IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received December 20, 2017, accepted February 6, 2018, date of publication February 13, 2018, date of current version March 15, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2805831

Learning Individual Moving Preference and Social Interaction for Location Prediction

RUIZHI WU, GUANGCHUN LUO, (Member, IEEE), QINLI YANG, AND JUNMING SHAO, (Member, IEEE)

School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China

Corresponding author: Guangchun Luo (gcluo.uestc@gmail.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61403062, Grant 41601025, and Grant 61433014, in part by the Science-Technology Foundation for Young Scientists of SiChuan Province under Grant 2016JQ0007, in part by the National Key Research and Development Program under Grant 2016YFB0502300, and in part by the Postdoctoral Science Foundation of China under Grant 2014M552344, Grant 2015M580786, and Grant 2015T80973.

ABSTRACT Location prediction has attracted increasing attention in diverse fields due to its wide applications, such as traffic planning and control, weather forecasting, homeland security, and travel recommendation. Many existing algorithms forecast a user's next location by learning that user's past moving patterns. However, the individual moving patterns in many practical applications (e.g., the moving trajectory of a taxi driver) tend to be random, which poses a big challenge for location prediction. In this paper, we propose a new robust location prediction model that considers both individual preferences and social interactions (PSI) at a group level to alleviate the effect of randomness and improve the location prediction performance. Specifically, we first extract hot places of interesting (POIs) and normal POIs, respectively, via a two-stage clustering approach. To characterize exterior social interactions, an associated group is identified, and an outline of group moving patterns is then extracted based on association rule mining. Finally, the next location is predicted by learning the individual's regular patterns and group moving patterns via a pair-wise ridge regression. In contrast to the traditional approaches, our proposed algorithm has several desirable characteristics: 1) PSI provides an intuitive and quantitative way to model human movement from two aspects: the individual's internal moving preferences and group-level exterior social interactions; 2) Building upon group-level pattern mining, PSI provides a more robust prediction model by learning both individual and group trend information simultaneously, alleviating the randomness of location prediction from individual historical trajectory data only; and 3) The experimental results demonstrate that PSI achieves a better prediction performance compared to the state-of-the-art methods.

INDEX TERMS Trajectory data, location prediction, data mining.

I. INTRODUCTION

Mobility data (e.g., GPS data, WiFi signals, bus-trip records [1], credit card transactions [2], and check-in data [3] from online social networks) are increasingly collected from devices such as mobile phones, smart cards and vehicular digital records. Tracking and mining the mobility patterns in these datasets has attracted a lot of attention, from both industry and the research community [4]–[8]. For example, the use of tens of thousands of taxis equipped with GPS sensors enable traffic administrators to perceive the city's traffic flow. The goal of location prediction, as a primary task for mobility data mining, is to learn human moving patterns from the historical data to forecast future locations. Typical applications include travel recommendations, city traffic flow control, location-aware advertisements and early warnings of potential public emergencies [9]. Over the past decade, numerous location prediction algorithms have been proposed. These existing studies suggest that human moving patterns are highly regular and periodic [10]–[13], usually limited to several frequented locations such as homes, offices and restaurants. However, human movement is not always regular; it often changes dynamically through interactions with exterior factors. Consider a taxi driver, whose moving trajectory appears to be random, because a taxi driver has no idea who he will pick up and where that customer will go. Even when observing the trajectory of a taxi for an entire day or even a month, regular moving patterns are rather rare. In such cases, predicting future movement is more challenging because it is difficult to learn mobility patterns by analyzing only that taxi's past data.

Over the past century, human behavior has also been extensively studied by behaviourists [14]–[17], who often consider human behavior from two aspects: an individual's internal preferences and his or her exterior social influences. For instance, when a person makes a decision, the outcome usually depends both on that individual's own knowledge, preferences and habits as well as direct or indirect social influences from the external environment, such as suggestions from friends. Although this exterior social interaction is difficult to observe and characterize directly, its influences are indirectly reflected by the moving patterns of social groups. To further explore the taxi driver example, on the day of a popular concert in the city, the probability that a taxi driver will travel to the concert location with his next customer is higher. This suggests that the external environment (i.e., the special event) can provide hints to help with location prediction. Fortunately, such external effects can also be reflected by group patterns (i.e., on the day of the concert, for a group of taxi drivers, a frequent mobility pattern exists—travel from many places to the concert location). Therefore, external interactions can be discovered by exploring frequent moving patterns in a group. Finally, any collected external interaction information is beneficial to individual movement prediction.

In this paper, we focus on predicting GPS data patterns that do not exhibit strong individual regular moving patterns. To deal with this challenging problem, we propose a robust location prediction model that explores both an individual's moving preferences and that individual's social interactions. Specifically, to quantify the effect of exterior social interaction on individual future movement, we first identify an associated group to which a person belongs via clustering. Then, the frequent moving patterns that may reflect external social interactions are extracted based on association rule mining. These frequent group moving patterns (which characterize the group's moving trends), together with the individual's past moving patterns are finally integrated to forecast the individual's next location. Moreover, the quantitative contributions of interior preference and exterior social influence on human behavior are learned by pairwise linear ridge regression. The main contributions of this paper are as follows:

- **Identification of Hot and Normal POIs.** To extract the key semantic information from trajectory data, a two-stage clustering method is proposed that discovers hot places of interesting (POIs) and normal POIs, respectively. In contrast to traditional approaches that identify only hot POIs, our POIs extraction approach better represents the trajectory and alleviates the information loss.
- **Intuitive Model.** Motivated by behaviorist theory, an intuitive model is introduced to characterize human movement from two aspects: internal moving preferences and group-based exterior social interactions. More importantly, this model provides a quantitative way to characterize the contributions of those two aspects.
- **High Performance.** By exploring an individual's moving preferences and social interactions, the PSI model can predict a user's next location more accurately. The rationale is that group-level frequent patterns alleviate the randomness of location predictions made solely from an individual's historical trajectory, and the time-aware learning strategy further filters out outdated patterns. Experimental results on several real datasets demonstrate the superiority of our PSI approach (cf. Section IV-C).

The remainder of this paper is organized as follows: The following section briefly surveys related work. Section III presents our algorithm in detail. Section IV contains an extensive experimental evaluation. Finally we provide a brief discussion and conclude the paper in Section V.

II. RELATED WORK

Over the past decades, many approaches have been proposed for location prediction (e.g., [2], [5], [18]–[23]). Here we review only the most highly related works. In addition, we introduce some related works concerning POI extraction.

A. LOCATION PREDICTION

Early location prediction studies often resorted to time-series analysis. The basic idea is to view trajectory data as location sequences, and then use traditional time series mining techniques such as the Markov chain to predict the next item in the sequence. For example, Gambs *et al.* [22] employs the Markov chain to first model the n previous locations visited by the user, called the mobility Markov chain (MMC) model, and then predicts the next location via the transfer probability between different locations. To consider spatial movement constraints, Cheng *et al.* [4] proposes a model that built personalized Markov chains by utilizing the user's location history sequence. They introduce a spatial constraint on the localized region and factorize the transfer probability matrix of personalized Markov chains to predict the user's next movement. Although Markov chains are widely used in sequence analysis, this approach does not sufficiently consider the temporal information in GPS data. By exploring follow-up research on student card consumption trajectories in camps, Barabasi [10] and Brockmann *et al.* [11] demonstrate that individual human mobility behavior sometimes shows regular spatial and temporal rules. Based on this assumption, Morzy [19], [24] extracts regular moving rules by mining users' frequent trajectories to predict location [25]. To better integrate temporal information, Giannotti *et al.* present a model that considers the time interval between two successive user locations and builds a decision tree to model these associated rules (called a T-pattern) between the temporal and spatial spaces. Relying on the T-patterns, location prediction strategies are further proposed in [13] and [26]. It is worth mentioning that including semantic information about human movements in the trajectories has raised some concerns, and various works have attempted to discern geographically triggered, temporally triggered, or semantically triggered intentions [5], [27]. For instance, Noulas *et al.* [3], [28] aims to capture the spatiotemporal characteristics from trajectories and build a semantic trajectory pattern tree to forecast a user's next moving location. Although individual mobility pattern mining is an essential driving factor for predicting a user's future location, it is not the only pertinent factor. In fact, user movement is highly susceptible to exterior influence. For example, similar people may have similar mobility patterns. Motivated by this social phenomenon, Cho *et al.* [29] proposes a time-aware Gaussian mixture model that considers users' social activities. Jia *et al.* [30] selects the top N friends of a user using a temporal-spatial Bayesian model to learn the dynamics of friends' influences on an individual's mobility patterns, and then predict a user's future location. Moreover, an increasing group of studies have used technologies related to collaborative filtering for location prediction [23], [31]–[34]. Wang [34] proposes the RCH model by integrating a user's regularity term and conformity term; this model adopts matrix factorization to capture the influence of intimate friends. Moreover, a Bayesian inference framework is another effective strategy for location prediction. Xue *et al.* [35] proposes the STS model, which utilizes sub-trajectories to build a global transfer probability matrix and adopts a Bayesian framework to infer a user's future location.

In summary, most existing approaches that use GPS data implicitly or explicitly assume that there are some regular patterns in individual movements. However, due to the randomness of individual trajectories in real-world applications (e.g., recall the example of the movement of the taxi driver), learning regular patterns from individual trajectories with traditional approaches is a non-trivial task. More importantly, an individual's moving preferences may change dynamically over time. In light of these problems, we introduce a new model, PSI, that uses a time-aware transfer matrix to describe individual preferences and extract the skeleton information from a group associated with the user to simultaneously characterize social influence and, finally, learn the contributions of the two factors in a simple yet quantitative way via pairwise ridge regression.

B. POIS IDENTIFICATION

Trajectory data produced by GPS devices often contains huge amounts of redundant information. Extracting POIs from the trajectory data is essential for practical mobility pattern mining [20]. The mainstream approach to POI extraction is to discover POIs based on clustering. For instance, Palma *et al.* [36] first extracted features such as speed and time from trajectory data, and then found the POIs via spatial-temporal clustering. Zheng *et al.* [37] discovered the interesting locations and travel sequences in a given geo-spatial region from GPS trajectories by first clustering the non-moving points into groups; each group represents a geographic region where a user stayed over a certain time interval, and then designed a scoring system to rank the clusters. However, most existing approaches mainly focus on the attractive areas and ignore places that users do not usually visit, which is also an aspect of a user's mobility information. In this study, we introduce a two-stage clustering strategy that identifies both hot POIs and normal POIs in urban areas to extract all the semantic locations from the GPS trajectory data.

III. PSI MODEL

In this section, we introduce PSI, a robust model to predict human movement.

A. INTUITION AND OVERVIEW

Inspired by behaviorist theory, we construct a human mobility model based on an individual's moving preferences and social interactions using a simple and intuitive yet quantitative approach. The key point is to find the exterior interaction influences from the frequent mobility patterns in grouped trajectories. These group-level patterns can typically provide hints concerning the influences of external events on human mobility. Therefore, relying on a trajectory similarity measure, we first identify trajectory groups. Then, we extract the frequent grouped mobility patterns (i.e., hot and important moving patterns). Considering the varying importance of moving preferences over time, a time-aware strategy is applied. Finally, the individual moving preferences are integrated with the social interaction influence to perform location prediction. For illustration, Fig. 1 gives an overview of our prediction model. The algorithm starts by extracting the POIs (Fig. 1, left). Then, the users are grouped based on spectral clustering (Fig. 1, middle), and only the frequently grouped moving patterns are considered (e.g., those with high visiting frequencies, depicted in blue) to reflect the external social influence. Finally, the individual's moving preferences and the group-level patterns are learnt by ridge regression to predict the user's next location (Fig. 1 right). In the following, we first describe how to extract the semantic trajectory for the GPS datasets, and then elaborate on how to learn the patterns from each driving factor and combine them together.

B. EXTRACTING A SEMANTIC TRAJECTORY

GPS-based trajectory data is represented by sets of points consisting of latitude, longitude and a time stamp; however, investigating human mobility patterns from raw GPS data directly is not the best approach. A more intuitive method is to group all the GPS points into a smaller number of semantic locations. Currently, doing this involves two main strategies: grid-based partitioning and semantic POI extraction. The first approach simply partitions the study area into multiple grid cells of either equal or different sizes. However, because this type of approach ignores the relative importance of different locations, it does not well represent the semantic locations in the trajectory data. The second approach extracts the semantic locations (such as a home, office, or shopping mall) based on their visiting frequencies in the trajectory data. Here we propose a two-stage clustering

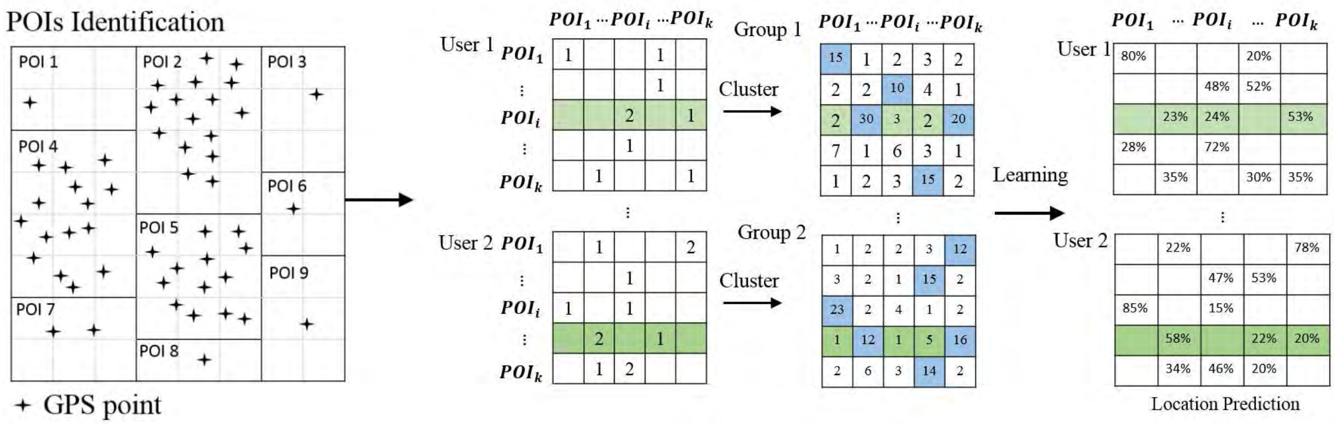


FIGURE 1. An overview of the PSI model. The first stage extracts semantic POIs from trajectory data. The second stage models an individual's moving preferences. The third stage applies the external social interaction model, in which users are first grouped and then group-trend moving patterns (e.g., the high frequent visiting pair-wise patterns) are discovered. The fourth stage performs quantitative learning via a pair-wise ridge regression algorithm to obtain the final location prediction.

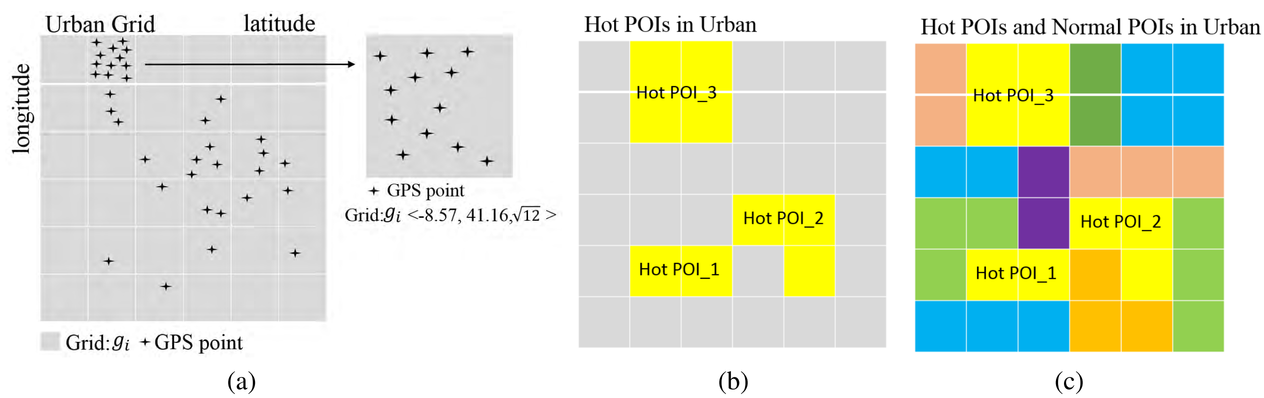
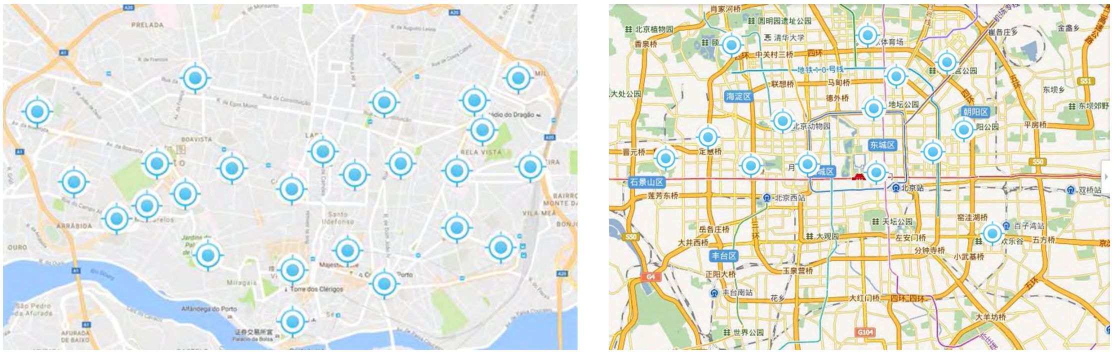


FIGURE 2. The framework to identify POIs via a two-stage clustering. (a) A trajectory in an urban grid, the cell in the right figure is an example; (b) The hot POIs (yellow areas in the figure) are identified using the DBSCAN algorithm; (c) The normal POIs (all colored areas except the yellow areas) in urban areas are identified using the K-means algorithm and then the hot and normal POIs are integrated.

strategy to group these GPS data points into hot POIs and normal POIs. Specifically, we first partition the study area into $n \times n$ (e.g., $n = 100$) grid cells. Then, we calculate the visiting frequency of each cell. Because some places are visited often while others are rarely visited, we regard the square root of the visiting frequency as the cell weight. Finally, each cell is represented by $g_i = (x_i, y_i, w_i)$, where w_i is the cell weight, and x_i and y_i are the average latitude and longitude of the points in the cell, respectively. Building upon this grid-based representation, we apply a typical density-based clustering algorithm, DBSCAN, [38] to the grid data. DBSCAN was selected due to its popularity and its ability to find arbitrarily-shaped clusters. More importantly, it is suitable for finding the dense regions (characterized as the hot POIs). Because each cell is associated with a weight, we used the weighted DBSCAN method, where w_i is used as a weight when computing the core points. After clustering, the hot POIs can be extracted. However, some cells are regarded as noise due to their low frequency. Therefore, we further apply the K-means to cluster the remaining 'noisy' cells into groups. These groups usually represent some common places that we term "normal" POIs. K-means is applied in the second stage for two reasons: (1) it maintains the complete trajectory information, and (2) unlike DBSCAN, K-means fits the sparse spatial data clustering as it allows partitioning the data space into Voronoi cells. The procedure to extract POIs is illustrated in Fig. 2. Fig. 3 further plots the extracted hot POIs for two real trajectory datasets. In contrast to the traditional approaches, our method finds both hot POIs and normal POIs; thus, it maintains more complete trajectory information. Finally, we represent each trajectory using the extracted POIs. Namely, $T_i = (t_i, poi_1, \cdots, poi_n)$, where t_i is the time of the trajectory, and poi_i is the i-th POI the user visited. Algorithm 1 shows the pseudocode for extracting POIs.

C. TIME-AWARE INDIVIDUAL MOVING PREFERENCE MODELING

Similar to traditional approaches, we extract an individual's moving preference based on frequent mobility patterns in



(a) (b)

FIGURE 3. Hot POIs in the urban area after the first clustering algorithm based on the Porto taxi dataset and the Geo-life dataset: (a) Hot POIs in Porto; (b) Hot POIs in Beijing.

Algorithm 1 Algorithm to Extract a Semantic Trajectory

Input: GPS trajectory data, K, ϵ, $minpts$.
Output: $ST = \{T_1, T_2, T_3...\}$ //ST is a semantic trajectory dataset.
Create a grid in the urban area;
for *each grid cell* g_i *in the urban grid* **do**
  //calculate g_i, P is a GPS point in cell g_i;
  $x_i = mean(P_{longitude})$;
  $y_i = mean(P_{latitude})$;
  $w_i = \sqrt{N(P)}$; //$N(P)$ is the number of points in cell g_i;
  $g_i = (x_i, y_i, w_i)$;
end
//POIs identification ;
//POI_h denotes a hot POI and POI_n denotes a normal POI;
$POI_h = DBSCAN(G(g_1, g_2, ...g_{n\times n}), \epsilon, minpts)$;
//$N(POI_h)$ is the number of hot POIs;
G_n is the noisy cells in DBSCAN algorithm;
$POI_n = K - means(G_n, (K - N(POI_h)))$;
$POIs = (POI_h, POI_n)$;
for *each trajectory* **do**
  //update semantic trajectory;
  $T_i = (t_i, poi_1, \cdots, poi_n)$;
end
$ST = \{T_1, T_2, T_3...\}$

---

historical individual trajectory data. Formally, let $f(POI_A)$ be the visiting frequency starting from the place of interest A but not visiting A for a given user, and $f(POI_A \rightarrow POI_B)$ be the number of times the user travelled from place A to place B. We define the individual's moving preference $Pr_{imp}(POI_A \rightarrow POI_B)$ as follows.

$$Pr(POI_A \rightarrow POI_B) = \frac{f(POI_A \rightarrow POI_B)}{f(POI_A)} \tag{1}$$

Considering that individual moving preferences change dynamically, more recent moving patterns are usually more important. Therefore, we introduce a decay function to characterize the relative importance of moving patterns over time. Finally, an individual's moving preference $Pr_{imp}(POI_A \rightarrow POI_B)$ is redefined as follows:

$$Pr_{imp}(POI_A \rightarrow POI_B) = \frac{\sum_i^M e^{-\gamma(t_i^{cur} - t_i)}}{\sum_j^N e^{-\gamma(t_j^{cur} - t_j)}} \tag{2}$$

where M represents the number of times the user moved from place A to place B, N is the number of times the user visited place A, t_i^{cur} is the current time when the i-th instance of one pattern occurred, and t_i is the start time of the i-th instance of that pattern. Here, γ is a constant used to control the time effect on moving preference modeling. In this study, we set $\gamma = 0.5$.

D. EXTERNAL SOCIAL INTERACTION MODELING

As stated above, human mobility is mainly driven by the individual's moving preferences and external social interactions. In this section, we elaborate on how to use the group-level sketching patterns to model the social interaction and why it works.

The social characteristics of human beings is important; through social interactions, people influence others and are, in turn, influenced by others. In the context of human mobility, individual movements are affected by the movement patterns in a community. Moreover, external environment conditions such as holidays and events are also reflected by group mobility patterns. Therefore, to enhance the predictability of human movement, we also consider group-level patterns. The group-level patterns differ from other traditional methods in two ways: (1) we consider group patterns rather than global patterns because the external effects tend to be local. In addition, global patterns may introduce considerable noise. (2) Only the most frequent moving patterns that represent the group trends are considered. The rational is that human mobility is usually affected only by the most important ideas, suggestions, or trends. Group patterns with low frequency usually characterize the diversity of mobility patterns in the group. Therefore, we identify the associated group for each user by performing trajectory clustering.

We integrate all the trajectories of each individual user in chronological order build a unique trajectory for each user. Because different trajectories often have different sampling rates and lengths, it is a non-trivial task to use traditional similarity measures such as Euclidean distance. Here, we employ the Jensen-Shannon divergence to assess trajectory similarity because it is both symmetric and allows comparing two distributions:

$$JSD(T_i||T_j) = \frac{1}{2}\Big(D_{KL}\big(T_i\|T^*\big) + D_{KL}\big(T_j\|T^*\big)\Big) \tag{3}$$

where

$$T^* = \frac{1}{2}(T_i + T_j)$$
$$D_{KL}\big(T_i\|T^*\big) = \sum_{poi} P^i(poi)log\Big(\frac{P^i(poi)}{P^*(poi)}\Big) \tag{4}$$

Here, T_i and T_j represent the i-th and j-th trajectory, respectively, poi represents a POI in the trajectories of T_i and T_j, and P^i and P^* indicate the probability distributions of trajectory of T_i and $T_i + T_j$, respectively.

Relying on the Jensen-Shannon divergence, a typical spectral clustering [39] approach is applied to find the C groups ($C = 10$ in the Porto taxi dataset, and $C = 5$ in the Geo-life dataset). Subsequently, we apply the prefix-span algorithm to each trajectory group to find the frequent pairwise moving patterns. Formally, prefix-span is a sequential pattern mining algorithm that explores prefix-projection. Unlike the Apriori algorithm used to create candidate frequent patterns in databases, the basic idea of prefix-span is to extract the prefixes of sequence items to build a projected database, and it removes a sequence if the suffix is less than a given support rate s after scanning the projected database. Then for each remaining sequence in the projected database, its ending location is used as the beginning location for new scans [40]. In this study, we employ the prefix-span algorithm to extract frequent patterns such as $(POI_A \rightarrow POI_B)$ that have a support rate of s, and then calculate the confidence of each pattern, respectively. Because the extracted patterns represent frequent moving patterns using the confidence of frequent patterns, they provide a potential way to model external social interactions. Formally, we can write

$$\begin{aligned} Pr_{esi}(POI_A \rightarrow POI_B) &= confidence(POI_A \rightarrow POI_B)_{fre} \\ &= \frac{f(POI_A \rightarrow POI_B)_{fre}}{f(POI_A)_{fre}} \end{aligned} \tag{5}$$

where $f(POI_A \rightarrow POI_B)_{fre}$ denotes the number of the frequent moving pattern $(POI_A \rightarrow POI_B)$ in a specific group (i.e., a sub-trajectory from place A to place B). Here, $f(POI_A)_{fre}$ is the number of patterns associated with place A in the group. $Pr_{esi}(POI_A \rightarrow POI_B)$ characterizes the group preference of the moving pattern.

E. LOCATION PREDICTION VIA LEARNING INDIVIDUAL MOVING PREFERENCE AND SOCIAL INTERACTION

After modeling the individual moving preferences and external social interactions, we finally integrate the two driving factors to perform trajectory prediction. To provide an intuitive yet quantitative way to analyze the importance of these two factors, we use a linear regression model for each pairwise trajectory pattern $A \rightarrow B$. Formally, this can be written as follows:

$$\begin{aligned} Pr(POI_A \rightarrow POI_B) = \beta_0 &+ \beta_1 \times Pr_{imp}(POI_A \rightarrow POI_B) \\ &+ \beta_2 \times Pr_{esi}(POI_A \rightarrow POI_B) \end{aligned} \tag{6}$$

where $Pr(POI_A \rightarrow POI_B)$ is the next location preference matrix of a user, $P_{imp}(POI_A \rightarrow POI_B)$ is the individual moving preference for the next locations and $P_{esi}(POI_A \rightarrow POI_B)$ is the external social interaction. β_i represents the corresponding coefficients. To predict user's next future movement, we learn the different quantitative contributions (i.e., different β_i values) from the current place A to the next location B. Unlike other prediction methods, such as support vector regression (SVR) or the hidden Markov model (HMM), this approach allows the model to learn the quantitative contributions of the driving factors intuitively. The traditional way for computing β_i usually adopts a least squares (LS) method. However, LS estimates are not robust with ill-conditioned input data, and may not fit sufficiently well to the test data (although it fits to the training data well). Here we use ridge regression, an improved least squares method for linear models that avoids overfitting and improves model robust due to its use of a regularizer to control the model complexity:

$$\min_{\beta} \|y - \beta X\|^2 + \lambda \|\beta\|_2^2 \quad \lambda \geq 0 \tag{7}$$

where $\beta = [\beta_0 \ \beta_1 \ \beta_2]^T$, λ is a constant, and y is the observed pairwise movement pattern matrix for each user. $X = [X_{imp} \ X_{esi}]$ contains the individual moving preference matrix X_{imp} and external social interaction matrix X_{esi} for each user, respectively.

To train the model, we must determine the training dataset. In this study, we use a sliding window strategy where the data within a given window size (e.g., trajectory data for three months) is regarded as training data and used to predict the next location within a window size (e.g., in the next week or month). The selected window size depends on the trajectory data and the user's application requirements. In addition, considering that human movement usually differs substantially on workdays and weekends, in our PSI model, we analyze these two moving patterns separately.

F. PSI ALGORITHM

In this section, we describe the PSI algorithm, which involves the following steps:

1) **POI Identifications**. First, the hot POIs and normal POIs that represent the semantic locations in trajectory data are identified. In contrast to most traditional approaches, which extract only the hot POIs, we use

a two-stage clustering (i.e., density-based DBSCAN and partitioning-based K-means) approach to identify the hot and normal POIs, respectively. The maximum total number of POIs (i.e., K) is set to 100 in all the experiments, and the hot POIs are determined by the DBSCAN algorithm. Then, the number of normal POIs is set to K = (100 - number of hot POIs) for the K-means algorithm.

2) **Individual Moving Preference Modeling**. Considering that individual moving preferences change dynamically, an individual's moving preferences over time are characterized by Eq. (2).
3) **External Social Interaction Modeling**. To characterize the external social effect, the trajectory is first grouped into several clusters building upon the Jensen-Shannon divergence. Then, the frequent group-level patterns are mined by the prefix-span algorithm.
4) **Location Prediction**: Building upon the modeling of individual moving preferences and external social interactions, the ridge regression is introduced to predict the next location. More importantly, the contributions of both individual moving preferences and external social interactions are measured.

Finally, the pseudocode of the PSI algorithm is summarized in Algorithm 2.

G. COMPLEXITY ANALYSIS

To extract POIs, we need to perform the two-stage clustering (DBSCAN and K-means). Therefore, the time complexity depends on the number of cells considered (e.g., N= 100 ×100), whose complexity is $O(N \cdot log(N))$. The most time-consuming part of trajectory clustering is the SVD decomposition of spectral clustering, which is $O(N_1 \cdot d^2 \cdot log(d))$ of the time complexity of spectral clustering, where N_1 is the number of users. For frequent group pattern mining, the running time is approximately $O(N_1^2)$. Therefore, the theoretical total running time is approximately $O(N \cdot log(N) + N_1 \cdot d^2 \cdot log(d) + N_1^2)$.

Algorithm 2 PSI Prediction Algorithm

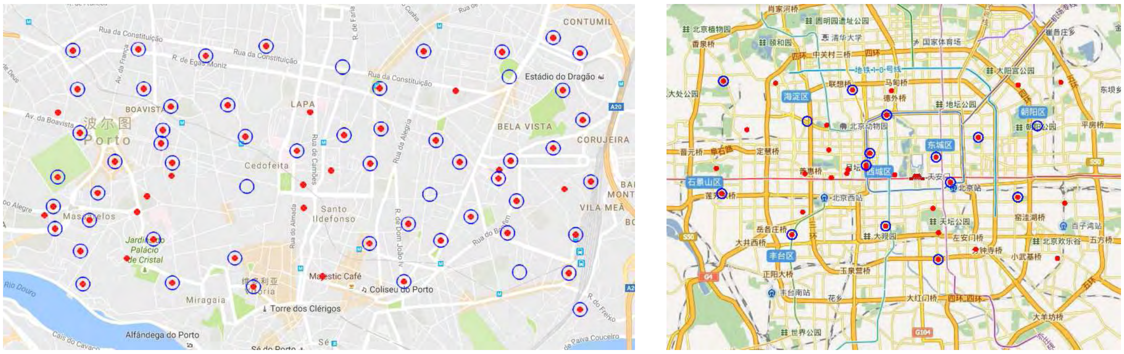
Input: $ST = \{T_1, T_2, T_3 ...\}, C, s, T, \gamma, \lambda$.
Output: Predicted location
//Part ST according T $(D_{train}, D_{test}) = Partition(ST, T)$;
for *each D in* (D_{train}, D_{test}) **do**
  //find individual moving preference;
  for *each user in dataset* **do**
    Find the individual moving preference, such as $Pr_{imp}(POI_A \rightarrow POI_B)$;
    calculating $Pr_{imp}(POI_A \rightarrow POI_B)$ using Eq. 2 and γ;
  end
  //Cluster users to find group patterns;
  for *each user in a dataset* **do**
    //calculating distance matrixDis;
    $Dis_{ij} = JSD(T_i || T_j)$;
    $JSD(T_i || T_j) = \frac{1}{2}\Big(D_{KL}\big(T_i \| T^*\big) + D_{KL}\big(T_j \| T^*\big)\Big)$;
  end
  //G is group;
  $G = spectralcluster(Dis, \mathrm{C})$;
  for *each G* **do**
    //mining frequent patterns from the group;
    $(POI_A \rightarrow POI_B)_{fre} = prefixspan(G, s)$;
    $Pr_{esi}(POI_A \rightarrow POI_B) = confidence(POI_A \rightarrow POI_B)_{fre}$;
  end
  $Pr(POI_A \rightarrow POI_B) = \beta_0 + \beta_1 \times Pr_{imp}(POI_A \rightarrow POI_B) + \beta_2 \times Pr_{esi}(POI_A \rightarrow POI_B)$;
  where $\beta = [\beta_0\ \beta_1\ \beta_2]^T$;
  Learning β using Eq.7 and λ;
end
Predicting location using β , Pr_{imp} and Pr_{esi};

IV. EXPERIMENT

To comprehensively study PSI's performance, we conducted experiments on two real-world datasets: Porto taxi GPS trajectory data and Geo-life data. We compared PSI with the Next-place model [22], the Prediction of Moving Object Location model (PMOL) [19], the Time-weight Collaborative Filter Model (TWFM) [31] and the Sub-trajectory Synthesis Model (STS) [35]. An introduction to and the parameter settings of these models is provided in Section IV-F. All the experiments were performed on a personal computer with a 3.5 GHz CPU and 8 GB of RAM. In general, in this study, we set the parameters as follows. C is the number of groups in the dataset; we set $C = 10$ in the Porto dataset and $C = 5$ in the Geo-life dataset. s is the support of the prefix-span algorithm; we set $s = 0.01$. T is the length of sliding time window, and γ is the decay-rate factor in Eq.(2). The effects of these parameters on the prediction performance will be further investigated in Section IV-G. We set the parameters of the compared algorithms to the values suggested by their authors.

A. DATASET

In this study, we focus on two real GPS trajectory datasets: Porto taxi data and Geo-life data [41], [42]. The Porto taxi dataset contains the trajectories of 442 taxis from July 2013 to June 2014 in the city of Porto, Portugal. This data represents approximately 1.7 million taxi rides. These mobile data terminals are installed in each vehicle and provide GPS localization and taximeter state information. Electronic dispatch systems make it easy to see where a taxi has been. One objective of this dataset is to predict the next destination of each taxi. For more details, please refer to the website (https://www.kaggle.com/c/pkdd-15-predict-taxi-service-trajectory-i). The Geo-life dataset is another GPS trajectory dataset that was collected by Microsoft Research Asia for 182 users over a period of more than five years (from



(a) (b)

FIGURE 4. An illustration of the location prediction for a given user. (a) A user in the Porto taxi data; (b) A user in the Geo-life dataset.

TABLE 1. The statistics of two trajectory datasets.

Dataset	Porto Taxi Dataset	Geo-life Dataset
Time Start	2013/07/01	2007/04/01
Time End	2014/06/30	2012/08/01
Number of Trajectories	1710670	19409

April 2007 to August 2012). Table 1 further lists the statistics of two trajectory datasets. We split both datasets into training data and test data at an 8:2 ratio, respectively.

B. EVALUATION METRICS

To quantitatively evaluate the PSI, we adopted the following performance metrics:

- *Prediction performance*. We evaluated the performance of the PSI model in terms of accuracy (including Acc, Acc@5, and Acc@all), precision (mean average precision (MAP)), rank (first place rank (FPR)), F1-score and record.
- *Sensitivity*. We tested the sensitivity of the prediction performance to parameter variation.

Acc@topP is the percentage of accurate predictions for a list of predictions with length P. We selected Acc@1, Acc@5 and Acc@all to describe the accuracy of the PSI model. FPR represents the prediction performance of the top-1 location rank. The formal definition is as follows.

$$FPR = (K - rank(L) + 1)/K \tag{8}$$

where K is the number of locations ($K = 100$ in this study), and $rank(L)$ is the position of the top-1 location in the predicted list. AFPR is the average FPR.

The MAP, derived from information retrieval, is an evaluation index that represents the relationships between all the predicted moving locations and the real moving locations. Formally,

$$MAP = (\sum_{r=1}^{n}(P(r) * rel(r)))/K \tag{9}$$

where K is the number of locations, $rel(\cdot)$ is a binary function on the relevance of a given rank, and $P(r)$ is the precision for a given rank.

The F1-score is an index that considers both the precision and the recall rate. It records the ratio between the number of true predictions of moving locations and the total number of real moving locations.

C. PSI'S LOCATION PREDICTION PERFORMANCE

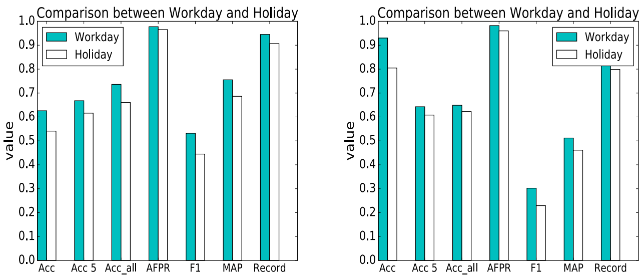
1) HIGH PREDICTION PERFORMANCE

To evaluate the performance of the PSI when predicting the next location, we first performed experiments on both Porto taxi data and Geo-life datasets with different settings: (1) different window sizes for training: three months, four months, five months and six months on the Porto dataset, eighteen months, twenty-four months and thirty months on the Geo-life dataset. The Geo-life dataset spans a longer time and contains sparser trajectory data (every user has a 20-day trajectory); and (2) with and without a decay function. Table 2 and Table 3 summarize the prediction performances in terms of the different evaluation measures on the two datasets (in Table 2 and table 3 M means months), respectively. Other parameters were set as follows: $k = 100$, $s = 0.01$, $\gamma = 0.5$, $\lambda = 0.5$. The time ranges of the predictions are one month for the Porto dataset and six months for the Geo-life dataset. From these tables, we can observe that PSI achieves a good prediction performance, especially for the top-1 location prediction (with an AFPR > 96%). In addition, we can see that the decay function has only a slight effect on the prediction performance. Regarding the selection of training data, when more data are used for training, the results show a slight improvement in the prediction performance. To better illustrate the results, Fig. 4 further plots the location prediction for one user from the Porto dataset and Geo-life dataset; the red points indicate the predicted the locations while the blue circles demonstrate the ground truth (i.e., the real moving locations). We can observe that the prediction is good because most of the predicted locations and the real moving locations match.

Furthermore, we evaluated the prediction performances for workdays and holidays, respectively. Human movements on workdays are usually more regular. Fig. 5 compares

TABLE 2. PSI's location prediction performance on the Porto taxi dataset.

Time Variation	Without Decay Function				With Decay Function			
Time Window	3M	4M	5M	6M	3M	4M	5M	6M
Acc	0.6093	0.6048	0.6033	0.6015	0.6110	0.6066	0.6049	0.6018
Acc@5	0.6841	0.6880	0.6908	0.6994	0.6891	0.6949	0.6984	0.7005
Acc@all	0.7400	0.7452	0.7490	0.7510	0.7429	0.7497	0.7544	0.7572
AFPR	0.9640	0.9697	0.9734	0.9761	0.9644	0.9699	0.9737	0.9763
MAP	0.7190	0.7381	0.7455	0.7549	0.7281	0.7307	0.7446	0.7538
F1	0.5004	0.5025	0.5035	0.5041	0.5008	0.5032	0.5040	0.5045
Record	0.9211	0.9328	0.9405	0.9462	0.9216	0.9333	0.9411	0.9468



(a) (b)

FIGURE 5. Comparison of PSI's prediction performance on workdays and holidays. (a) Porto taxi dataset (b) Geo-life dataset.

TABLE 3. PSI's location prediction performance on the Geo-life dataset.

Time Variation	Without Decay Function			With Decay Function		
Time Window	18M	24M	30M	18M	24M	30M
Acc	0.7489	0.8302	0.9300	0.7549	0.8306	0.9303
Acc@5	0.6262	0.6353	0.6364	0.6321	0.6381	0.6427
Acc@all	0.6346	0.6449	0.6486	0.6393	0.6415	0.6490
AFPR	0.9765	0.9803	0.9820	0.9785	0.9822	0.9903
MAP	0.3804	0.4650	0.5066	0.3872	0.4753	0.5119
F1	0.2688	0.2726	0.2984	0.2748	0.2782	0.3023
Record	0.7256	0.8080	0.8353	0.7242	0.8097	0.8381

the prediction performances for workdays and holidays. As expected, Fig. 5 shows that PSI achieves better prediction performance for workdays in terms of the different evaluation measures on both datasets.

2) LOCATION PREDICTION WITH DIFFERENT TIME RANGES

In addition to predicting the moving trajectory for either workdays or holidays, we further considered the time ranges when predicting future movements. Here, considering the two datasets, we predicted the next location of a given user in 10 days to 50 days when using the Porto dataset and 1 months, 3 months, 6 months and 9 months when using the Geo-life dataset. Table 4 and Table 5 summarizes the prediction performances. Clearly, the PSI captures both the individual moving preferences and social interaction influences; thus, it achieves high prediction performances over different time ranges.

TABLE 4. Location prediction performance with different time ranges on the Porto dataset.

Time	Acc@1	Acc@5	Acc@all	AFPR	F1	MAP	Record
10 days	0.4988	0.6383	0.6810	0.9729	0.3830	0.7103	0.9278
20 days	0.5644	0.6689	0.7194	0.9694	0.4510	0.7209	0.9283
30 days	0.6110	0.6891	0.7429	0.9644	0.5008	0.7281	0.9216
40 days	0.6126	0.6954	0.7523	0.9595	0.5146	0.6898	0.9097
50 days	0.6284	0.6988	0.7583	0.9574	0.5210	0.6761	0.9036

TABLE 5. Location prediction performance with different time ranges on the Geo-life dataset.

Time	Acc@1	Acc@5	Acc@all	AFPR	F1	MAP	Record
1 months	0.9278	0.6193	0.6209	0.9836	0.2908	0.4877	0.7987
3 months	0.9265	0.6441	0.6451	0.9894	0.2984	0.5002	0.8201
6 months	0.9303	0.6427	0.6490	0.9903	0.3023	0.5119	0.8381
9 months	0.9576	0.6714	0.6516	0.9875	0.3156	0.5131	0.8778

3) TREND PREDICTION

Here, we further evaluated the group trend movement by merging all the individual predicted locations together. Group movement provides a more comprehensive way to understand human moving patterns. In Fig. 6, we draw a heat map of the predicted location in a global manner, and compare it to the real movements for all users. It is interesting to note that the predicted group mobility patterns are highly successful.

D. QUANTITATIVE ANALYSIS OF INDIVIDUAL MOVING PREFERENCE AND EXTERNAL SOCIAL INTERACTION

In contrast to previous approaches, PSI provides a simple quantitative way to investigate the effect of individual moving preferences (IMP) and external social interactions (ESI) on human movement, respectively. For each user, we derive the coefficients β_1 and β_2, which characterize the contributions of the two driving factors, respectively. The contribution of IMP is calculated by $\frac{\beta_1}{\beta_1+\beta_2}$, and the contribution of ESI is calculated by $\frac{\beta_2}{\beta 1+\beta_2}$. Fig. 7(a) shows a plot of the portions of the individual moving preferences and social interactions for the movement of some taxis in Porto, for example, the red part shows the proportion of individual moving preferences and the yellow part shows the proportion for social interactions. Fig. 7(b) shows the individual moving preference of all the taxis in the dataset, and Fig. 7 (c) and (d) show the results

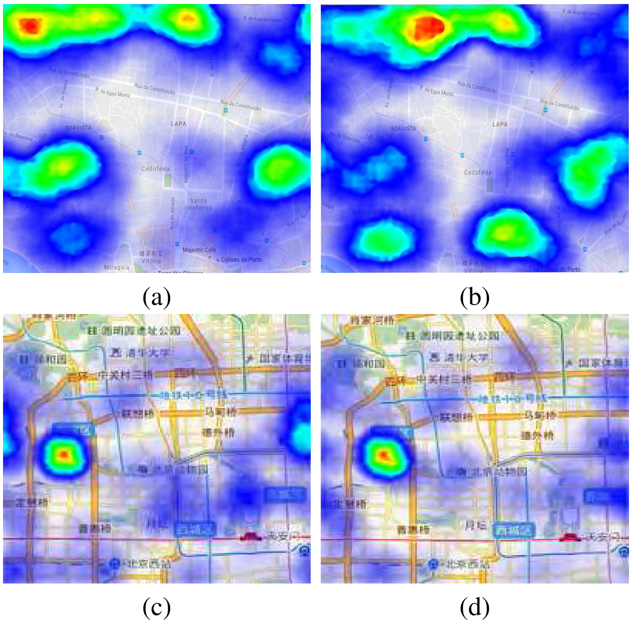


FIGURE 6. **The group movement distributions on two datasets. (a) Group prediction location distribution in Porto; (b) Group true location distribution in Porto; (c) Group prediction location distribution in Beijing; (d) group true location distribution in Beijing.**

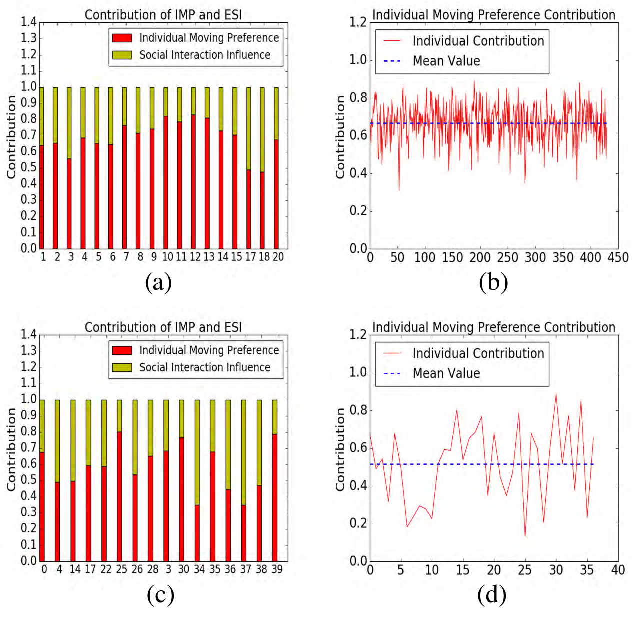


FIGURE 7. **Illustration of the contributions of individual moving preferences and external social interactions when determining human movement from the Porto and Geo-life datasets. a) Example users (Porto). (b) All users (Porto). (c) Example users (Beijing). (d) All users (Beijing).**

on the Geo-life dataset. From Fig. 7, we can easily see the external social influence on human mobility, where the average contribution of individual moving preference accounts for 66.9% in the Porto data and 51.2% in the Geo-life data. In general, some users have strong individual regular moving patterns, while some people tend to follow the group mobility. However, for both datasets, external social interactions play an important role in human mobility.

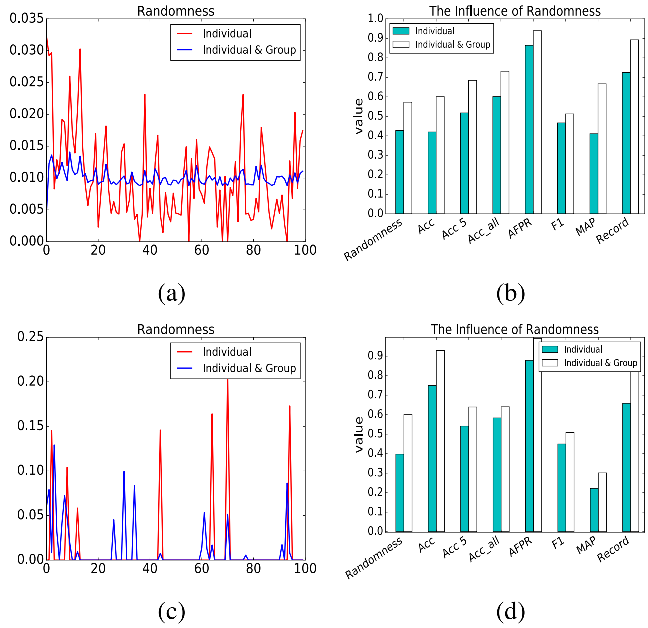


FIGURE 8. **Randomness analysis on two datasets. (a) The randomness of the next location by comparing individual patterns or individual & group patterns of a taxi in Porto; (b) The different prediction performances on the Porto dataset using the two different strategies: individual patterns and individual plus group patterns; (c), (d) The same types of plots using the Geo-life dataset.**

E. RANDOMNESS ANALYSIS

To further evaluate how PSI helps location predictions on irregular GPS trajectory data (i.e., the derived social interaction supports a robust prediction, which alleviates the random movement prediction based on historical individual movements only), we perform a randomness analysis in this section. Specifically, we use entropy (see. Eq. 10) to characterize the moving distribution to reflect the randomness of a given individual user and analyze the influence of mobility randomness on location prediction. For example, we assume that a user starts from a location called POI_A and predict that user's next location based on individual historical patterns alone and on the individual plus group information. Fig. 8(a) plots the randomness of movement with individual patterns and for individual & group patterns for a given taxi driver in Porto. Fig. 8(b) plots the prediction performances using the two different strategies, and Fig.8 (c) and (d) show the results on the Geo-life dataset. From Fig.8, we can observe that the randomness of the next location decreases when the algorithm also considers external social interactions. More importantly, the resulting decrease in randomness obtains a better prediction performance, which indicates that our PSI model captures the social interaction well and achieves a good location prediction on the GPS trajectory data even when each individual movement pattern is highly random.

$$H(POI_A) = -\sum_{i=1}^{K} p_i \log p_i \tag{10}$$

TABLE 6. Prediction performances of different algorithms on the Porto dataset.

Model	Acc@1	Acc@5	Acc@all	AFPR	F1	MAP	Record
PSI	**0.6110**	**0.6891**	**0.7429**	**0.9644**	**0.5008**	**0.7281**	**0.9216**
1-HMM	0.3902	0.4357	0.4895	0.7010	0.2973	0.4717	0.8010
2-HMM	0.4537	0.4709	0.5711	0.7977	0.3477	0.5410	0.8224
PMOL	0.4430	0.4810	0.5600	0.8091	0.3536	0.5013	0.7974
TWFM	0.3821	0.4217	0.4511	0.6813	0.2219	0.4673	0.6687
STS	0.4244	0.6577	0.7243	0.9121	0.1619	0.7133	0.8789

TABLE 7. Prediction performances of different algorithms on the Geo-life dataset.

Model	Acc@1	Acc@5	Acc@all	AFPR	F1	MAP	Record
PSI	**0.9303**	**0.6427**	**0.6490**	**0.9903**	**0.5119**	**0.3023**	**0.8381**
1-HMM	0.6243	0.4585	0.5145	0.7676	0.3041	0.2127	0.6669
2-HMM	0.7218	0.5375	0.5732	0.8063	0.4504	0.2524	0.7269
PMOL	0.7430	0.5214	0.5519	0.8517	0.4236	0.2213	0.6974
TWFM	0.6421	0.4817	0.4711	0.7236	0.2119	0.1713	0.6087
STS	0.6460	0.5126	0.5128	0.8791	0.2576	0.2501	0.6564

F. COMPARISONS WITH OTHER LOCATION PREDICTION APPROACHES

In this section, to further demonstrate the benefits of our location prediction model, we compare PSI with the Next-place model (including the one-order HMM and two-order HMM, abbreviated by 1-HMM and 2-HMM, respectively) [22], the Prediction of Moving Object Location model (PMOL) [19] based on frequent pattern mining, the Time Weight Collaborative Filter Model (TWFM) [31] and the Sub-trajectory Synthesis Model (STS) [35]. The 1-HMM, 2-HMM and PMOL models focus on individual moving preferences via a Markov model and on individual pattern mining, while the TWFM and STS models adopt external global mobility movement to infer the individual's next location via collaborative filtering and the Bayesian inference framework. It should be noted that the original STS model focused on destination prediction; therefore, we made a modification to the STS model and use a 2-length sub-trajectory to predict the user's next location. Table 6 and Table 7 summarize the prediction performances in terms of different measures on both real-world datasets (the number of states is 100 in the Next-place model, and the number of grids is 100*100 in the STS model). From these tables, we can observe that our model achieves the best results. These good performances may be due to PSI integrating the individual moving preferences and external social interactions, thus enhancing the predictability of human movement patterns. More importantly, the group-level information contains only the trend information; other "noisy" information is filtered out. Therefore, external social influence is well captured and the model achieves a high prediction performance.

G. SENSITIVITY TO PARAMETERS

In this section, we perform sensitivity analyses of PSI regarding the different parameters on the Porto dataset, including the number of POIs (K), the decay rate factor γ, the support

TABLE 8. A sensitivity analysis of K on the prediction performance.

K	Acc@1	Acc@5	Acc@all	AFPR	F1	MAP	Record
50	0.6413	0.6809	0.7910	0.9825	0.5411	0.7637	0.9621
75	0.6364	0.6851	0.7681	0.9714	0.5272	0.7528	0.9316
100	0.6110	0.6891	0.7429	0.9644	0.5008	0.7281	0.9216

TABLE 9. A sensitivity analysis of γ on the prediction performance.

γ	Acc@1	Acc@5	Acc@all	AFPR	F1	MAP	Record
0.2	0.6093	0.6846	0.7389	0.9640	0.5008	0.7186	0.9207
0.4	0.6110	0.6888	0.7429	0.9644	0.5008	0.7281	0.9215
0.5	0.6110	0.6891	0.7429	0.9644	0.5008	0.7281	0.9216
0.6	0.6089	0.6881	0.7393	0.9640	0.5006	0.7281	0.9211
0.8	0.6081	0.6823	0.7392	0.9640	0.5012	0.7184	0.9211

TABLE 10. A sensitivity analysis of s on the prediction performance.

s	Acc@1	Acc@5	Acc@all	AFPR	F1	MAP	Record
10%	0.5013	0.5321	0.6138	0.8891	0.4126	0.5713	0.8574
5%	0.5321	0.6017	0.6437	0.9013	0.4429	0.6013	0.8987
1%	0.6110	0.6891	0.7429	0.9644	0.5008	0.7281	0.9216
0.5%	0.4511	0.4722	0.5177	0.8013	0.3754	0.5103	0.6641

TABLE 11. A sensitivity analysis of λ on the prediction performance.

λ	Acc@1	Acc@5	Acc@all	AFPR	F1	MAP	Record
0.1	0.6110	0.6887	0.7421	0.9644	0.5008	0.7280	0.9216
0.3	0.6109	0.6886	0.7419	0.9644	0.5008	0.7280	0.9216
0.5	0.6110	0.6891	0.7429	0.9644	0.5008	0.7281	0.9216
0.7	0.6109	0.6884	0.7417	0.9644	0.5008	0.7281	0.9216
0.9	0.6110	0.6883	0.7416	0.9644	0.5008	0.7281	0.9216

parameter (s) for determining the frequent group-level patterns and the regularization parameter λ for ridge regression. Tables 8 to 11 show the prediction performances when varying the values of the different parameters. From these tables, we can see that PSI is quite robust to the number of POIs, the decay rate factor and to λ: the prediction performances with different values remain stable. However, the prediction performance is highly sensitive to the parameter s. Lower s values, derive more frequent patterns; thus, they help alleviate the randomness of patterns learned from historical individual moving preferences. However, the prediction performance decreases with very low s values, where almost no group information is used for prediction. These values tend to introduce noise, which negatively affects the final location prediction.

V. DISCUSSION AND CONCLUSION

The proposed PSI method builds upon both *Individual Moving Preference* and *External Social Interaction*. PSI utilizes external social interaction as a group-level filter to reduce the randomness of individual mobility patterns. Considering both driving factors when modeling human mobility is a natural fit, because both factors are important in determining human movements based on the studies of behaviourists. Most other location prediction algorithms assume that regular patterns exist in the individual trajectory data, which is challenging in some real-world scenarios. Although our PSI model can be viewed as a special combination of the internal and external factors for modeling human moving patterns, it largely differs from the traditional approaches. One main difference is that external social interactions are modeled using only the sketching group patterns instead of using the full information. Another attractive property of PSI is that it provides a quantitative way to evaluate the importance of the driving factors for human mobility at each user level. Through comprehensive experiments, we have shown that PSI outperforms some other location prediction methods and that it provides an intuitive way to analyze the results. In future work, we plan to focus on analyzing evolving human mobility patterns based on data stream mining techniques.

ACKNOWLEDGMENTS

The authors would like to express their gratitude to all those who have helped them during the writing of this paper. They thank the editors and reviewers for all their help.

REFERENCES

[1] Y. Zheng, L. Capra, O. Wolfson, and H. Yang, "Urban computing: Concepts, methodologies, and applications," *ACM Trans. Intell. Syst. Technol.*, vol. 5, no. 3, p. 38, 2014.

[2] N. J. Yuan, Y. Wang, F. Zhang, X. Xie, and G. Sun, "Reconstructing individual mobility from smart card transactions: A space alignment approach," in *Proc. IEEE 13th Int. Conf. Data Mining*, Dec. 2013, pp. 877–886.

[3] A. Noulas, S. Scellato, N. Lathia, and C. Mascolo, "Mining user mobility features for next place prediction in location-based services," in *Proc. IEEE 12th Int. Conf. Data Mining*, Dec. 2012, pp. 1038–1043.

[4] C. Cheng, H. Yang, M. R. Lyu, and I. King, "Where you like to go next: Successive point-of-interest recommendation," in *Proc. IJCAI*, vol. 13. 2013, pp. 2605–2611.

[5] J. J.-C. Ying, W.-C. Lee, T.-C. Weng, and V. S. Tseng, "Semantic trajectory mining for location prediction," in *Proc. 19th ACM SIGSPATIAL Int. Conf. Adv. Geograph. Inf. Syst.*, 2011, pp. 34–43.

[6] H. K. Pao, J. Fadlil, H. Y. Lin, and K. T. Chen, "Trajectory analysis for user verification and recognition," *Knowl.-Based Syst.*, vol. 34, no. 5, pp. 81–90, 2012.

[7] K. Zheng, Z. Huang, A. Zhou, and X. Zhou, "Discovering the most influential sites over uncertain data: A rank-based approach," *IEEE Trans. Knowl. Data Eng.*, vol. 24, no. 12, pp. 2156–2169, Dec. 2012.

[8] H. Su, K. Zheng, J. Huang, H. Wang, and X. Zhou, "Calibrating trajectory data for spatio-temporal similarity analysis," *VLDB J.*, vol. 24, no. 1, pp. 93–116, Feb. 2015.

[9] L. Zhao, Q. Sun, J. Ye, F. Chen, C.-T. Lu, and N. Ramakrishnan, "Multi-task learning for spatio-temporal event forecasting," in *Proc. 21th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2015, pp. 1503–1512.

[10] A.-L. Barabasi, "The origin of bursts and heavy tails in human dynamics," *Nature*, vol. 435, no. 7039, pp. 207–211, May 2005.

[11] D. Brockmann, L. Hufnagel, and T. Geisel, "The scaling laws of human travel," *Nature*, vol. 439, no. 7075, pp. 462–465, 2006.

[12] T. Kim, Y. Yue, S. Taylor, and I. Matthews, "A decision tree framework for spatiotemporal sequence prediction," in *Proc. 21th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2015, pp. 577–586.

[13] F. Giannotti, M. Nanni, F. Pinelli, and D. Pedreschi, "Trajectory pattern mining," in *Proc. 13th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2007, pp. 330–339.

[14] G. A. Miller and F. C. Frick, "Statistical behavioristics and sequences of responses," *Psychol. Rev.*, vol. 56, no. 6, p. 311, 1949.

[15] B. K. Scarborough, T. Z. Like-Haislip, K. J. Novak, W. L. Lucas, and L. F. Alarid, "Assessing the relationship between individual characteristics, neighborhood context, and fear of crime," *J. Criminal Justice*, vol. 38, no. 4, pp. 819–826, 2010.

[16] Q. Ke and B. J. Oommen, "Logistic neural networks: Their chaotic and pattern recognition properties," *Neurocomputing*, vol. 125, no. 3, pp. 184–194, 2014.

[17] J. D. Sterman, "Deterministic chaos in models of human behavior: Methodological issues and experimental results," *Syst. Dyn. Rev.*, vol. 4, nos. 1–2, pp. 148–178, 1988.

[18] G. Yavaş, D. Katsaros, Ö. Ulusoy, and Y. Manolopoulos, "A data mining approach for location prediction in mobile environments," *Data Knowl. Eng.*, vol. 54, no. 2, pp. 121–146, 2005.

[19] M. Morzy, "Prediction of moving object location based on frequent trajectories," in *Proc. Int. Symp. Comput. Inf. Sci.*, 2006, pp. 583–592.

[20] Y. Zheng, "Trajectory data mining: An overview," *ACM Trans. Intell. Syst. Technol.*, vol. 6, no. 3, p. 29, 2015.

[21] S. Rendle, C. Freudenthaler, and L. Schmidt-Thieme, "Factorizing personalized Markov chains for next-basket recommendation," in *Proc. 19th Int. Conf. World Wide Web*, 2010, pp. 811–820.

[22] S. Gambs, M.-O. Killijian, and M. N. del Prado Cortez, "Next place prediction using mobility Markov chains," in *Proc. 1st Workshop Meas., Privacy, Mobility*, 2012, p. 3.

[23] D. Lian, X. Xie, V. W. Zheng, N. J. Yuan, F. Zhang, and E. Chen, "CEPR: A collaborative exploration and periodically returning model for location prediction," *ACM Trans. Intell. Syst. Technol. (TIST)*, vol. 6, no. 1, p. 8, 2015.

[24] M. Morzy, "Mining frequent trajectories of moving objects for location prediction," in *Proc. Int. Workshop Mach. Learn. Data Mining Pattern Recognit.*, 2007, pp. 667–680.

[25] L. Wang, K. Hu, K. Tao, and X. Yan, "Mining frequent trajectory pattern based on vague space partition," *Knowl.-Based Syst.*, vol. 50, no. 3, pp. 100–111, 2013.

[26] A. Monreale, F. Pinelli, R. Trasarti, and F. Giannotti, "WhereNext: A location predictor on trajectory pattern mining," in *Proc. 15th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2009, pp. 637–646.

[27] J. J.-C. Ying, W.-C. Lee, and V. S. Tseng, "Mining geographic-temporal-semantic patterns in trajectories for location prediction," *ACM Trans. Intell. Syst. Technol*, vol. 5, no. 1, p. 2, 2013.

[28] A. Noulas, S. Scellato, C. Mascolo, and M. Pontil, "An empirical study of geographic user activity patterns in foursquare," in *Proc. ICWSM*, vol. 11. 2011, pp. 570–573.

[29] E. Cho, S. A. Myers, and J. Leskovec, "Friendship and mobility: User movement in location-based social networks," in *Proc. 17th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2011, pp. 1082–1090.

[30] Y. Jia, Y. Wang, X. Jin, and X. Cheng, "Location prediction: A temporal-spatial Bayesian model," *ACM Trans. Intell. Syst. Technol.*, vol. 7, no. 3, p. 31, 2016.

[31] Y. Ding and X. Li, "Time weight collaborative filtering," in *Proc. 14th ACM Int. Conf. Inf. Knowl. Manage.*, 2005, pp. 485–492.

[32] R. Salakhutdinov and A. Mnih, "Probabilistic matrix factorization," in *Proc. NIPS*, vol. 20. 2011, pp. 1–8.

[33] L. Xiong *et al.*, "Temporal collaborative filtering with Bayesian probabilistic tensor factorization," in *Proc. SIAM Int. Conf. Data Mining*, Columbus, OH, USA, Apr./May 2010, pp. 211–222.

[34] Y. Wang *et al.*, "Regularity and conformity: Location prediction using heterogeneous mobility data," in *Proc. 21th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2015, pp. 1275–1284.

[35] A. Y. Xue, R. Zhang, Y. Zheng, X. Xie, J. Huang, and Z. Xu, "Destination prediction by sub-trajectory synthesis and privacy protection against such prediction," in *Proc. IEEE 29th Int. Conf. Data Eng. (ICDE)*, Apr. 2013, pp. 254–265.

[36] A. T. Palma, V. Bogorny, B. Kuijpers, and L. O. Alvares, "A clustering-based approach for discovering interesting places in trajectories," in *Proc. ACM Symp. Appl. Comput.*, 2008, pp. 863–868.

[37] Y. Zheng, L. Zhang, X. Xie, and W.-Y. Ma, "Mining interesting locations and travel sequences from GPS trajectories," in *Proc. 18th Int. Conf. World Wide Web*, 2009, pp. 791–800.

[38] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Proc. KDD*, vol. 96. 1996, pp. 226–231.

[39] U. von Luxburg, "A tutorial on spectral clustering," *Statist. Comput.*, vol. 17, no. 4, pp. 395–416, 2007.

[40] J. Pei *et al.*, "Mining sequential patterns by pattern-growth: The PrefixSpan approach," *IEEE Trans. Knowl. Data Eng.*, vol. 16, no. 11, pp. 1424–1440, Nov. 2004.
[41] Y. Zheng, Q. Li, Y. Chen, X. Xie, and W.-Y. Ma, "Understanding mobility based on gps data," in *Proc. 10th Int. Conf. Ubiquitous Comput.*, 2008, pp. 312–321.
[42] Y. Zheng, X. Xie, and W.-Y. Ma, "GeoLife: A collaborative social networking service among user, location and trajectory," *IEEE Data Eng. Bull.*, vol. 33, no. 2, pp. 32–39, Jun. 2010.

RUIZHI WU received the B.S. degree from Hangzhou Dianzi University in 2012. He is currently pursuing the Ph.D. degree with the School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu, China. His research interests include spatial-temporal data mining, data mining, and machine learning.

GUANGCHUN LUO (M'06) received the B.S., M.S., and Ph.D. degrees in computer science from the University of Electronic Science and Technology of China, Chengdu, China, in 1995, 1999, and 2004, respectively. He is currently a Professor of computer science with the University of Electronic Science and Technology of China. His research interests include computer networking, cloud computing, and big data.

QINLI YANG received the Ph.D. degree from The University of Edinburgh, U.K. She is currently with the University of Electronic Science and Technology of China. She has published many papers in prestigious journals like *Water Research*, *Environmental Modelling and Software*, the *Journal of Environmental Management*, and several papers in the field of data mining. Her current research interests include data mining driven water resources research.

JUNMING SHAO (M'17) received the Ph.D. degree *(summa cum laude)* from the University of Munich, Germany, in 2011. He became the Alexander von Humboldt Fellow in 2012. He not only published papers for top-level data mining conferences like KDD, ICDM, and SDM (two of those papers have won the Best Paper Award), but also published data mining-related interdisciplinary work in leading journals, including *Brain*, the *Neurobiology of Aging*, and *Water Research*. His research interests include data mining and neuroimaging.