Research Article

A Spatial-Temporal-Semantic Method for Location Prediction in Indoor Spaces

Peng Wang , Jing Yang , and Jianpei Zhang

College of Computer Science and Technology, Harbin Engineering University, Harbin 150000, China

Correspondence should be addressed to Jing Yang; yangjing@hrbeu.edu.cn

Received 6 July 2021; Accepted 16 February 2022; Published 3 March 2022

Academic Editor: Enrico M. Vitucci

Copyright © 2022 Peng Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

While a large number of works concentrated on forecasting trajectories in the outdoor environment, predicting the movement of users in indoor settings has attracted much more attention recently because of the development of smartphones and maturity of Wi-Fi services, e.g., in office buildings. Predicting a user's movement in indoor spaces can not only help better understand his/her intentions but also improve his/her living experience. While most of the prediction approaches to date tackle the problem by constructing the mathematical models to learn the mobility of objects, they cannot efficiently model the movement of indoor users in the constraint but filled with spatial-temporal-semantic info settings. In order to solve this issue, we propose a frequent subtrajectory-based Markov model that incorporates the spatial location, the temporal aspect, and the shop category context into a unified framework. We first present the frequent subtrajectory algorithm to model and predict adjacent moving points from physical movement perspective, and then, by taking the duration of stay at a specific location into account, we further improve the prediction precision. Finally, by taking location context in the indoor environment (e.g., shop categories) into consideration, we successfully model and predict the user's future visiting points from the semantic perspective. To validate the effectiveness of our model, we conduct a complete evaluation on a large-scale real-world dataset with more than 261,269 trajectories collected from over 120,000 customers in a shopping mall. The experiment results demonstrate that our method performs significantly superior prediction performance comparing the state-of-the-art models.

1. Introduction

During the past decades, a large number of studies have concentrated on predicting the further whereabouts of human beings [1–4]. However, with the prevalence of global positioning system (GPS), the majority of works are mainly focused on outdoor trajectory prediction, but researches show that people tend to spend over 87% of their lifetime in indoor settings, for instance, conference rooms, shopping centers, transition terminals, and private homes [5, 6]. What is more, indoor localization products such as Bluetooth, Infrared, RFID, Wi-Fi [7], and iBeacon [8] have been popular in commercial markets [9, 10]; more and more people started to use the Wi-Fi-enabled smartphones; as a direct and promising application based on indoor positioning system [10], indoor trajectory prediction gradually attracts much more attention which is a crucial part for the location-based services for human beings.

Two main reasons lead to the insufficiency of study on indoor trajectory prediction. The first one is that without accurate indoor positioning technology, it is hard to collect usable and sufficient experimental data [11]; because the techniques used in outdoor localization such as GPS or GLONASS (Global Navigation Satellite System) perform badly in the indoor settings, the signal attenuation will aggravate significantly in the indoor settings because of the building materials [12]. The other reason is that the indoor environment is much more complicated compared to that of outdoor settings. People just need to track the road map to their destinations in the outdoor environment, but there will be multiple choices for the customers even in a small region due to the complex elements in indoor settings such as gates and elevators which make it hard to accurately predict the further whereabouts of indoor users. Note that we do not predict a specific shop the customer will visit next, in our study, due to the unique feature of our dataset; what we predict is the next Wi-Fi access point (AP) the customer will connect.

Nowadays, many indoor establishments are providing free Wi-Fi services to improve user experiences in their facilities. For instance, in the British supermarket Tesco, hundreds of its shops have provided Wi-Fi connections. Wi-Fi is almost an indispensable and fundamental service for large indoor buildings. By taking advantage of this service, building managers can efficiently capture the physical movements of workers or customers [13–16] to provide better services for customers or accomplish efficient management of workers. And researches on accurate Wi-Fi-based positioning have been carried out; for example, in [17], the authors tackle the problem of insufficient RSS measurements for Wi-Fi-based localization.

In this article, a large-scale real-world dataset containing over 261,369 indoor trajectories was collected from a large shopping mall with seven floors across 90,000 square meters. The mall operator has deployed 67 Wi-Fi access points in the shopping mall; other information such as floor plans, deployment of the Wi-Fi access points, and the shop categories are provided by the proprietor of the building. The data offers us a special chance to study human movement in indoor space and investigate the correlation between human physical movement and contextual movement in indoor settings. To enhance the quality of trajectory prediction in the indoor scenarios, we present a subtrajectory-based Markov model by incorporating the spatial, temporal, and semantic information into a unified model. The major contributions of this work can be briefed as follows:

- (1) From the spatial point, a frequent subtrajectory-based forecasting approach is proposed which not only improves the prediction precision but also restrains the exponential increase in probability transition matrix and solves the data sparsity problem; what is more, the proposed method only needs the most recent movements to do the prediction which saves much storage space
- (2) From the temporal perspective, we investigate the influence of users' retention time at a specific location and the starting time of a trajectory on the prediction performance
- (3) From the semantic perspective, we investigate the semantic information in the indoor environment, e.g., shop categories in the shopping mall; we verify to which extent we can accurately predict which semantic section the users will go next; we also propose a unified model to integrate semantic section prediction with location prediction; it further improves the prediction performance
- (4) We conduct comprehensive experiments to test the prediction accuracy of our algorithm using a real-world dataset obtained from a shopping mall. Experimental outcomes show that the newly proposed method performs better than the baseline techniques including Markov model and hidden Markov model

Shopper and store managers can benefit a lot from indoor trajectory prediction. For instance, when managers

can predict the customers' movement in advance, they can compare the predicted location with users' history locations; if the managers find abnormal movements of a specific user, they can make contingency plans to avoid dangerous events. Furthermore, retailers can push related promotion information to a target customer through online advertising system if they know that shopper will physically approach their store; it will not only boost sales but also improve the customers' shopping experience.

The rest of this study is structured as follows. First, we present literature review in Section 2. In Section 3, we will introduce the preliminaries on indoor trajectory prediction. In Section 4, we present the details about the proposed prediction model which integrates spatial, temporal, and semantic information into a unified framework. In Section 5, experiments were carried out to verify the effectiveness of our algorithm. Finally, we conclude the paper in Section 6.

2. Related Work

In this part, we highlight related studies on indoor trajectory prediction, and differences between these works and our study are discussed.

2.1. Trajectory Prediction. During the past decade, extensive research has been carried out in the field of trajectory prediction; for example, in the book, the authors provide the latest, cutting-edge research on neural networks and machine learning methods to solve the data prediction problem [18]. Based on the type of experimental data, the methodology of trajectory prediction can be divided into two groups: continuous and discrete. In the group of continuous, the continuous coordinate values are generated through the global positioning system, whereas in the group of discrete, the dataset was generated through wireless sensors where the coordinate values denote the identifications of the sensors [19]. The deviation between the predicted locations and the corresponding real locations can be used to evaluate the accuracy of continuous forecasting, and the performance of discrete prediction can be evaluated as the percentage of the actual IDs in the returned predicted IDs. Since our dataset is collected from the shopping mall and the trajectory is represented as a sequence of time-ordered Wi-Fi access points, our work belongs to the category of discrete trajectory prediction.

Based on the prediction methods, existing work can be classified into (1) frequent pattern mining and (2) model building.

2.2. Frequent Pattern Mining. A large number of studies have been carried out by extracting frequent patterns to predict users' further locations [20]. In the representative work [21], an indexing scheme that exploits periodic pattern information to organize historical spatiotemporal data was proposed. The authors combined the prefix tree and FP tree to predict the movement rule of objects [22].

While most of the studies are carried out on the basis of geographical characteristics, temporal and contextual information have also been investigated recently [23]. In [24],

an Indoor-WhereNext approach for trajectory prediction in the shopping mall is proposed; the authors first compute the similarities of location sequences based on the aspects of spatial and semantic, and then, similarity user groups are obtained by employing the AP algorithm which is used to train the prediction models. In [25], a spatial-temporal model to forecast the trajectories in the indoor setting for the blinds has been studied to help them to avoid obstacles. While the aforementioned works adopt the trajectory patterns for prediction directly, they did not consider the situation of reoccurrences of a pattern behind human movements and they did not provide a solution for continuous updating frequent patterns.

2.3. Model Construction. When considering the approaches of model construction, researchers put forward machine learning algorithms such as dynamic Bayesian network [26, 27] and recurrent neural network (RNN) to forecast human further whereabouts by constructing predicting models based on the historical trajectory data.

In Jin et al.'s study [28], an augmented-intention recurrent neural network (AI-RNN) model to forecast locations in diverse trajectories is proposed. However, in our background study, the most commonly utilized models are Markov model, hidden Markov model, and improvements based on these models.

The Markov transition matrix was calculated according to cell organization of a goal region, and the Markov chain model was presented in [29]. To predict the further positions of a student on campus, in the recent work [3], the concept of time was considered in the prediction algorithm.

HMM- (hidden Markov model-) based trajectory prediction algorithms can be categorized into two groups: (1) parameter learning and (2) structure modeling [30]. For parameter learning, the authors proposed a mixed hidden Markov model (MHMM) [31]; and an HMM-based trajectory prediction model that the primary parameters could be altered autonomously was proposed [32]. For structure learning, based on HMM (hidden Markov model), the temporal structure of moving trajectories was modeled and recognized [33]; and the authors presented an algorithm to analyze pedestrian behaviors by constructing probabilistic models [34].

On behalf of the variations of the Markov model, the authors proposed an intermediate model coined as mix Markov model (MMM), which takes advantage of both the individual movement history and the other peoples' movements. The algorithm first clusters the users into different groups based on their historical movements, and then, the authors constructed a specific Markov model for each group; experiments show an improvement of prediction accuracy compared to the Markov model and HMM (hidden Markov model) [19]. Though the algorithm can accurately predict further locations, it is time-consuming to construct models for each group. The authors proposed a prediction model that is based on the users' history locations over some time and his/her recent locations [35]. The authors extended a prediction model named mobility Markov chain (MMC) to record the n previous locations and coined it as n -MMC. It

is a variance of the Markov model in which the future whereabouts did not only rely on the present location but also depended on the n previous locations.

Existing work trying to predict indoor movements employed the technique of pattern mining and variance of Markov chain [1, 2, 27, 35–37], but the temporal information and semantic section information in the trajectory were not used. What is more, the aforementioned methods were only validated on small or synthetic datasets (7 POIs in Lam et al. [37], only 72 hours' data in [1], only 48 hours' data in [38]). And in [35], the authors controlled the experiment environment to collect the data that the participants were aware of the experiments. Unlike the existing studies, our dataset was obtained from the general public in a large inner-city shopping center, and our proposed method considers both the spatial movement of frequent subtrajectory and shop categories which have not been studied before.

3. Preliminaries

In this section, we will introduce some basic concepts and notations used in the work are illustrated in Notations.

Definition 1 (indoor space). In general, the indoor space can be depicted as $\text{IndoorSpace} = (\text{Cell}, \text{AP}, \text{Deployment})$.

Based on the previous work [39], the indoor settings can be divided into multiple cells denoting different function areas. AP is a set of Wi-Fi access points deployed in the indoor space. Deployment stores the location details of the access points in the building.

Definition 2 (semantic section). In the indoor environment, the floor plan can be partitioned into various sections according to its functions: for example, lobby, resting, guest room, and restaurant in the hotels; men's footwear, women's fashion, food court, and entertainment sections in the shopping mall; consulting area, registration area, and outpatient area in the hospital. In the following parts of the study, we will use SC to denote semantic sections for short.

Definition 3 (AP point). In general, AP point is the abbreviation of Wi-Fi access point in the indoor building, each with a unique value separating it from the others, and it covers several semantic sections. We denote an access point AP as $\text{AP}_i = \{i, \text{Sub}_{\text{SC}_i}\}$, where Sub_{SC_i} is a subset of all the semantic sections (SC).

An instance of AP_i in our dataset is shown in Table 1; the AP point AP_{16} envelops 4 semantic sections, namely, Cosmetics, Jewelry, Cafe, and Beverage; thus, we denote AP_{16} as $\text{AP}_{16} = \{16, [\text{Cosmetics}, \text{Jewelry}, \text{Cafe}, \text{Beverage}]\}$.

Definition 4 (trajectory point). An indoor trajectory point is denoted as $P = \{\text{AP}_i, t_{\text{start}}, t_{\text{end}}\}$.

AP_i refers to the specific Wi-Fi access point, t_{start} is the time when the user walks into the Wi-Fi access point's activation range, and t_{end} is the time when the user walks out of the Wi-Fi access point's activation range.

TABLE 1: Access point illustration.

Access point	Shop	Semantic section
AP ₁₆	Estee Lauder	Cosmetics
	LANCOME	Cosmetics
	Cartier	Jewelry
	Buccellati	Jewelry
	Starbucks	Cafe
	Heytea	Beverage

Definition 5 (indoor trajectory). The indoor trajectory T is made up of a sequence of time-ordered Wi-Fi access points, where

$$T = \{P_1, P_2, \dots, P_n\}. \quad (1)$$

Definition 6 (indoor trajectory database D). Indoor trajectory database D is used to save the customers' movement information in the indoor space, $D = \{T_1, T_2, \dots, T_k\}$, $\|D\| = k$ is the count of trajectories stored in the database.

Definition 7 (indoor trajectory prediction). Given an indoor trajectory $T = \{P_1, P_2, \dots, P_n\}$, the objective is to predict the location of P_{n+1} in the $(n+1)^{\text{th}}$ instant according to the past n^{th} timestamps.

In the rest of the paper, the terms indoor trajectory and trajectory are used interchangeably unless otherwise specified.

4. Trajectory Prediction Model

In this part, our unified spatial-temporal-semantic indoor trajectory prediction model will be discussed in detail. Figure 1 is the outline of our approach. At its most basic level, the model first processes the raw trajectories into FS- (frequent subtrajectory-) based trajectories, and then, the prediction is carried out by combining the semantic information in indoor scenario and the temporal information, namely, duration of stay into a unified paradigm.

4.1. Spatial Perspective. In this section, we will introduce our prediction approach from the spatial, namely, physical movement perspective. Intuitions and the detailed frequent subtrajectory-based method are illustrated here. To begin with, we will briefly describe the basic concept of Markov process and its related features.

4.1.1. Markov Process. The underlying technology used in our work is the Markov process; the Markov process is a random process controlled by the Markov property that the future states depend on current states only, regardless of the states in the past. The Markov process is extensively utilized in discrete-time state forecasting problems. The feature is noted as the Markov property or no aftereffect. In the Markov process, the chance of shifting to a new state counts on the present state and the transition matrix only. So predicting the probability of a customer going to a new position at $(n+1)^{\text{th}}$ timestamp given his/her history location sequences can be depicted as the formula below.

$$\begin{aligned} P(P_{n+1} = \text{AP}_{n+1} \mid P_n = \text{AP}_n, P_{n-1} = \text{AP}_{n-1}, \dots, P_1 = \text{AP}_1) \\ = P(P_{n+1} = \text{AP}_{n+1} \mid P_n = \text{AP}_n). \end{aligned} \quad (2)$$

It is also called first-order Markov process, which supposes that the probability of exploring a new location depends on the last attended position only.

4.1.2. Higher-Order Markov Process. In reality, the probability of visiting a new location may depend on all the visited historical locations; hence, we are inspired to exploit the higher-order Markov process. Note that the higher-order Markov process is an extension of the Markov process such that the next latent state is stochastically dependent on the latent states multiple time steps in the past. This extension yields a richer model that can better capture the dynamics of the trajectories.

However, the higher-order Markov process is impractical for even a moderate number of time steps look back, due to the exponential increase in the size of the transition matrix and the data sparsity problem where if the set of history points taken by a user in k previous steps does not match any of other users, and then, there will be no candidate locations for him/her to recommend. In this study, we adopt a second-order Markov process that the further whereabouts to be visited are stochastically dependent on both current and previous locations.

4.1.3. Subtrajectory-Based Markov Model. We hold the assumption that the indoor trajectory data collected from users contains meaningful movement patterns that depict common behaviors of them. To be specific, trajectories repeatedly followed by users can be extracted to reveal their movement patterns. Investigating the movement patterns of a large number of people can disclose the most efficient commuter route; on the other hand, analyzing the movement patterns of every user can know more about his/her lifestyle choices and associated changes. Thus, movement patterns that represent the most frequently traveled routes are employed in this study to depict and forecast the whereabouts of customers.

A novel method called subtrajectory-based Markov model is proposed in our study to address the data sparsity and the exponential increase in the size of the transition matrix problems caused by employing higher-order Markov model. The trend of spatial granularity (the notion of spatial granularity in this study refers to the number of states in the Markov process) and the density of transition matrix when we employ new subtrajectory patterns into our model is shown in Figure 2. The granularity gradually decreases as more patterns are added into the model while the density of the transition matrix increases.

What is more, compared with the outdoor environment, the indoor setting which is composed of many of indoor elements, such as apartments, gates, stairways, walls, corridors, and floors, provides much richer constraint information. The geometric restrictions among the indoor elements are much more complex than those of outdoor environment.

FIGURE 1: Overview of spatial-temporal-semantic approach for trajectory prediction.

(a) Trend of granularity with respect to number of patterns (b) Trend of density with respect to number of patterns

FIGURE 2: Granularity vs. density.

Moreover, topological constraints, such as adjacency and inclusion correlations, are much more general in the indoor scenario. So there will be much meaningful information held in indoor trajectories. In general, taking advantage of the potential useful information will effectively enhance the accuracy of indoor trajectory tracking. Next, we will describe our subtrajectory-based Markov model in detail.

To begin with, we decompose all the trajectories into subtrajectories comprising two to five neighboring trajectory points. Then, we rank the extracted subtrajectory patterns in descending order in terms of frequency. We argue that the top-ranked subtrajectory patterns are important and meaningful, and we merge the trajectory points in the frequent subtrajectory patterns into a supernode. For example, considering an indoor trajectory sequence $A \rightarrow C \rightarrow B \rightarrow D \rightarrow C \rightarrow B$, the set of subtrajectory with size two is $\{(A \rightarrow C), (C \rightarrow B), (B \rightarrow D), (D \rightarrow C), (C \rightarrow B)\}$, and then, as it is shown in Table 2, we count the number of the subtrajectory patterns and rank them in descending order, respectively; the most frequent pattern in the example is $(C \rightarrow B)$.

After computing the top frequent subtrajectory patterns, we will deal with the raw indoor trajectories and transform

them into “frequent subtrajectory”-based trajectories. First, we determine frequent subtrajectory patterns according to their frequency, and then, we combine the trajectory points into a supernode. The process is illustrated in Figure 3.

4.2. Temporal Perspective

4.2.1. Duration of Stay. ST (duration of stay) in wireless LAN environment is an important parameter that can be used to estimate the duration a mobile user spends at the range of a particular access point AP_i . We hold the assumption that the time a customer consumes on a specific position has influence to the place where he is going to visit in the future. Consider the situation of a customer in a shopping mall; if he has already spent large time on a specific food-oriented store, it is highly possible that he will not visit a food store in the near future; otherwise, if he has only spent a few minutes on a series of food stores, we can assume that he is trying to find his favorite restaurant and we can predict the next place he is going to visit is also a food store with high confidence.

Based on the aforementioned assumption, we further enhance our model with the duration of stay at a specific

TABLE 2: An example of frequent patterns with size two.

Item	Value
(C→B)	2
(A→C)	1
(B→D)	1
(D→C)	1

FIGURE 3: Converting raw trajectories into subtrajectory-based trajectories.

access point AP_i . We analyzed the duration of stay ST of all users' login data; the result is shown in Figure 4; results show that most of the users tend to spend 5 minutes in a specific access point's range; thus, we generate several candidate lists for the next location to go based on the duration of stay ST for a specific access point AP_i . Our methodology is as follows: when performing the prediction, we first check out which access point the user currently stays, and then, we look into how long he has stayed in the range of that access point; finally, based on the candidate list generated from training data, we return the access points with top- k highest probabilities as prediction outcomes. The procedure is illustrated in Figure 5.

4.2.2. Data Processing. Before constructing the transition probability matrix, the data cleaning procedure is performed. We remove the trajectories which have less than four transitions because we hold the assumption that if the length of the trajectories is too short there will be no specific meanings between the transitions, and they will introduce inevitable errors.

4.2.3. Transition Matrix Construction. Based on Equation (2), the chance of a customer going to a new position AP_{i+1} by considering the first-order Markov Chain can be represented as follows:

$$\Pr(AP_j | AP_i) = \frac{F(AP_i \rightarrow AP_{i+1})}{\sum_{k=1}^n F(AP_i \rightarrow AP_k)}. \quad (3)$$

The numerator denotes the number of unique transform from AP_i to AP_{i+1} in the training dataset, and the divisor

FIGURE 4: Distribution: duration of stay.

is the total number of the transit trajectories that are started from AP_i ; n is the total number of Wi-Fi access points.

As for the second-order Markov model, we will build a larger transition matrix because there will be much more combinations between the last two history locations, and the probability of moving to a new location considering the last two locations is as follows:

$$\Pr(AP_{i+1} | AP_i, AP_{i-1}) = \frac{F(AP_{i-1}, AP_i \rightarrow AP_{i+1})}{\sum_{k=1}^n F(AP_{i-1}, AP_i \rightarrow AP_k)}. \quad (4)$$

4.3. Semantic Augmented Trajectory Prediction. To better enhance the quality of our prediction model, we incorporate semantic section (SC) aspect in the indoor environment into our framework.

4.3.1. Semantic Section. In real-life settings, based on the functions, the indoor building could be separated into various areas. For example, the hotel can be partitioned into check-in, resting, lobby, restaurant, entertainment, and sports sections; and the supermarket can be divided into cooked food, fruits, meat, beverage, and daily necessities sections. We hold the assumption that user's physical movements have relationship with his/her current locations. It is clear that in the train station, the majority of passengers will follow the route of buying tickets, security checking, waiting, and boarding; we assume that in the shopping mall, customers will also follow some specific patterns; for instance, and as shown in Figure 6, after strolling for a pair of shoes, the shopper may spend some time for eating and then buy a ticket for the big hits movie. So we can conclude that, in fact, people's movement shows sequential patterns [40–42]. And we propose a semantic section-based Markov model to validate our hypothesis of how accurate can we predict the semantic section a user visits.

Definition 8 (semantic section prediction based on Markov model). Consider a shopper has visited n semantic sections in an indoor building, namely, $\{Section_1, Section_2, \dots, Section_n\}$; our goal is to predict what is the next semantic section SC the customer will attend.

FIGURE 5: Working mechanism about duration of stay.

FIGURE 6: Illustration of customer's movement in the shopping mall.

Given a trajectory sequence $T = \{P_1, P_2, \dots, P_n\}$, the number of semantic section sequences is

$$N = \prod_{i=1}^n |\text{Sub}_{\text{SC}_i}|, \quad (5)$$

where $|\text{Sub}_{\text{section}_i}|$ denotes the count of semantic sections that access point AP_i covers. For instance, given a trajectory $T = \{P_1, P_2, P_3, P_4\}$ in the shopping mall, each trajectory point covers several shop categories, as it is shown in Table 3. AP_1 's activation range overlaps four shop categories of a , c , d , and e , AP_2 's activation range overlaps three shop categories of b , c , and d , and AP_3 's activation range overlaps three shop categories of a , d , and e . When predicting the next semantic section (shop category) the user is going to visit, firstly, all the possibly category chains are generated according to the previous $(n-1)$ access points; in the example, a total of $4 * 3 * 3 = 36$ possible store category sequences are formed. Then, we will use the newly generated data to train the Markov model and predict the possible sections to be explored further.

To enhance the accuracy of our prediction model, we incorporate the SC (semantic section) prediction with trajectory prediction. Each candidate's semantic section weight is

TABLE 3: An example of shop category.

	AP_1	AP_2	AP_3	AP_4
	a	b	a	a
Shop category (semantic section)	c	c	d	d
the AP covers	d	d	e	
	e			

calculated based on Algorithm 1: candidate list $L2$ is used to store possible semantic sections to be explored later, and for the candidate Wi-Fi access point AP_i , if there is a semantic section it overlaps that is in $L2$, we allocate β to its semantic section weight, where $\beta = 1/\text{length}(L2)$; lastly, we use Formula (6) to compute the unified score $\text{Pro}(\text{AP}_i)$ of each Wi-Fi access point AP_i to be explored by the customer; α is the weight coefficient we signed to semantic section.

$$\text{Pro}(\text{AP}_i) = (1 - \alpha)\text{Pro}(\text{AP}_i) + \alpha W_{\text{section}}[i]. \quad (6)$$

So far, we can predict the whereabouts for a user by giving his top- k positions P_{n+1} with the highest merged score.

5. Experiment Evaluation

In this section, we will evaluate and compare our model with baseline algorithms. The algorithms were implemented with Python 3.5. Fivefold cross-validations were carried out, and the indoor trajectory dataset was split into two sets: training dataset and testing dataset. The training dataset was utilized to construct the predicting models, and the testing set was used to evaluate the accuracy of the prediction model.

5.1. Setup

5.1.1. Dataset. The dataset we used was collected from a large shopping center with seven floors and 67 Wi-Fi access points across 90,000 square meters. The shopping center contains more than 200 shops which belong to over 34 shop categories. Table 4 shows the details of the dataset.

```

Data:  $L1$ : candidate access point list,  $L2$ : predicted shop category list
Result: each candidate  $AP_i$ 's SC weight
1  $W_{\text{section}} = []$ ,  $\beta = 1/\text{len}(L2)$ ;
2 for  $AP_i$  in  $L1$ 
3   for semantic section in  $\text{Sub}_{SC_i}$  do
4     while semantic section in  $L2$  do
5        $W_{\text{section}}[i] += \beta$ 
6     end
7   end
8 end
9 return  $W_{\text{section}}$ 

```

ALGORITHM 1: Computing SC weight.

TABLE 4: Overview of the dataset.

Item	Number
Wi-Fi points	67
Connections	907,084
Stores	200
Store categories	34
Trajectories	261,369

TABLE 5: Visiting time w.r.t. store category.

Item	% of connecting time	Average time (h)
Food	23% (2.06% each AP)	1.39
Retail	70% (1.52% each AP)	2.29
Navigational	7% (0.68% each AP)	1.00
Total	100% (1.49% each AP)	2.77

According to the layout of the mall, the nearby regions of APs can be partitioned into three major categories: food court, retail, and navigational which includes 11 APs, 46 APs, and 10 APs, respectively. The distribution of retention time with reference to each category is depicted in Table 5. People spend seven percent of their times in navigational areas, 23% for eating foods, and 70% in a retail context.

5.1.2. Parameters. Table 6 lists all the used parameters in our experiments. We set a parameter to the default value when we investigate other parameters.

5.1.3. Measurement Metrics and Baselines. To be specific, considering a target customer, our indoor trajectory predicting model calculates a score for each possible location (i.e., AP point in this study) and returns access points with the top- k maximum values as forecasting outcomes to the target customer. To evaluate the accuracy of the proposed algorithm, we use the following three metrics:

- (1) $p@1$ which denotes the percentage that we predict the correct next position
- (2) $p@5$ which denotes the percentage of times that the correct next location was in the top-five most probable locations

- (3) Mean reciprocal rank which denotes the average of the reciprocal ranking positions for the correct next location

$$P@k = \frac{\text{corrected predictions}}{\text{total number of predictions}}, \quad (7)$$

$$\text{MRR} = \frac{1}{n} \sum_{i=1}^n \frac{1}{\text{ran } k_i}.$$

Note that we did not employ the mean absolute error (MAE) as our performance metric in the experiment. The reason is that MAE is a metric for the average magnitude of the errors in a set of predictions. However, in our problem, users have no ratings about the next location and the metric is not suitable for our problem.

To validate the effectiveness of our proposed algorithm, we compare our method (STS) with the following baselines:

- (1) HMM has been proven to be effective in predicting future locations [43]
- (2) LHMM performs well in modeling the movement of users in spatial-temporal space [38]
- (3) MMC can represent the mobility behavior of an individual in a compact yet precise way [44]
- (4) TPA takes the concept of different time in a day into consideration and performs well in forecasting the students' trajectory in campus [3]
- (5) STM considers the similarity between trajectories in terms of spatial and contextual aspects [45]

5.2. Experiment Results. In this section, we will analyze the extensive experimental results. The influence of trajectories' length is first investigated; then, inspired by Wang et al. [3], we investigate the effect of trajectories' starting time to our model in prediction accuracy. Next, we check the effect of frequent subtrajectory (FS)'s size and number to the performance of our framework. After that, we study the effect of duration of stay at a specific location to the prediction performance. Then, we check the prediction accuracy of semantic section using the Markov model; after which, we

TABLE 6: Parameters.

Parameter	Values	Default
Size of {FS} (s)	2, 3, 4, 5	2
Number of {FS} (n)	1, 2, 3, ..., 20	20
Weight of shop category (α)	0.9, 0.8, ..., 0.1	0.8

investigate the location prediction accuracy of our unified framework and compare it with baseline methods. Finally, we discuss some important findings in trajectory prediction problem.

5.2.1. Influence of Trajectory Length. In this section, we will check the influence of trajectory length to the prediction precision by employing the first-order Markov model. According to Figure 7, we verified the length of trajectories from four to ten which represent the majority of dataset. As can be seen from Table 7, our evaluation matrices did not vary too much when the length of trajectories varies; the results perform stable around 0.14, 0.39, and 0.27 for $p@1$, $p@5$, and MRR, respectively. We can conclude that the performance of the Markov model is stable against the different lengths of trajectories, so in the following sections, we do not verify trajectory length again.

5.2.2. Different Starting Time. The authors take the concept of different starting time into consideration in their prediction algorithm [3]; they divided the daytime into multiple parts and construct the prediction model for each part, respectively; experiment results show that compared to basic Markov model their prediction accuracy improves nearly 100%. Inspired by the paper, we also test this idea in our study; the dataset is split into different parts according to the starting time of a trajectory; Figure 8 depicts the distribution of association times of all the trajectories in the shopping mall; as can be seen from the figure, people start to pour into the shopping mall in the middle of the day and the number of customers going into the shopping mall peaks at 14:00. Table 8 shows the prediction results considering different starting time of a trajectory. We find that the method considering the notion of time in shopping mall did not outperform the basic Markov model too much; the reason may be that unlike the special movement behaviors of a student in campus where his daily location transition is affected badly by the time when the movement occurs, e.g., in the morning, the students will probably go to the classroom, users' movements in the shopping mall are pretty flexible; people do not have restriction constraints about where to go next at a specific time. So in the rest of the study, our approach does not consider the notion of starting time.

5.2.3. Frequent Subtrajectory. In this section, we study the influence of frequent subtrajectory in terms of its size and number on prediction performance. In the experiment, we vary the size of frequent subtrajectory from two to five. Figure 9 shows the results; when the number of frequent subtrajectories increases, we get better prediction results. And an interesting finding is that adopting a smaller size

FIGURE 7: Number of trajectories w.r.t. the length.

TABLE 7: Performance of the Markov model on different lengths of trajectories.

Trajectory length	$p@1$	$p@5$	MRR
4	0.1315	0.3882	0.2630
5	0.1354	0.3996	0.2687
6	0.1390	0.3976	0.2722
7	0.1427	0.4092	0.2790
8	0.1466	0.4125	0.2786
9	0.1579	0.4211	0.2930
10	0.1441	0.4041	0.2791

FIGURE 8: Association times.

of frequent subtrajectory can achieve a better result. The explanation is that, in general, when increasing the size of the patterns, there will be fewer patterns appearing in the last two super nodes which can be referenced for the prediction.

5.2.4. Duration of Stay. Next, we investigate the influence of stay time on trajectory prediction. Figure 10 shows the results; as it can be concluded from the figure, when considering stay time, we get better results compared to that without considering stay time, that is to say, temporal aspect will affect the prediction precision.

TABLE 8: Performance of different starting times.

t_{start}	$p@1$	$p@5$	MRR
9:00–9:30	0.1354	0.4194	0.2806
9:30–10:00	0.1417	0.4467	0.2897
10:00–10:30	0.1468	0.4758	0.3007
10:30–11:00	0.1637	0.4790	0.3136
11:00–11:30	0.1669	0.4740	0.3151
11:30–12:00	0.1571	0.4604	0.3036
12:00–12:30	0.1604	0.4432	0.3001
12:30–13:00	0.1496	0.4393	0.2938
13:00–13:30	0.1580	0.4482	0.2993
13:30–14:00	0.1565	0.4407	0.2977
14:00–14:30	0.1594	0.4455	0.3017
14:30–15:00	0.1538	0.4440	0.2969
15:00–15:30	0.1638	0.4544	0.3060
15:30–16:00	0.1654	0.4603	0.3102
16:00–16:30	0.1664	0.4697	0.3126
16:30–17:00	0.1688	0.4534	0.3100
17:00–17:30	0.1702	0.4708	0.3146
17:30–18:00	0.1806	0.4673	0.3209
Without considering starting time	0.1425	0.4050	0.2765

5.2.5. Semantic Section Prediction. In this part, experiments will be conducted to validate whether the semantic sections (shop categories) visited by customers in the shopping mall are predictable or not. As depicted before, we use the generated data to train the prediction model; the results are shown in Table 9. In our experiment, we use the trajectories with length ranging from four to ten that generate 17,399 to 162,008 store category sequences, respectively, and the forecasting accuracy of the model for $p@5$ is around 50%. So we could conclude that customers' movements in the shopping mall do have certain patterns. And we can surmise that integrating the concept of shop category into our prediction model would strengthen the accuracy of location prediction.

5.2.6. Semantic Augmented Prediction and Comparisons with Baseline Method. In this part, experiments will be carried out to validate the correlation between the AP's original weight and semantic weight. Figure 11 depicts the results when the weight parameter α ranges from 0 to 1; as can be seen from the figure, the prediction accuracy shows a trend of increase in the beginning for $p@1$, $p@5$, and MRR, respectively; the best prediction results are achieved when the weight parameter α is around 0.8; after which the accuracy of our model slowly declines. Through the results, we can say that employing semantic information in the indoor environment can help improve the prediction accuracy of our model.

Table 10 shows the overall prediction results by HMM, LHMM, MM, TPA, and our method; as can be concluded from the table, HMM- and layered HMM-based approaches perform the worst; this may be resulted in inappropriate hidden states which can be investigated in future work. Markov-based approaches perform better which achieve

FIGURE 9: Prediction accuracy with respect to number of frequent subtrajectories.

FIGURE 10: Prediction precision evaluation: considering duration of stay.

TABLE 9: Performance of MM on shop category prediction.

Length	$p@1$	$p@5$	MRR
4	0.1426	0.4888	0.3108
5	0.1418	0.4920	0.3114
6	0.1418	0.4947	0.3125
7	0.1399	0.4893	0.3097
8	0.1421	0.4929	0.3116
9	0.1439	0.4991	0.3154
10	0.1450	0.5040	0.3164

the prediction accuracy of 0.405 and 0.455 for $p@5$ in terms of Markov chain model and TPA model, respectively. The similarity-based method STM achieves a pretty high prediction accuracy comparing to the aforementioned methods meaning that grouping the trajectories and constructing

FIGURE 11: Impact of weight coefficient α .

TABLE 10: Comparing with baseline methods.

Algorithms	$p@1$	$p@5$	MRR
HMM	0.0782	0.3205	0.2038
LHMM	0.0865	0.2880	0.2011
MM	0.1425	0.4050	0.2765
TPA	0.1591	0.4551	0.3037
STM	0.2937	0.5829	0.4728
Our (STS)	0.3941	0.8293	0.5797

models for each subset of trajectories will help improve the prediction accuracy. Our proposed approach performs the best which improves the prediction accuracy of 84.6%, 28%, and 55.6% against STM in terms of $p@1$, $p@5$, and MRR, respectively.

6. Conclusion and Further Work

Predicting the trajectory of users in the indoor setting has gradually drawn much attention recently with the development of Wi-Fi-enabled smartphones and free Wi-Fi services in the shopping malls. Due to the complicated indoor environment such as doors, elevators, and rooms, it is hard to perform the prediction task even in a small area. In order

to better understand the movement mobility of indoor users, we propose a unified framework that takes the spatial, temporal, and contextual information into consideration. In this study, we first adopted the idea of “supernode” that we grouped the frequent subtrajectories into a new node which can enhance the ability of our model in trajectory prediction in indoor environments and solve the data sparsity problem. Next, we studied the influence of temporal factor called duration; it further improved the performance. Then, we combined the results of location prediction and semantic section prediction into a unified framework which improved the overall prediction accuracy showing that human movements have certain patterns. Finally, extensive experiments were carried out to test the performance of our proposed approach by employing a large real-world dataset collected from an inner-city shopping center and compared it with baseline algorithms; outcomes showed that our method outperformed the baseline methods considerably in forecasting the human movement in indoor settings.

By this end, we conclude the strong points of our proposed method as follows: (1) by employing the subtrajectory-based second-order Markov model, we avoid the problem of exponential increase in transition matrix; (2) the data sparsity problem is solved by introducing the concept of “supernode”; (3) compared to the existing methods, our model provides a solution for continuous updating of the frequent patterns;

(4) semantic information is utilized for the predicting problem which is crucial in the indoor environment. Despite the strong points, there are limitations in this work: (1) although the semantic aspect is employed here, the Markov-based algorithm cannot provide similarity information between the trajectories which may denote the closeness of customers and can preserve more information; (2) we used a dataset in the shopping mall; however, the effectiveness of our method on other indoor environments is unclear. Further investigations will be carried out on various indoor datasets such as airport terminals and hospitals to validate the robustness of our approach. Moreover, the semantic aspect of the indoor environment can be evaluated by employing the tree-like model or graph structure to preserve more information.

Notations

AP_i :	Wi-Fi access point in indoor settings
SC:	Semantic section in the indoor environment
t_{start} :	Time stamp association with AP_i
t_{end} :	Time stamp disassociation with AP_i
$P = \{AP_i, t_{start}, t_{end}\}$:	Indoor trajectory point which combines access point and its association and disassociation time
$T = \{P_1, P_2, \dots, P_n\}$:	Indoor trajectory which is a series of time-ordered Wi-Fi access points
$D = \{T_1, T_2, \dots, T_k\}$:	Indoor trajectory database
$F(AP_i \rightarrow AP_j)$:	Frequency of movement from AP_i to AP_j
$Pr(AP_j AP_i)$:	The transition probability of visiting a new location
$A(a_{ij})$:	Transition matrix which contains all the transition probabilities
T_{sub} :	Subtrajectory which is composed of adjacent trajectory points
FS:	Frequent subtrajectory
ST:	Duration of stay in a location
Δ :	Threshold for ST
$W_{section}$:	The weight of a semantic section in the indoor building.

Data Availability

The trajectory data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the Natural Science Foundation of Heilongjiang Province of China under Grant No. F2015030.

References

- [1] B.-K. Ang, D. Dahlmeier, Z. Lin, J. Huang, M.-L. Seeto, and H. Shi, "Indoor next location prediction with Wi-Fi," in *The Fourth International Conference on Digital Information Processing and Communications (ICDIPC2014)*, pp. 107–113, Kuala Lumpur, Malaysia, 2014.
- [2] J. Huang, D. Dahlmeier, Z. Lin, B.-K. Ang, M.-L. Seeto, and H. Shi, "Wi-fi based indoor next location prediction using mixed state-weighted markov-chain model," *International Journal of Machine Learning and Computing*, vol. 4, no. 6, pp. 505–509, 2014.
- [3] B. Wang, Y. Hu, G. Shou, and Z. Guo, "Trajectory prediction in campus based on Markov chains," in *International Conference on Big Data Computing and Communications*, pp. 145–154, Shenyang, China, 2016.
- [4] J. Ying, H. Chen, K. W. Lin et al., "Semantic trajectory-based high utility item recommendation system," *Expert Systems with Applications*, vol. 41, pp. 4762–4776, 2014.
- [5] P. L. Jenkins, T. J. Phillips, E. J. Mulberg, and S. P. Hui, "Activity patterns of Californians: use of and proximity to indoor pollutant sources," *Atmospheric Environment. Part A. General Topics*, vol. 26, no. 12, pp. 2141–2148, 1992.
- [6] N. E. Klepeis, W. C. Nelson, W. R. Ott et al., "The National Human Activity Pattern Survey (NHAPS): a resource for assessing exposure to environmental pollutants," *Epidemiology*, vol. 11, no. 3, pp. 231–252, 2001.
- [7] N. Hernández, J. M. Alonso, and M. Ocaña, "Fuzzy classifier ensembles for hierarchical wifi-based semantic indoor localization," *Expert Systems with Applications*, vol. 90, pp. 394–404, 2017.
- [8] P. Martin, B. J. Ho, N. Grupen, S. Munoz, and M. Srivastava, "An ibeacon primer for indoor localization: demo abstract," in *Proceedings of the 1st ACM Conference on Embedded Systems for Energy-Efficient Buildings, BuildSys 2014*, pp. 190–191, Memphis, TN, USA, 2014.
- [9] C. S. Jensen, H. Lu, and B. Yang, "Indoor-a new data management frontier," *IEEE Data Engineering Bulletin*, vol. 33, pp. 12–17, 2010.
- [10] J. Yim, J. Joo, and C. Park, "A Kalman filter updating method for the indoor moving object database," *Expert Systems with Applications*, vol. 38, no. 12, pp. 15075–15083, 2011.
- [11] H. Liu, H. Darabi, P. P. Banerjee, and J. Liu, "Survey of wireless indoor positioning techniques and systems," *IEEE Transactions on Systems, Man, and Cybernetics, Part C*, vol. 37, pp. 1067–1080, 2007.
- [12] H. Xu, Y. Ding, P. Li, R. Wang, and Y. Li, "An RFID indoor positioning algorithm based on bayesian probability and K-nearest neighbor," *Sensors*, vol. 17, no. 8, p. 1806, 2017.
- [13] C. Bettstetter, "Mobility modeling in wireless networks," *ACM SIGMO787 BILE Mobile Computing and Communications Review*, vol. 5, no. 3, pp. 55–66, 2001.
- [14] R. S. Campos, L. Lovisolo, and M. L. R. de Campos, "Wi-fi multi-floor indoor positioning considering architectural aspects and controlled computational complexity," *Expert Systems with Applications*, vol. 41, no. 14, pp. 6211–6223, 2014.
- [15] T. Liu, P. Bahl, and I. Chlamtac, "Mobility modeling, location tracking, and trajectory prediction in wireless atm networks," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 6, pp. 922–936, 1998.

- [16] P. S. Prasad and P. Agrawal, "Mobility prediction for wireless network resource management," in *2009 41st Southeastern Symposium on System Theory*, pp. 98–102, Tullahoma, TN, USA, 2009.
- [17] K. Dong, Z. Ling, X. Xia, H. Ye, W. Wu, and M. Yang, "Dealing with insufficient location fingerprints in Wi-Fi based indoor location fingerprinting," *Wireless Communications and Mobile Computing*, vol. 2017, Article ID 1268515, 11 pages, 2017.
- [18] P. Samui, S. S. Roy, and V. E. Balas, *Handbook of Neural Computation*, Academic Press, 2017.
- [19] A. Asahara, K. Maruyama, A. Sato, and K. Seto, "Pedestrian movement prediction based on mixed Markov-chain model," in *Proceedings of the 19th ACM SIGSPATIAL international conference on advances in geographic information systems*, pp. 25–33, Chicago Illinois, 2011.
- [20] Q. Zhao and S. S. Bhowmick, *Sequential Pattern Mining: A Survey*, ITechnical report CAIS Nayang technological university, Singapore, 2003.
- [21] N. Mamoulis, H. Cao, G. Kollios, M. Hadjieleftheriou, Y. Tao, and D. W. Cheung, "Mining, indexing, and querying historical spatiotemporal data," in *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 236–245, 2004.
- [22] M. Morzy, "Mining frequent trajectories of moving objects for location prediction," in *Machine Learning and Data Mining in Pattern Recognition*, pp. 667–680, Springer, 2007.
- [23] J. J.-C. Ying, E. H.-C. Lu, W.-C. Lee, T.-C. Weng, and V. S. Tseng, "Mining user similarity from semantic trajectories," in *Proceedings of the 2nd ACM SIGSPATIAL International Workshop on Location Based Social Networks*, pp. 19–26, San Jose California, 2010.
- [24] P. Wang, S. Wu, H. Zhang, and F. Lu, "Indoor location prediction method for shopping malls based on location sequence similarity," *ISPRS International Journal of Geo-Information*, vol. 8, p. 517, 2019.
- [25] Z. Lv, J. Li, H. Li, Z. Xu, and Y. Wang, "Blind travel prediction based on obstacle avoidance in indoor scene," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 5536386, 14 pages, 2021.
- [26] L. Nguyen, H.-T. Cheng, P. Wu, S. Buthpitiya, and Y. Zhang, *Pnlum: System for Prediction of Next Location for Users with Mobility*, vol. 2, In Nokia Mobile Data Challenge 2012 Workshop, 2012.
- [27] J. Petzold, A. Pietzowski, F. Bagci, W. Trumler, and T. Ungerer, "Prediction of indoor movements using Bayesian networks," in *international symposium on location-and context-awareness*, pp. 211–222, Springer, 2005.
- [28] C. Jin, Z. Lin, and M. Wu, "Augmented intention model for next-location prediction from graphical trajectory context," *Wireless Communications and Mobile Computing*, vol. 2019, Article ID 2860165, 2 pages, 2019.
- [29] Y. Ishikawa, Y. Tsukamoto, and H. Kitagawa, "Extracting mobility statistics from indexed spatio-temporal datasets," in *Spatio-Temporal Database Management*, pp. 9–16, 2nd International Workshop STDBM'04, Toronto, Canada, 2004.
- [30] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–286, 1989.
- [31] R. M. Altman, "Mixed hidden Markov models," *Journal of the American Statistical Association*, vol. 102, no. 477, pp. 201–210, 2007.
- [32] S. Qiao, D. Shen, X. Wang, N. Han, and W. Zhu, "A self-adaptive parameter selection trajectory prediction approach via hidden Markov models," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 1, pp. 284–296, 2015.
- [33] S. Gong, M. Walter, and A. Psarrou, "Recognition of temporal structures: learning prior and propagating observation augmented densities via hidden markov states. In computer vision, 1999," in *Proceedings of the Seventh IEEE International Conference on Computer Vision*, pp. 157–162, Kerkyra, Greece, 1999.
- [34] D. Makris and T. Ellis, "Spatial and probabilistic modelling of pedestrian behaviour," in *Proceedings of the British Machine Vision Conference 2002 (BMVC)*, pp. 1–10, British Machine Vision Association, Cardiff, UK, 2002.
- [35] S. Gambs, M.-O. Killijian, and M. N. del Prado Cortez, "Next place prediction using mobility Markov chains," in *Proceedings of the First Workshop on Measurement, Privacy, and Mobility*, p. 3, Bern Switzerland, 2012.
- [36] O. O. Barzaïq, S. W. Loke, and H. Lu, "On trajectory prediction in indoor retail environments for mobile advertising using selected self-histories," in *Ubiquitous Intelligence and Computing and 2015 IEEE 12th Intl Conf on Autonomic and Trusted Computing and 2015 IEEE 15th Intl Conf on Scalable Computing and Communications and Its Associated Workshops (UIC-ATC-ScalCom)*, pp. 304–307, Beijing, China, 2015.
- [37] L. D. Lam, A. Tang, and J. Grundy, "Predicting indoor spatial movement using data mining and movement patterns," in *2017 IEEE International Conference on Big Data and Smart Computing (BigComp)*, pp. 223–230, Jeju, Korea, 2017.
- [38] Q. Li and H. C. Lau, "A layered hidden markov model for predicting human trajectories in a multi-floor building," in *2015 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT)*, pp. 344–351, Singapore, 2015.
- [39] N. Wang, P. Jin, Y. Xiong, and L. Yue, "A multi-granularity grid-based graph model for indoor space," *International Journal of Multimedia and Ubiquitous Engineering*, vol. 9, no. 4, pp. 157–170, 2014.
- [40] E. Cho, S. A. Myers, and J. Leskovec, "Friendship and mobility: user movement in location-based social networks," in *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 1082–1090, San Diego California USA, 2011.
- [41] M. C. Gonzalez, C. A. Hidalgo, and A.-L. Barabasi, "Understanding individual human mobility patterns," *Nature*, vol. 453, no. 7196, pp. 779–782, 2008.
- [42] C. Song, Z. Qu, N. Blumm, and A. L. Barabási, "Limits of predictability in human mobility," *Science*, vol. 327, no. 5968, pp. 1018–1021, 2010.
- [43] W. Mathew, R. Raposo, and B. Martins, "Predicting future locations with hidden Markov models," in *Proceedings of the 2012 ACM conference on ubiquitous computing*, pp. 911–918, Pittsburgh Pennsylvania, 2012.
- [44] S. Gambs, M.-O. Killijian, and M. N. del Prado Cortez, "Show me how you move and I will tell you who you are," in *Proceedings of the 3rd ACM SIGSPATIAL International Workshop on Security and Privacy in GIS and LBS*, pp. 34–41, San Jose California, 2010.
- [45] P. Wang, J. Yang, and J. Zhang, "Location prediction for indoor spaces based on trajectory similarity," in *DSIT 2021: 4th International Conference on Data Science and Information Technology*, pp. 402–407, Shanghai, China, 2021.