particular, the spatial-temporal context at a trip's onset as well as the type of transportation mode used are the most important features for destination prediction. This acquired information on predicted destination at travel onset converted into route prediction improves trajectory prediction based on the public transportation vehicle location data obtained from a public registry.

The focus of our proposed approach is on the individual user rather than the transportation vehicle, for which real-time information of route progression and stop locations are available. Ultimately, in our proposed approach, once the customer boards the transportation vehicle, the final destination of the user is predicted by means of the present contextual information using a tree-based model, whose output then acts as input for a Markov chain model to further refine the prediction during travel progression. In the future, this approach may be complemented with micromobility data to achieve precise route predictions between the start and end location of a user, rather than the start and end location of a user's public transportation trip. Although there exists information that one can easily infer from a trip (origin location, day of the week, or time of the day) this provides only a static approach and limits the accuracy of destination prediction. Ideally, destination prediction should be continuously refined as the travel progresses and more information is acquired, such as knowledge of transportation mode as revealed by algorithms applied on gathered smartphone sensor data. This additional information would then boost the destination accuracy. In other words, significant points, i.e., frequently visited locations during a round trip can be used to infer the destination, as in summation, these points represent a trajectory.

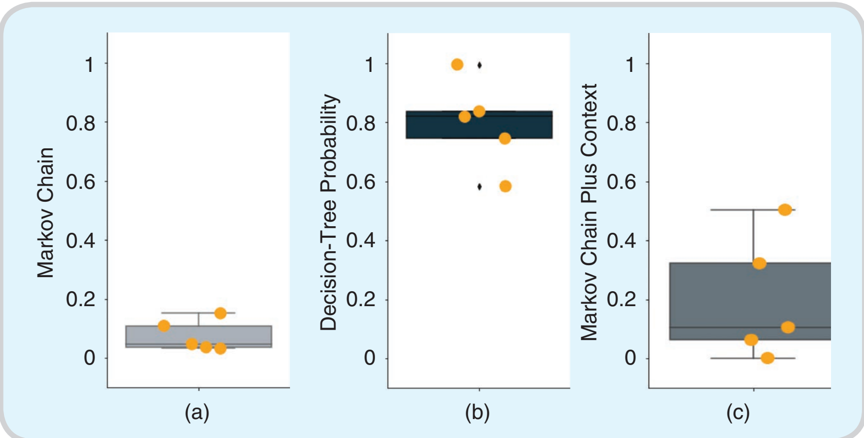

FIG 6 A Markov chain analysis for trajectory prediction. (a) A boxplot showing the product of next-location probabilities to obtain a likely travel destination estimate. (b) A boxplot showing the probability of travel destination obtained by a multiclass random forest classifier. (c) The same as (a) but incorporating the likely destination and travel-path-length form (b) into the Markov chain analysis.

User's Journey Counts Are Dependent on Day of the Week and Month of the Year.

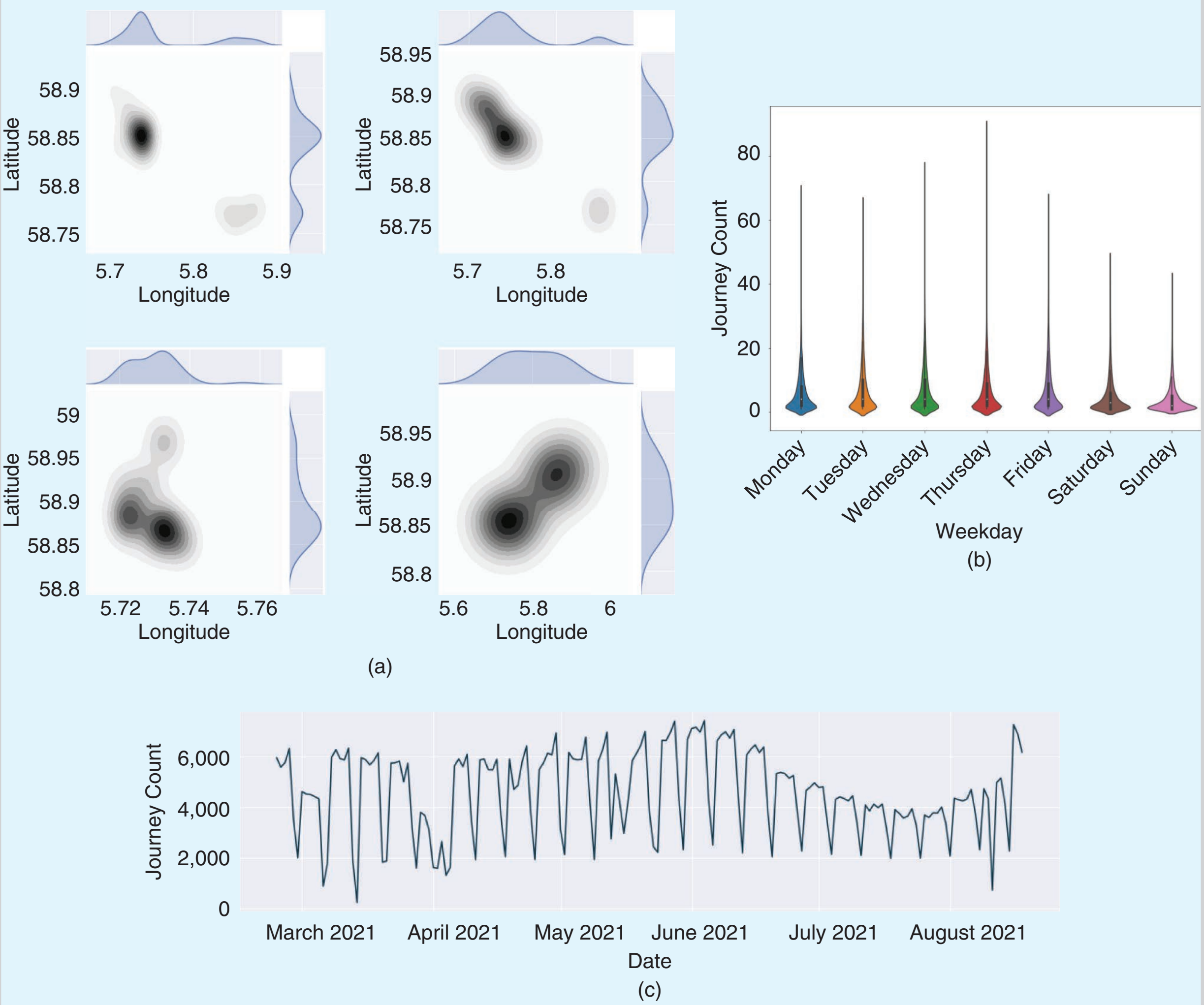

FIG S1 (a) Joint plots showing the density of origins/destinations of journeys for four individual users. (b) Violin plots showing the distribution of the sum of journeys each user has undertaken per day of the week. (c) A line plot showing the total journey count per day across 175 days.

However, this approach is hampered in the case of route changes due to congestion, route work, and so on. Thus, a high accuracy of destination prediction is required at early stages of travel, which we would have shown can be overcome with our suggested approach.

The challenges related to destination prediction often revolve around three issues. First, data sparsity, i.e., the available historical trajectories are far from being able to cover all possible trajectories needed for highly accurate destination prediction. Related to this, although the mere use of historical trajectories allows for the extraction of patterns and activities, they lack the contextual semantics required for understanding the intended activities of the user. The second issue deals with the collection and combination of various data sources related to human behavior, which is increasingly being discussed in the context of user privacy. The third issue concerns the longevity of human mobility patterns and the applicability of destination prediction on gathered mobility data, which may be subject to change over time, e.g., when a person resides in a new location.

Using location data as the main source of a user's context is often related to matters such as privacy concerns, Users Display Different Origin-Destination Patterns.

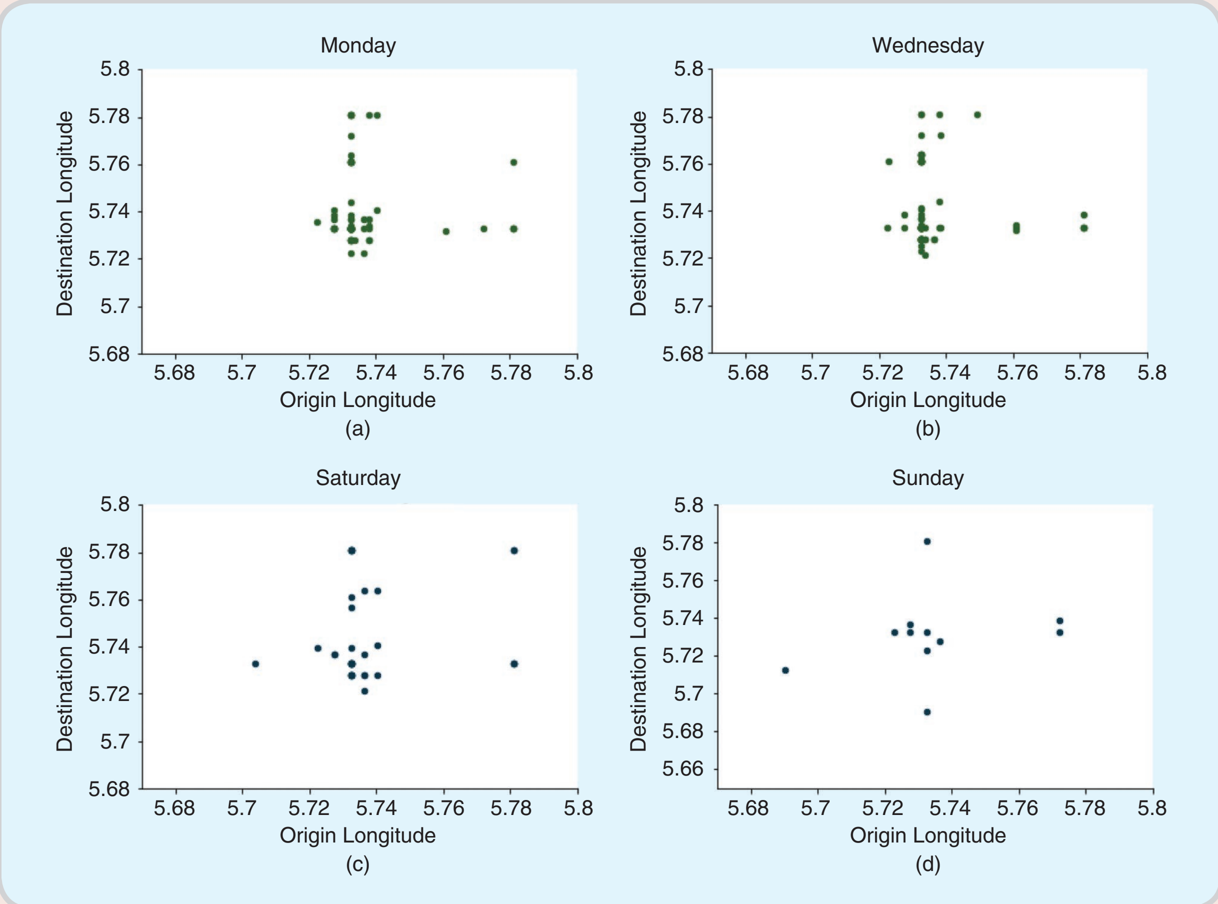

FIG S2 (a) and (b) Scatterplots showing the longitude of origins and destinations of each journey of one user during the week. (c) and (d) Scatterplots showing the longitude of origins and destinations of each journey of one user during the weekend.

budget constraints, or limited availability [35], [41], [42]. As a consequence, available trajectories are limited not sufficient to cover all possible query trajectories, which in turn negatively impacts destination predictability. Several methods have been put forward to overcome this challenge. For example, Wang et al. [43] examine the changes in distances from sampling locations to a final destination on a trajectory instead of searching similar trajectories in a sparse data set. The underlying idea is that the shorter the distance, the closer the selected location is to the destination, and thus the higher the probability is of reaching the destination. Alternatively, Xue et al. [44] propose a method called *SubSyn*, which allows for an expansion of a trajectory data set by decomposition of a trajectory into the subtrajectories of two adjacent locations, which are then merged again to create several synthesized trajectories.

Enhancing User Privacy by Focusing on Sparse Location Data of Public Transportation Vehicles

In our method, we made use of coarse location information at the onset of a trip, as represented by an ID of a subdivision of a municipality at the onset of a trip. We examined whether this information, in combination with additional contextual features (such as time of the day, the used transportation type, whether the origin is a city's hotspot, or whether the user is departing from home or work), is sufficient for destination prediction. We investigated the applicability of contextual destination prediction based solely on the information available at a trip's onset by comparing it to trajectory prediction based on a spatial-temporal sequence of GPS data of the public transportation vehicles during a trip's progression. However, we found that with our Markov chain analysis, destination prediction based on location information during trip progression is not superior to destination prediction based on

The Quantity of Journeys a User Is Undertaking Is Not Related to Prediction Accuracy.

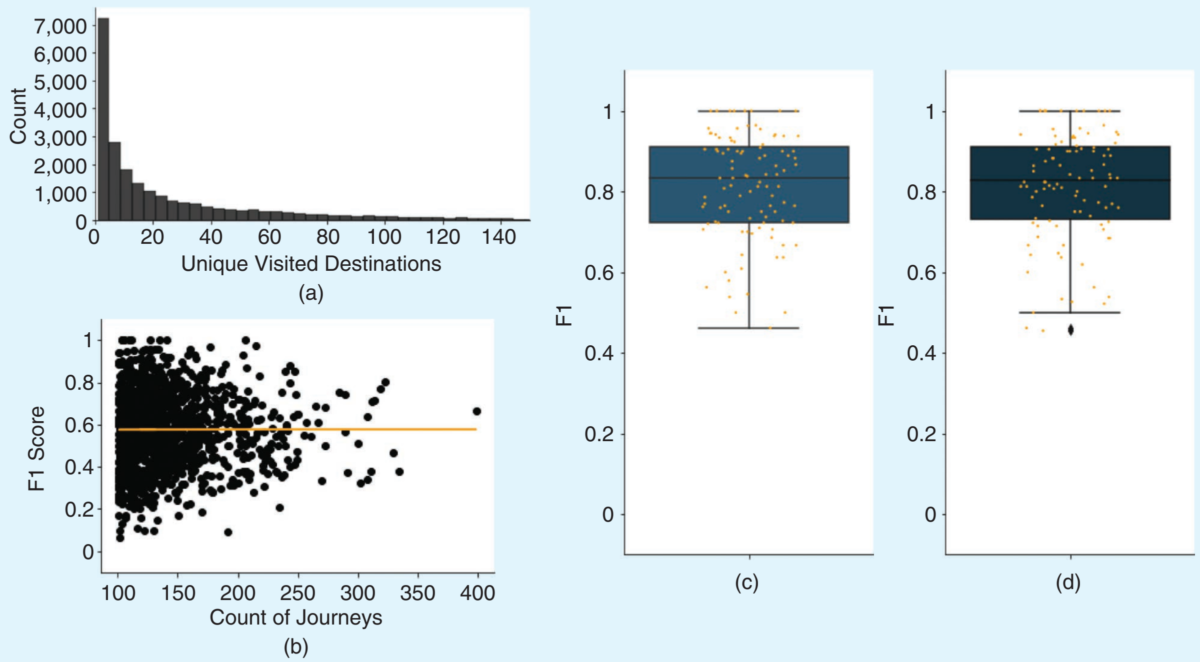

FIG S3 (a) A histogram displaying the count of users and their unique destinations visited. (b) A scatterplot showing the relationship of the obtained F1-measure and journeys taken during the 175-day period. The orange line represents a polynomial fit. (c) A boxplot showing a next-place prediction accuracy random forest classifier for 200 randomly selected users. (d) The same as (d) but for a hyperparameter-tuned classifier.

contextual features at the onset of a trip. As a matter of fact, the performance of our Markov chains was worse but can be marginally improved by restricting possible destination and trip lengths to the destinations and trip lengths predicted by our random forest classifier. Although the Markov chain trajectory prediction was not stellar, we provide a novel solution of destination and trajectory prediction. In particular, we have solely used the coordinates of the public transportation stations where the user is at a given timepoint as inputs for trajectory prediction. This approach may entail complications related to data sparsity given that the sequence of a trip is based only on the location of public transportation stations, however, this can be overcome with a higher sampling of location points. Ultimately, we opted for this novel approach to show the possibilities that lie within trajectory and destination prediction by the use of publicly available location information of a public transportation vehicle a user is in.

Replacing Movement Trajectories With Contextual Information at Trip Onset for Destination Prediction

To minimize the use and collection of personal information over longer time intervals, we have constrained our destination prediction based on contextual information that was solely available at the onset of a public transportation ride. In particular, we have used both spatial-temporal information of the user (place, time of the day, and day of the week) as well as contextual information, such as transportation mode (bus, ferry, and train), dynamic properties of the city, and the BSU that might represent the home/work location. Little research exists that relies on (temporal) contextual information beyond spatial information, and even less work has incorporated the mode of transportation into destination prediction [19]. We obtained this information by matching the assigned passenger on a transportation vehicle to the location data, thereby receiving location information only when a user is matched to a certain transportation vehicle. In fact, several approaches exist to infer a user's mode of transportation based on mobile phone sensors with high accuracy (>90%) [45], [46]. However, these approaches often require that users opt-in for the collection of sensor data of their mobile devices.

In line with the fact that people restrict their movements to selected pathways and stations [14], [31], and that they possibly use the same transportation type to

Additional Contextual Information Suitable as Input for Destination Prediction Models.

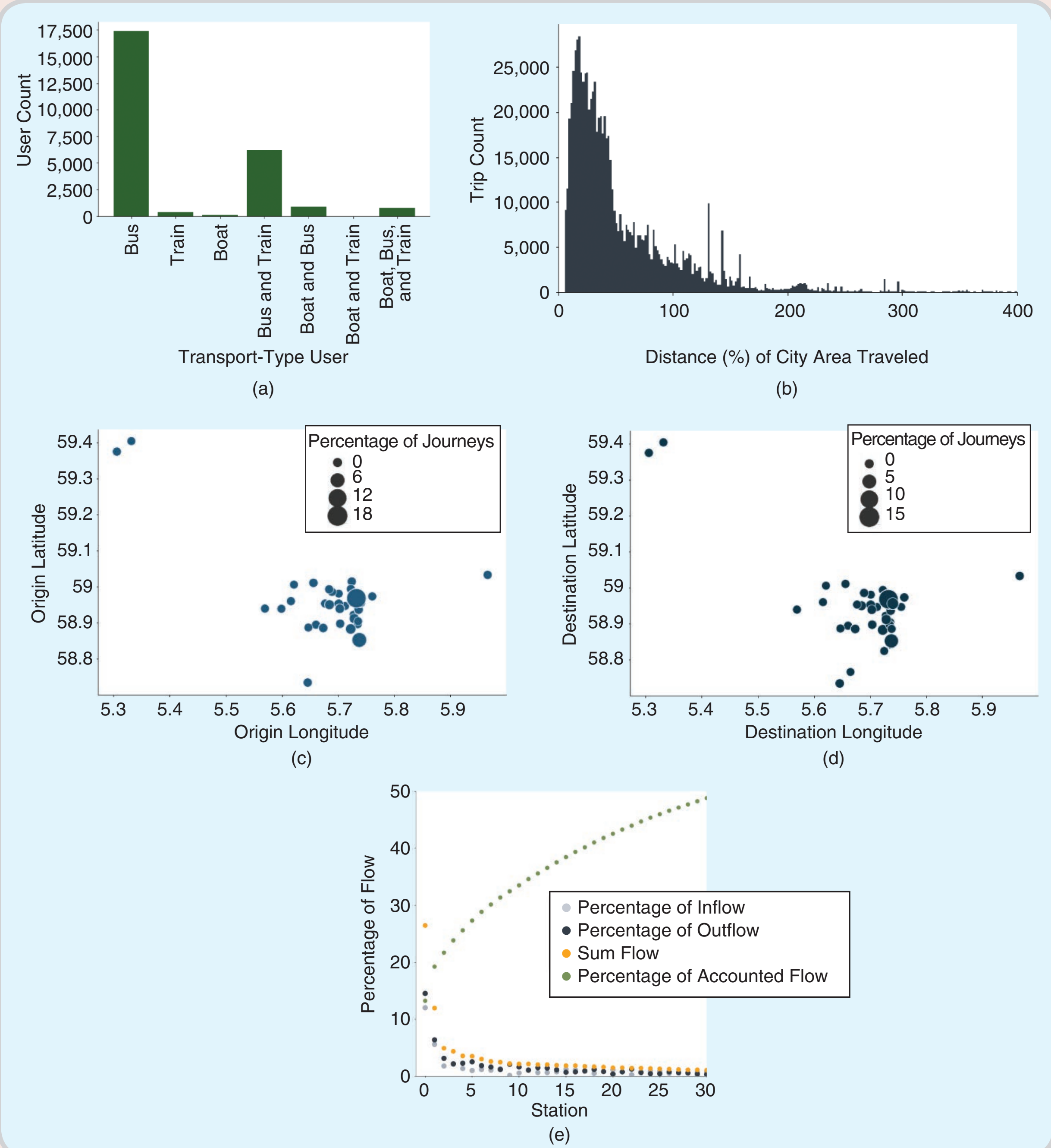

FIG S4 (a) A bar graph showing the count of traveler types. (b) A bar graph showing the linear distance of each trip as measured by diameter of traversed distance of the urban area (100% = 8.85 km). (c) A scatterplot displaying the location of origin hotspots in the city. The size of the circle represents the percentage of all journeys. (d) The same as (c) but for destination hotspots. (e) A dotted-line plot showing the inward (gray) and outward (blue) movement of people to stations as well as the sum of inward and outward movements (orange) and the accounted overall flow (light green).

cover selected distances, we show that transportation mode is a particularly strong predictor of a user's destination. A further measure we selected to decrease the use of personal information while at the same time boosting the accuracy of destination prediction by our random forest classifier is the use of features of a city. These features are collectively being defined by the movements of its inhabitants. For example, the changes in mobility of users during the day has an impact on the overall urban environment [47]. It has been shown that the spatial structure of cities is dynamic, displaying both stability as well as variations over time during the day [40]. For example, the persistence of hotspots, i.e., the physical location of a public transportation station showing an accumulation of departures and/or arrivals, can be permanent or intermittent depending on location and time of the day. By these means, we merge individual and collective movement patterns for destination prediction, as has previously been performed [48]. However, although spatial-temporal information entails some context and possible causes of humans' intentions to move, it is still limited. One way of enhancing our knowledge of users' intentions is to incorporate semantic information [38].

Longevity of Human Mobility and the Effect of Temporal Change

Human movement appears to be random from an outsider's perspective, however, human movement follows strict, regular patterns. As mentioned in the previous section, one of the prime investigations into human mobility showed that people restrict their movements among a few selected locations at high temporal regularity, e.g., the daily commute to and from the workplace [31]. In stark contrast to work published prior to the outbreak of the COVID-19 global pandemic in 2020–2021, the findings described here are based on mobility data collected in a period dominated by mobility restrictions. Nevertheless, even though the mobility of people and their use of public transportation has been reduced, we show that accurate destination prediction is possible under these circumstances. Depending on national guidelines, on the one hand, a reduced set of public transportation data might consist predominantly of people commuting to and from work. This would increase the regularity of travel patterns. On the other hand, people worked mainly remotely when possible and used public transportation for other activities. Although it is challenging to derive trip intentions based solely on the available information, we assume that the trip intentions described here mainly comprise trips to and from work, where it was not possible to work from home, or trips undertaken for social/leisure activities. However, once restrictions are removed, people will most likely increase their public transportation usage, which in turn may transform the collected trajectory patterns on which our predictions may not hold any longer.

The interplay of spatial-temporal context has been investigated by Chon et al. [49]. The authors found that spatial and temporal context are tightly connected, i.e., a location-dependent predictor is better than a location-independent predictor for predicting the temporal behavior of individual users. A further challenge of route and destination prediction based on contextual data might be that users have different intentions, leading to various movement patterns at different spatial and temporal scales [48]. Destination prediction models must be concerned with these. For instance, commuters have a more stable trajectory history over time, while tourists are more exploratory in their mobility behavior, and the movement patterns can change often and quickly in short periods of time.

Conclusions

Overall, we have shown the importance of contextual information at early stages of travel for the accurate prediction of a journey's destination. In particular, we considered contextual information related to individual users as well as information that is defined by the collective behavior of a city's inhabitants. We found that transportation mode is a critical determinant of future whereabouts of users. Furthermore, we have demonstrated that obtained predictions from contextual information enhances trajectory predictions. Future work must address how destination and trajectory prediction may be merged for LBSs. Furthermore, additional investigations are required into how the decay and dynamic change of mobility patterns over time can be addressed in destination predictions. Overall, this is particularly important to maintain the utility of human movement patterns for LBSs.

Acknowledgments

The work was funded by the Norwegian Research Council under grant 309716 and by fluxLoop to Tor-Morten Grønli.

About the Authors

Malte Bieler (malte.bieler@kristiania.no) earned his Ph.D. degree in neuroscience from the University of Hamburg in 2017. He is currently a postdoctoral researcher in the Research Group for Applied Computer Science and the Mobile Technology Lab at the Institute of Technology, Kristiania University College, Oslo, 0152, Norway. His primary research interests lie within urban informatics, particularly in the analysis of mobility data for the extraction of various types of patterns or models.

Raghava Rao Mukkamala (rrm.digi@cbs.dk) earned his Ph.D. degree in computer science from the IT University of Copenhagen, Denmark. He is currently an associate professor with the Centre for Business Data Analytics, Department of Digitalization, Copenhagen Business School (CBS), Frederiksberg, 2000, Denmark. He is director for the Masters in Data Science program at CBS

and teaches several courses on machine learning. He is also with the Department of Technology, Kristiania University College, Oslo, 0152, Norway. His current research interests focus on big data analytics, data science, blockchain technologies, and cybersecurity. Combining formal/mathematical modelling approaches with machine learning techniques, his current research program seeks to develop new algorithms for big data analytics.

Tor-Morten Grønli (tor-morten.gronli@kristiania.no) earned his Ph.D. degree in computer science from Brunel University. He is currently a professor in the Department of Technology, Kristiania University College, Oslo, 0152, Norway, and the founding director of the Research Group for Applied Computer Science and the Mobile Technology Lab. He is also director for the Ph.D. in Information Technology program at Kristiania University College and teaches courses on software architecture, emerging technologies, and mobile/Internet of Things (IoT) computing. His primary research focus on mobile computing, the IoT, applied machine learning, and software architectures.

References

[1] R. Abdalla, *Mobile GIS and Location-Based Services (LBS)*. Cham, Switzerland: Springer, 2016, pp. 83–103.
[2] J. Nord, K. Synnes, and P. Parnes, “An architecture for location aware applications,” in *Proc. 35th Annu. Hawaii Int. Conf. Syst. Sci.*, 2002, pp. 3805–3810, doi: 10.1109/HICSS.2002.994513.
[3] F. Wu and Z. Li, “Where did you go: Personalized annotation of mobility records,” in *Proc. 25th ACM Int. Conf. Inf. Knowl. Manage.*, 2016, pp. 589–598, doi: 10.1145/2983323.2983845.
[4] S. Eubank *et al.*, “Modelling disease outbreaks in realistic urban social networks,” *Nature*, vol. 429, no. 6988, pp. 180–184, 2004, doi: 10.1038/nature02541.
[5] H. A. Makse, S. Havlin, and H. E. Stanley, “Modelling urban growth patterns,” *Nature*, vol. 377, no. 6550, pp. 608–612, 1995, doi: 10.1038/377608a0.
[6] H. D. Rozenfeld, D. Rybski, J. S. Andrade, M. Batty, H. E. Stanley, and H. A. Makse, “Laws of population growth,” *Proc. Nat. Acad. Sci. USA*, vol. 105, no. 48, pp. 18,702–18,707, 2008, doi: 10.1073/pnas.0807435105.
[7] F. Gu *et al.*, “Indoor localization improved by spatial context—A survey,” *ACM Comput. Surv.*, vol. 52, no. 3, pp. 1–35, 2019, doi: 10.1145/3322241.
[8] T. H. Silva *et al.*, “Urban computing leveraging location-based social network data: A survey,” *ACM Comput. Surv.*, vol. 52, no. 1, pp. 1–39, 2019, doi: 10.1145/3301284.
[9] S. Winter, M. Tomko, M. Vasardani, K.-F. Richter, K. Khoshelham, and M. Kalantari, “Infrastructure-independent indoor localization navigation, *ACM Comput. Surv.*, vol. 52, no. 3, pp. 1–24, 2019, doi: 10.1145/3321516.
[10] Z. Ding, X. Li, C. Jiang, and M. Zhou, “Objectives and state-of-the-art of location-based social network recommender systems,” *ACM Comput. Surv.*, vol. 51, no. 1, pp. 1–28, 2018, doi: 10.1145/3154526.
[11] A. Rudenko, L. Palmieri, M. Herman, K. M. Kitani, D. M. Gavrila, and K. O. Arras, “Human motion trajectory prediction: A survey,” *Int. J. Robot. Res.*, vol. 39, no. 8, pp. 895–935, 2020, doi: 10.1177/0278364920917446.
[12] M. Alzantot and M. Youssef, “UPTIME: Ubiquitous pedestrian tracking using mobile phones,” in *Proc. 2012 IEEE Wireless Commun. Netw. Conf. (WCNC)*, pp. 3204–3209, doi: 10.1109/WCNC.2012.6214359.
[13] M. B. Kjærgaard, S. Bhattacharya, H. Blunck, and P. Nurmi, “Energy-efficient trajectory tracking for mobile devices,” in *Proc. 9th Int. Conf. Mobile Syst., Appl., Services (MobiSys ‘11)*, 2011, pp. 307–320, doi: 10.1145/1999995.2000025.
[14] C. Song, Z. Qu, N. Blumm, and A.-L. Barabasi, “Limits of predictability in human mobility,” *Science*, vol. 327, no. 5968, pp. 1018–1021, 2010, doi: 10.1126/science.1177170.
[15] C. Song, T. Koren, P. Wang, and A.-L. Barabási, “Modelling the scaling properties of human mobility,” *Nature Phys.*, vol. 6, no. 10, pp. 818–823, 2010, doi: 10.1038/nphys1760.
[16] M. Chen, X. Yu, and Y. Liu, “Mining moving patterns for predicting next location,” *Inf. Syst.*, vol. 54, pp. 156–168, Dec. 2015, doi: 10.1016/j.is.2015.07.001.
[17] N. Mamoulis, H. Cao, G. Kollios, M. Hadjieleftheriou, Y. Tao, and D. W. Cheung, “Mining, indexing, and querying historical spatiotemporal data,” in *Proc. 2004 ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining (KDD ‘04)*, pp. 236–245, doi: 10.1145/1014052.1014080.
[18] J. Yuan *et al.*, “T-drive: Driving directions based on taxi trajectories,” in *Proc. 18th SIGSPATIAL Int. Conf. Adv. Geographic Inf. Syst. (GIS ‘10)*, 2010, pp. 99–100, doi: 10.1145/1869790.1869807.
[19] C. Schreckenberger, S. Beckmann, and C. Bartelt, “Next place prediction: A systematic literature review,” in *Proc. 2018 2nd ACM SIGSPATIAL Workshop Prediction Human Mobility (Predict GIS)*, pp. 37–45, doi: 10.1145/3283590.3283596.
[20] F. Araújo, F. Araújo, K. Machado, D. Rosário, E. Cerqueira, and L. A. Villas, “Ensemble mobility predictor based on random forest and Markovian property using LBSN data,” *J. Internet Services Appl.*, vol. 11, no. 1, p. 7, 2020, doi: 10.1186/s13174-020-00130-7.
[21] J.-P. Rodrigue, C. Comtois, and B. Slack, “The geography of transport systems,” in *The Geography Transport Systems*. London, U.K.: Routledge, 2017.
[22] Z. A. Almusaylim and N. Jhanjhi, “Comprehensive review: Privacy protection of user in location-aware services of mobile cloud computing,” *Wireless Pers. Commun.*, vol. 111, no. 1, pp. 541–564, 2020, doi: 10.1007/s11277-019-06872-3.
[23] M. Siddula, Y. Li, X. Cheng, Z. Tian, and Z. Cai, “Privacy-enhancing preferential LBS query for mobile social network users,” *Wireless Commun. Mobile Comput.*, vol. 2020, pp. 1–13, Sep. 2020, doi: 10.1155/2020/8892321.
[24] R. Imai, K. Tsubouchi, T. Konishi, and M. Shimosaka, “Early destination prediction with Spatio-temporal user behavior patterns,” in *Proc. ACM Interact., Mobile, Wearable Ubiquitous Technol.*, 2017, vol. 1, pp. 1–19, doi: 10.1145/3161197.
[25] A. Y. Xue, R. Zhang, Y. Zheng, X. Xie, J. Huang, and X. Zhenghua, “Destination prediction by sub-trajectory synthesis and privacy protection against such prediction,” in *Proc. 2013 IEEE 29th Int. Conf. Data Eng. (ICDE)*, pp. 254–265, doi: 10.1109/ICDE.2013.6544830.
[26] J. Alvarez-Garcia, J. Ortega, L. Gonzalez-Abril, and F. Velasco, “Trip destination prediction based on past GPS log using a hidden Markov model,” *Expert Syst. Appl.*, vol. 37, no. 12, pp. 8166–8171, 2010, doi: 10.1016/j.eswa.2010.05.070.
[27] C. Manasseh and R. Sengupta, “Predicting driver destination using machine learning techniques,” in *Proc. IEEE 16th Int. Conf. Intell. Transport. Syst. (ITSC 2013)*, 2013, pp. 142–147, doi: 10.1109/ITSC.2013.6728224.
[28] J. Zhao, J. Xu, R. Zhou, P. Zhao, C. Liu, and F. Zhu, “On prediction of user destination by sub-trajectory understanding: A deep learning based approach,” in *Proc. 27th ACM Int. Conf. Inf. Knowl. Manage.*, 2018, pp. 1413–1422, doi: 10.1145/3269206.3271708.
[29] J. Krumm and E. Horvitz, “Predestination: Inferring destinations from partial trajectories,” in *UbiComp 2006: Ubiquitous Computing*, vol. 4206, P. Dourish and A. Friday, Eds. Berlin: Springer-Verlag, pp. 243–260.
[30] D. Brockmann, L. Hufnagel, and T. Geisel, “Scaling laws of human travel,” *Nature*, vol. 439, no. 7075, pp. 462–465, 2006, doi: 10.1038/nature04292.
[31] M. C. González, C. A. Hidalgo, and A.-L. Barabási, “Understanding individual human mobility patterns,” *Nature*, vol. 453, no. 7196, pp. 779–782, 2008, doi: 10.1038/nature06958.
[32] A. Tsiligkaridis, J. Zhang, H. Taguchi, and D. Nikovski, “Personalized destination prediction using transformers in a contextless data setting,” in *Proc. 2020 IEEE Int. Joint Conf. Neural Netw. (IJCNN)*, pp. 1–7, doi: 10.1109/IJCNN48605.2020.9207514.
[33] J. Xu, J. Zhao, R. Zhou, C. Liu, P. Zhao, and L. Zhao, “Predicting destinations by a deep learning based approach,” *IEEE Trans. Knowl. Data Eng.*, vol. 33, no. 2, pp. 651–666, 2021, doi: 10.1109/TKDE.2019.2932984.
[34] Y. Lu and Y. Liu, “Pervasive location acquisition technologies: Opportunities and challenges for geospatial studies,” *Comput., Environ. Urban Syst.*, vol. 36, no. 2, pp. 105–108, 2012, doi: 10.1016/j.compenvurbsys.2012.02.002.
[35] D. Ashbrook and T. Starner, “Using GPS to learn significant locations and predict movement across multiple users,” *Pers. Ubiquitous Comput.*, vol. 7, no. 5, pp. 275–286, 2003, doi: 10.1007/s00779-003-0240-0.
[36] B. D. Ziebart, A. L. Maas, A. K. Dey, and J. A. Bagnell, “Navigate like a cabbie: Probabilistic reasoning from observed context-aware behavior,” in *Proc. 10th Int. Conf. Ubiquitous Comput. (UbiComp ‘08)*, 2008, pp. 322–331, doi: 10.1145/1409635.1409678.
[37] W. Huang, L. Man, H. Weisong, G. Song, and K. Xie, “Hierarchical destination prediction based on GPS history,” in *Proc. 2013 IEEE 10th Int.*

Conf. Fuzzy Syst. Knowl. Discovery (FSKD), pp. 972–977, doi: 10.1109/FSKD.2013.6816336.

[38] L. Xia, Q. Huang, and D. Wu, “Decision tree-based contextual location prediction from mobile device logs,” *Mobile Inf. Syst.*, vol. 2018, p. 1,852,861, Apr. 2018, doi: 10.1155/2018/1852861.

[39] L. E. Baum and T. Petrie, “Statistical inference for probabilistic functions of finite state Markov chains,” *Ann. Math. Statist.*, vol. 37, no. 6, pp. 1554–1563, doi: 10.1214/aoms/1177699147.

[40] T. Louail *et al.*, “From mobile phone data to the spatial structure of cities,” *Scientific Rep.*, vol. 4, no. 1, p. 5276, 2014, doi: 10.1038/srep05276.

[41] N. Marmasse and C. Schmandt, “A user-centered location model,” *Pers. Ubiquitous Comput.*, vol. 6, no. 5, pp. 318–321, doi: 10.1007/s007790200035.

[42] S. Wang, Z. Bao, J. S. Culpepper, and G. Cong, “A survey on trajectory data management, analytics, and learning,” *ACM Comput. Surv.*, vol. 54, no. 2, pp. 1–36, 2020, doi: 10.1145/3440207.

[43] L. Wang, Z. Yu, B. Guo, T. Ku, and F. Yi, “Moving destination prediction using sparse dataset: A mobility gradient descent approach,” *ACM Trans. Knowl. Discovery Data*, vol. 11, no. 3, pp. 1–33, 2017, doi: 10.1145/3051128.

[44] A. Y. Xue, J. Qi, X. Xie, R. Zhang, J. Huang, and Y. Li, “Solving the data sparsity problem in destination prediction,” *VLDB J.*, vol. 24, no. 2, pp. 219–243, 2015, doi: 10.1007/s00778-014-0369-7.

[45] S.-H. Fang, Y.-X. Fei, Z. Xu, and Y. Tsao, “Learning transportation modes from smartphone sensors based on deep neural network,” *IEEE Sensors J.*, vol. 17, no. 18, pp. 6111–6118, 2017, doi: 10.1109/JSEN.2017.2737825.

[46] L. Stenneth, O. Wolfson, P. S. Yu, and B. Xu, “Transportation mode detection using mobile phones and GIS information,” in *Proc. 19th ACM SIGSPATIAL Int. Conf. Adv. Geographic Inf. Syst., GIS ‘11*, ACM Press, 2011, pp. 54–63, doi: 10.1145/2093973.2093982.

[47] C. Roth, S. M. Kang, M. Batty, and M. Barthélemy, “Structure of urban movements: Polycentric activity and entangled hierarchical flows,” *PLoS One*, vol. 6, no. 1, p. e15923, doi: 10.1371/journal.pone.0015923.

[48] N. C. Chen, W. Xie, R. E. Welsch, K. Larson, and J. Xie, “Comprehensive predictions of tourists’ next visit location based on call detail records using machine learning and deep learning methods,” in *Proc. 2017 IEEE Int. Congr. Big Data (BigData Congress)*, pp. 1–6, doi: 10.1109/BigDataCongress.2017.10.

[49] Y. Chon, H. Shin, E. Talipov, and H. Cha, “Evaluating mobility models for temporal prediction with high-granularity mobility data,” in *Proc. 2012 IEEE Int. Conf. Pervasive Comput. Commun.*, pp. 206–212, doi: 10.1109/PerCom.2012.6199868.

ITS