电子科技大学
University of Electronic Science and Technology of China

Regularity and Conformity: Location Prediction Using Heterogeneous Mobility Data

Ruizhi Wu

Data Mining Lab, Big Data Research Center, UESTC
Email：nxznwurz@163.com

Contents

Introduction of Prediction in Trajectory Data

Mobility prediction enables appealing proactive experiences for location-aware services and offers essential intelligence to business and governments.

methods	target			feature					
	CI	GPS	Wifi	SMP	TC	IT	SR	CF	HT
PSMM [6]	√			√		√	√		
W^4 [40]	√			√	√	√			
M5Tree[25]	√					√	√		
CEPR [19]	√			√		√		√	
SHM [12]	√					√	√		
gSCorr [11]	√					√	√		
DBN [26]	√					√	√		
NextPlace [27]		√	√			√			
WhereNext [23]		√				√			
Markov [1]		√							
RCH(Our Model)	√			√		√		√	√

CI: check-in, SMP: spatial mobility pattern, TC: text content
IT: individual temporal patterns, SR: social relationship
CF: collaborative filtering, HT: heterogeneous mobility datasets

RCH model motivation and overview

Motivation:
Recent studies suggest that human mobility is highly regular and predictable. Additionally, social conformity theory indicates that people's movements are influenced by others.

Goal:
Given visited venues of a group of users, this paper goal is to predict their future locations at a certain time.

A user ui's visit to a venue vj can be driven by either regularity and conformity.

$$R_{ij}(t) = R_{ij}^{(r)}(t) + R_{ij}^{(c)}(t),$$

Let R(t) ∈ RM×N be the preference matrix of U to V at time t, Rij(t) indicates ui's preference to vj at t.

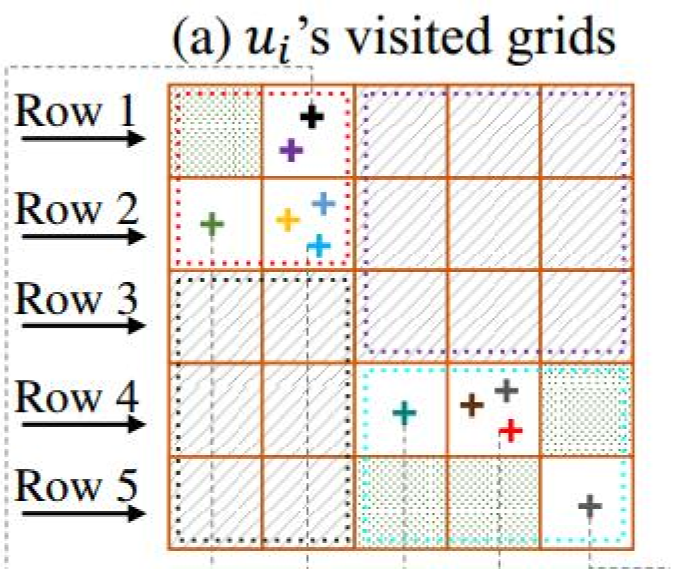
They categorize days into two types, workdays and holidays, and let T = {t1, t2, ..., tT } represent the T time slots in the two classes of days

Regularity Term

Assuming the **Markov property** of users' transitions between grids.

A stochastic process has the Markov property if the conditional probability distribution of future states of the process (conditional on both past and present states) depends only upon the present state, not on the sequence of events that preceded it

Let C = {d1, d2, ..., dI} be the C geographical grid cells (e.g., 100m × 100m) discretizing the whole geospatial space of a city. Each venue vj belonging to a certain grid dkj .



Regularity Term

Consider the probability that ui visits vj in terms of regularity, denoted as Pr(vj|ui).

We assume that vj belongs to a grid dkj , and ui travels from a grid dk to vj (note that it is possible that dk = dkj).

$$\Pr(v_j|u_i) \propto \sum_{k=1}^{I} \Pr(d_k|u_i)\Pr(v_j|d_k)$$
$$= \sum_{k=1}^{I} \Pr(d_k|u_i)\Pr(d_{k_j}|d_k)\Pr(v_j|d_{k_j})$$

Regularity Term

$$\Pr(v_j|u_i) \propto \sum_{k=1}^{I} \Pr(d_k|u_i)\Pr(v_j|d_k)$$

$$= \sum_{k=1}^{I} \Pr(d_k|u_i)\Pr(d_{k_j}|d_k)\Pr(v_j|d_{k_j})$$

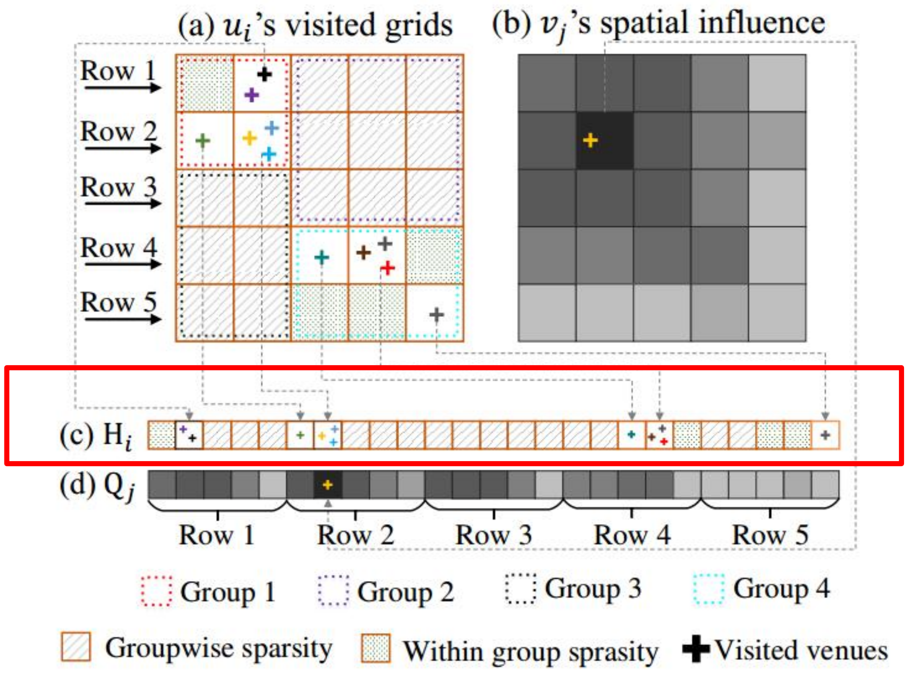


Figure 1: Learning the regularity term

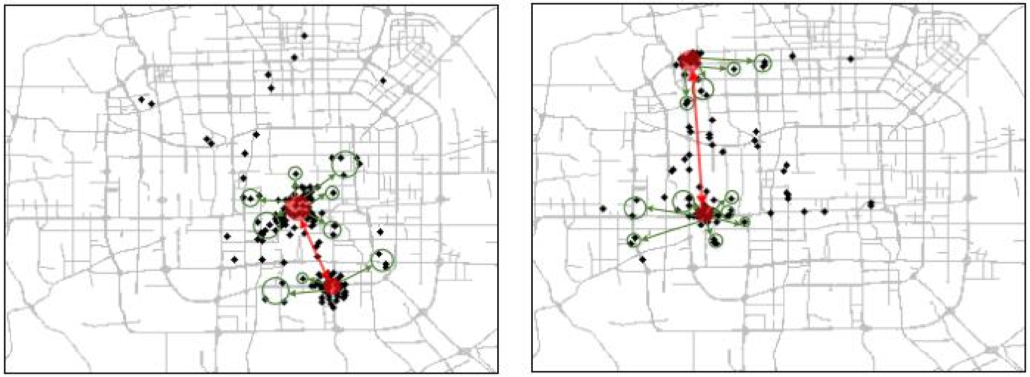
Hik is the visiting frequency of ui to grid dk, approximating Pr(dk|ui), and we term H the *hub matrix* of U

The second factor Pr(dkj |dk) is the transition probability from dk to dkj, which is learned based on a gravity model using hetereneous mobility datasets. The third factor Pr(vj|dkj) can be estimated using the visiting frequency of vj in grid dkj.

Regularity Term

We combine the second and third factor together as Qjk, which represents the spatial influence of vj to grid dk. Qjk indicates the degree of influence that attracts users from grid dk to vj, and we refer to Q as the *spatial influence matrix.*

$$\Pr(v_j|u_i) = H_i Q_j^T$$



(a) check-ins of user A (b) check-ins of user B

The I grid cells are clustered into G groups G = {g1, g2, . . . , gG}.

$$R_{i,j}^{(r)} = \sum_g \mathbf{H}_i^{(g)} \left(\mathbf{Q}_j^{(g)}\right)^\top ,$$

Conformity Term

Motivation:
The social conformity theory suggests that users who have similar backgrounds, interests, and social statuses often behave similar to each other, which is also the psychological root of collaborative filtering.

R(c) into two low dimensional latent matrices U and V, where Ui and Vj are latent factors of user i and POI j, both with dimension K.

$$R_{i,j}^{(c)} = U_i V_j^T$$

We add the time changing part of user latent factor Ui(t) to describe changeable preferences, and Ui is the stationary part for unchangeable interests for venues.

$$R_{i,j}^{(c)}(t) = (\mathbf{U}_i + \mathbf{U}_i(t))\mathbf{V}_j^\top .$$

Gravity model

Gravity models are used in various social sciences to predict and describe certain behaviors that mimic gravitational interaction as described in Isaac Newton's law of gravity. Generally, the social science models contain some elements of mass and distance, which lends them to the metaphor of physical gravity.

$$T_{i,j}^{*} = c\frac{(\mathcal{O}_i^{*})^a(\mathcal{D}_j^{*})^b}{\exp(r \cdot dis_{i,j})}, \quad * \in \{B, A, C\},$$

Two task:

1) estimating the *spatial influence matrix* Q ∈ RN×I, where the element in the jth row and kth column Qjk is the spatial influence of venue vj to grid dk; and

2) learning the group structure of the hub matrix H.

Gravity model

Let Oi ∗ be the number of individuals leaving grid di, for i = 1, 2, ..., I, and Dj ∗ be the number of people going towards grid dj, for j = 1, 2, . . . , I, where ∗ ∈ P indicates a certain type of mobility. In this work, we consider three types of mobility: B(bus), C(check-in), and A(taxi), i.e., P = {B, A, C}.

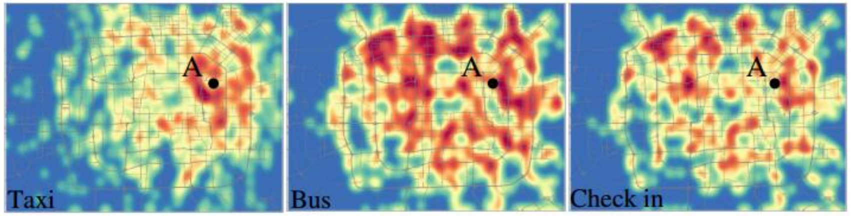
$$T^*_{i,j} = c\frac{(\mathcal{O}^*_i)^a(\mathcal{D}^*_j)^b}{\exp(r \cdot dis_{i,j})}, \quad * \in \{B, A, C\},$$

Our goal is to estimate the coefficients a, b, r by fitting this model using observed mobility data. We achieve it using the multivariate regression method.

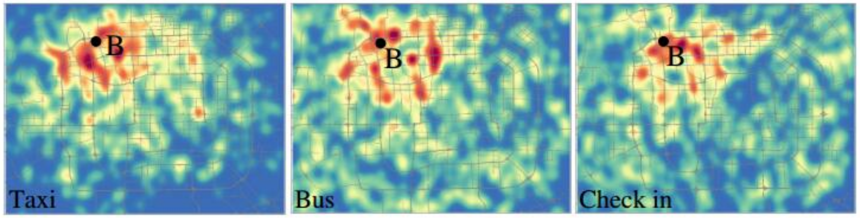
$$\ln T^*_{i,j} = a \ln \mathcal{O}^*_i + b \ln \mathcal{D}^*_j - r \cdot dis_{i,j} + \ln c.$$

Combined with the estimated visiting frequency of vj in dkj, denoted as ρj, we have the final spatial influence.

$$Q^*_{jk} = \rho_j \hat{T}^*_{k,k_j}.$$



(a) spatial influence to grid A (a bar district) from other grids



(b) spatial influence to grid B (an IT district) from other grids

Next, we learn the group structure of users' major hubs. Given the estimated transition matrix T ˆ ∗ for ∗ ∈ P, we employ the Dirichlet Multinomial Regression (DMR) to learn the group structure (also known as the *functional zones*) of a city. Specifically, for grid di, i = 1, . . . , I, we extract all out-going transitions T ˆ ∗ i· and in-coming transitions T ˆ ·i

Final objective function and Optimization

We learn the predictor Ri,j(t) using a supervised learning approach by constructing an optimization problem. As explained in, if we only consider the conformity term, our problem can be solved using a time-aware matrix factorization mode.

$$\min_{\mathbf{U},\mathbf{U}(t),\mathbf{V}} \sum_{t\in\mathcal{T}} \|\mathbf{R}(t) - (\mathbf{U}+\mathbf{U}(t))\mathbf{V}^\top\|_F^2 + \gamma(\|\mathbf{U}\|_F^2 + \|\mathbf{V}\|_F^2) + \beta \sum_{t\in\mathcal{T}} \|\mathbf{U}(t)\|_F^2, \quad (13)$$

Final objective function:

$$\begin{aligned}
&\Theta(\mathbf{U},\mathbf{U}(t),\mathbf{V},\mathbf{H}(t),\theta^B,\theta^A)\\
&= \sum_{t\in\mathcal{T}} \|\mathbf{R}(t) - \sum_{g\in\mathcal{G}} \mathbf{H}^{(g)}(t)\left(\sum_{*\in\mathcal{P}} \theta^* \mathbf{Q}^{*(g)}\right)^\top - (\mathbf{U}+\mathbf{U}(t))\,\mathbf{V}^\top\|_F^2\\
&+ \sum_{t\in\mathcal{T}} ((1-\alpha)\sigma \sum_{j=1}^{M}\sum_{g\in\mathcal{G}} \|\mathbf{H}_j^{(g)}(t)\|_2 + \alpha\sigma \sum_{j=1}^{M} \|\mathbf{H}_j(t)\|_1)\\
&+ \gamma(\|\mathbf{U}\|_F^2 + \|\mathbf{V}\|_F^2) + \beta \sum_{t\in\mathcal{T}} \|\mathbf{U}(t)\|_F^2, \qquad (14)
\end{aligned}$$

Final objective function and Optimization

Algorithm 1: Optimization of RCH Model

Input: $\alpha, \beta, \gamma, \sigma, \tilde{\mathbf{R}}(t)(t \in \mathcal{T}), \mathbf{Q}^{*}(* \in \mathcal{P})$
Output: $\mathbf{U}$, $\mathbf{U}(t)$, $\mathbf{V}$, $\mathbf{H}(t)$, θ^{B} and θ^{A} minimizing Θ in (32)

$\mathbf{U}, \mathbf{U}(t), \mathbf{V}, \mathbf{H}(t), \theta^{B}, \theta^{A} \Leftarrow \mathbf{U}_0, \mathbf{U}_0(t), \mathbf{V}_0, \mathbf{H}_0(t), \theta^{B}_0, \theta^{A}_0$;
repeat
    update $\mathbf{U}$ with (17);
    update $\mathbf{V}$ with (21);
    update θ^{B} and θ^{A} with (27);
    for $\tau = 1, 2, ..., T$ **do**
        update $\mathbf{U}(t)$ with (20);
        for $j = 1, 2, ..., M$ **do**
            for $g = 1, 2, 3, ..., G$ **do**
                if $\|\mathscr{F}(\tilde{\boldsymbol{R}}_j^{(-g)}(\tau)\tilde{\boldsymbol{Q}}^{(g)}, \alpha\sigma)\|_2 \leq (1-\alpha)\,\sigma$ **then**
                    $\mathbf{H}_j^{(g)}(\tau) = \mathbf{0}$;
                else
                    $\mathbf{H}_j^{(g)}(\tau) \Leftarrow \mathscr{H}\left(\mathbf{H}_j^{(g)}(\tau)\right)$
until *convergence*;
return $\mathbf{U}$, $\mathbf{U}(t)$, $\mathbf{V}$, $\mathbf{H}(t)$, θ^{B} and θ^{A}

Experiment

We divide the check-in data into two parts in a chronological order: 70% for training portion and 30% for testing portion.

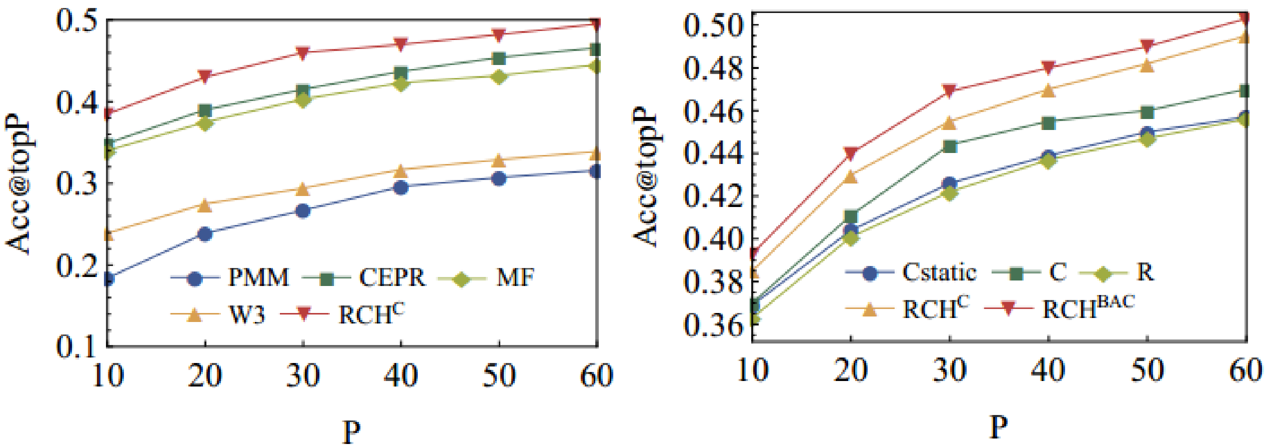
Check-in Dataset   Bus Dataset   Taxi Dataset

We use two metrics to measure the performance of location prediction: prediction accuracy (Acc@topP) and the average percentile rank (APR) of the actually visited venues.

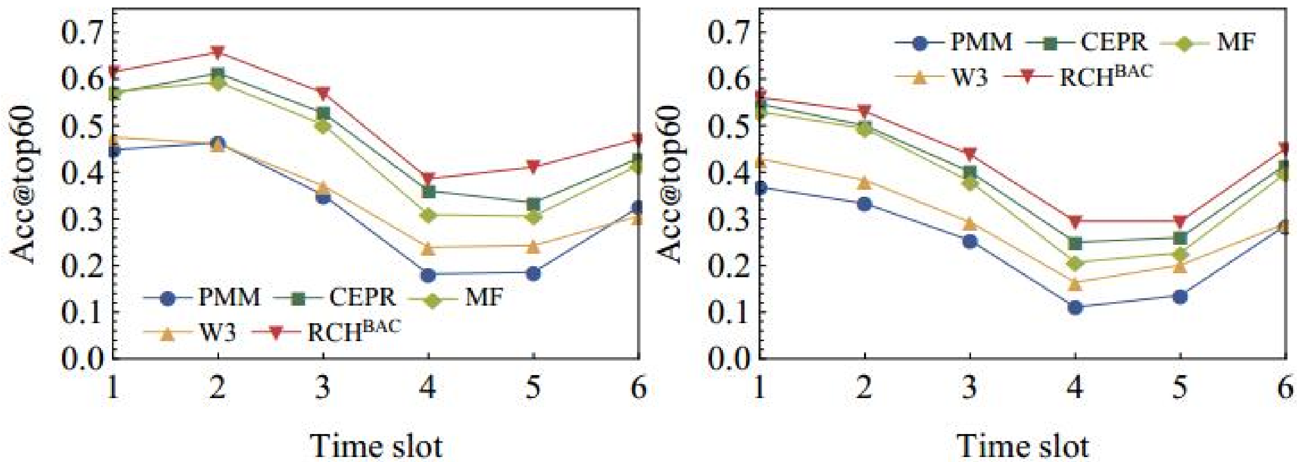
The percentile rank of prediction for venue vj . where rank(vj) is the position of venue vj in the predicted list and N is the number of venues. It is clear that PR is 1 if the true venue is ranked as the first. Average Percentile Rank (APR) is the average of PR over the testing set.

$$PR = \frac{N - rank(v_j) + 1}{N},$$

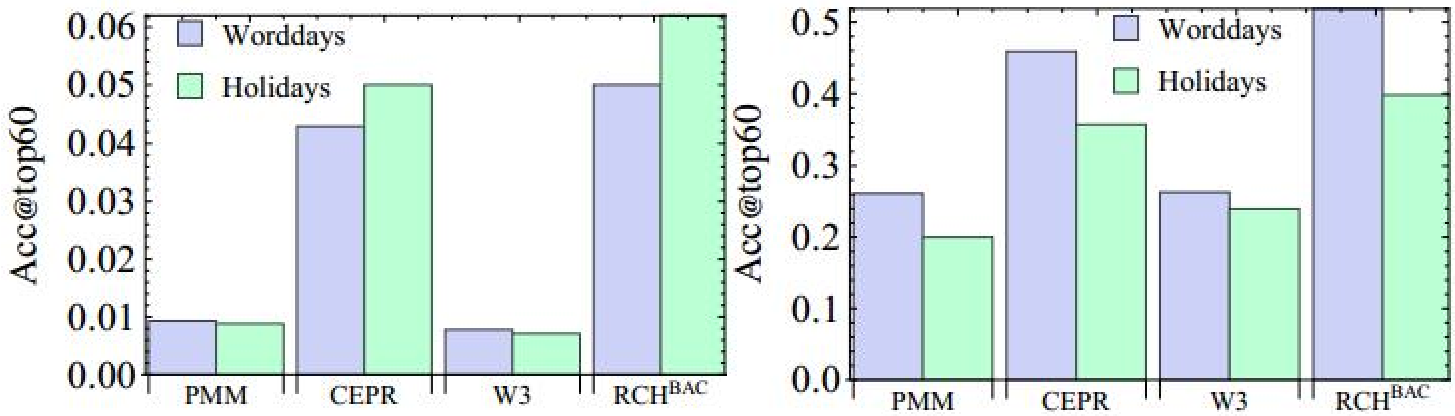
Experiment



(a) different models (b) different variations of RCH



(a) workdays (b) holidays



(a) unvisited venues (b) visited venues

Models	Workdays						Holidays					
t	12-4am	4-8am	8am-12pm	12-4pm	4-8pm	8pm-12am	12-4am	4-8am	8am-12pm	12-4pm	4-8pm	8pm-12am
C_{static}	0.884	0.899	0.865	0.799	0.801	0.832	0.872	0.848	0.807	0.753	0.761	0.840
C	0.885	0.908	0.869	0.820	0.825	0.848	0.868	0.854	0.818	0.781	0.787	0.844
R	0.887	0.884	0.873	0.826	0.831	0.860	0.859	0.843	0.814	0.768	0.790	0.837
RCH^{C}	0.896	0.911	0.880	**0.835**	0.838	0.870	0.881	0.859	0.823	0.781	0.793	0.849
RCH^{BAC}	**0.899**	**0.912**	**0.883**	**0.835**	**0.840**	**0.871**	**0.890**	**0.863**	**0.829**	**0.786**	**0.795**	**0.850**

Table 3: APR of our models in different time slots.

Interior and Exterior are indispensable.