

CityCoupling: Bridging Intercity Human Mobility

Zipei Fan¹, Xuan Song¹, Ryosuke Shibasaki¹, Tao Li², Hodaka Kaneda³

¹Center for Spatial Information Science, the University of Tokyo, Japan

²School of Computer Science, Florida International University, Miami, U.S.A.

³Zenrin DataCom Co'Ltd, Tokyo, Japan

fanzipei@iis.u-tokyo.ac.jp, {songxuan, shiba}@csis.u-tokyo.ac.jp,

taoli@cs.fiu.edu, h_kaneda@zenrin-datacom.net

ABSTRACT

There are two broad categories of citywide human mobility, routine, composed of daily or periodic travel, and rare, which occurs during events such as the Olympic Games or natural disasters. State-of-the-art studies have shown that routine mobility patterns can be modeled stochastically, while rare human mobility modeling, essential to a variety of urban computing scenarios, such as emergency management and traffic regulation, is a much more challenging and understudied problem. Instead of training a rare-event-specific human mobility model, which suffers from the particularity of the rare events, in this paper we provide a new insight into rare events and propose a novel algorithm, CityCoupling, which establishes an intercity spatial mapping that uses human mobility in one city as input and reproduces human mobility in another city. More intuitively, we attempt to answer the question "What if this rare event happened in another city?". To find the optimal intercity spatial mapping, we utilize an expectation-maximization algorithm to estimate a probabilistic geographical correspondence matrix by regarding intercity trajectory matching as latent variables. Thereafter, a Gibbs sampling-based multiple hidden Markov model generates simulated trajectories. We apply our approach to a large mobile phone GPS dataset in Japan and determine the spatial mapping between Tokyo and Osaka to transfer the human mobility at the Great Eastern Japan Earthquake in Tokyo, which was heavily affected, to simulate what might have occurred if Osaka had been struck by the earthquake. We conduct the evaluation by assuming that New Year's Countdown is a rare event that occurs simultaneously in both Tokyo and Osaka, and thus we quantitatively compare our simulation with the ground truth in Osaka.

ACM Classification Keywords

H.4 Information Systems Applications: Miscellaneous; H.2.8 Database Management: Database Applications - data mining, Spatial databases and GIS

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

UbiComp '16, September 12 - 16, 2016, Heidelberg, Germany

Copyright is held by the owner/author(s). Publication rights licensed to ACM.

ACM 978-1-4503-4461-6/16/09 \$15.00

DOI: <http://dx.doi.org/10.1145/2971648.2971737>

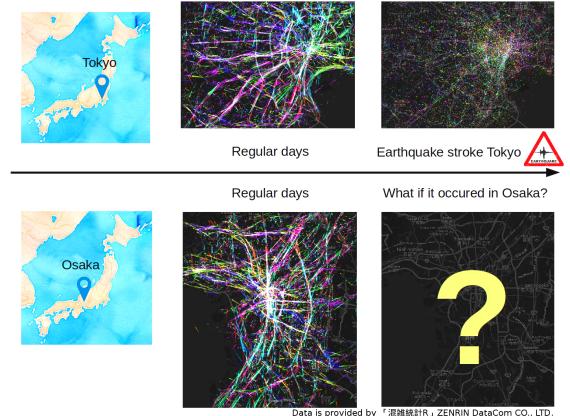


Figure 1. ¹Tokyo was struck by the Great Eastern Japan Earthquake, while Osaka, far from the epicenter, was only slightly affected. To learn lessons and gain experience from Tokyo, we may ask "What if Osaka was also struck by the earthquake?"

Author Keywords

human mobility; urban computing; emergency management; big data

INTRODUCTION

When we turn on the TV, we can find reports of major events [1], both joyful (e.g. the Olympics Games or the New Year's celebrations), and disastrous [2] (e.g. earthquakes or tsunamis) from around the world. Being affected by the happiness or agony of people experiencing these events, we may wonder: "What if this happened in my city?" The same question also intrigues city planners, emergency managers, and traffic safety bureaus. For these groups, gaining experience and learning the lessons from other cities hosting a major event or struck by a tremendous natural disaster can serve as critical examples for reflecting on the design of their own cities.

Owing to the popularization of mobile phones and advances in localization technology, long-term continuous recording of large-scale human mobility can be measured at negligible cost. Such human mobility data give us an objective, accurate, and informative recording of movement for a long period of

¹We used "Konzatsu-Tokei (R)" from ZENRIN DataCom CO., LTD. "Konzatsu-Tokei (R)" is generated based on user agreement of GPS mobile phones. The statistics data is anonymized and aggregated so that individual information can never be retrieved. The anonymization and aggregation are made by NTT Docomo based on the request of Zenrin DataCom co'ltd. In this study, the proposed methodology is applied to raw GPS data by NTT DOCOMO INC.

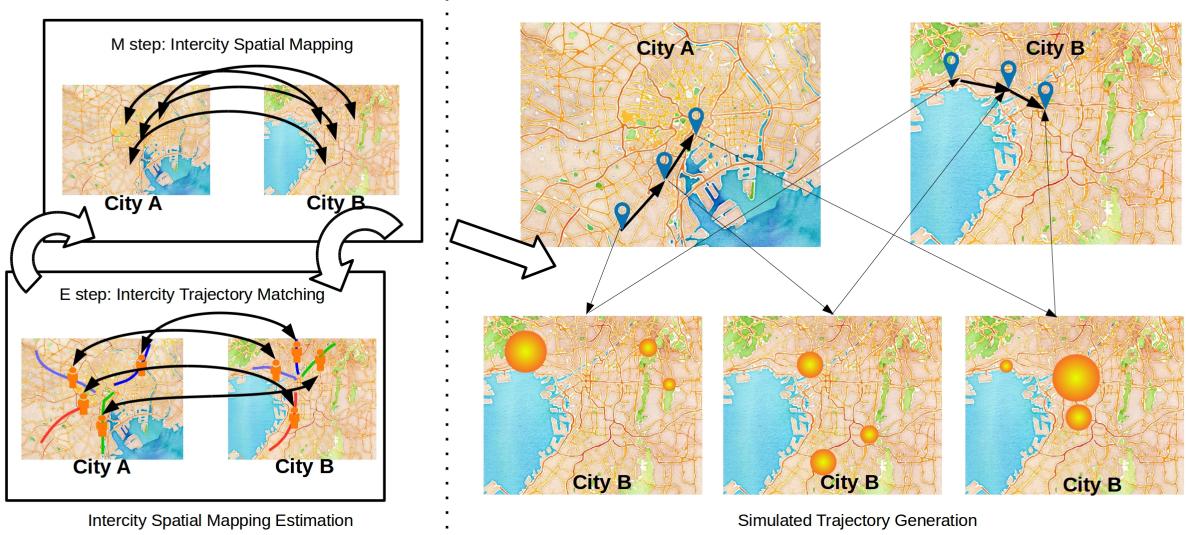


Figure 2. Overview of our CityCoupling algorithm. Our algorithm can be roughly divided into two phases: an EM algorithm to estimate intercity spatial mapping, depicted in the left panel, and an HMM algorithm to generate simulated trajectories, depicted in the right panel.

time and therefore enable a variety of emerging ubiquitous computing applications, especially on modeling and generalizing rare events, which can hardly be sensed and analyzed in a traditional way.

Routine human mobility has predictable characteristics, but each rare event is unique. Considering two examples of the greatest earthquakes in Japan (the Great Hanshin Earthquake in 1995, which occurred in the early morning near Osaka and led to the collapse of a large number of buildings, and the Great Tohoku Earthquake in 2011, which occurred in the afternoon in the eastern part of Japan and was followed by a severe tsunami), we find that human mobility during these two earthquakes differ considerably between them because of the disparity in characteristics of the two earthquakes. Thus, we must pay special attention to determining whether we have sufficient information and proper assumptions to make generalizations related to each of these factors.

Aside from the particularity of each rare event, the particularity of city layouts also prevents us from directly transferring knowledge from one city to another. For example, when we learn lessons from one city in which a transportation hub has been struck by a tremendous earthquake, we must know the details of the corresponding transportation hub in the other city. In previous research, this is done mainly by empirical knowledge with some supportive evidence (e.g. the largest railway station in one city corresponds to the largest railway station in another city); however it is unrealistic to determine city layout alignment manually, because of complex city topologies and multi-modality in intercity layout alignments (that is, one place in a city may correspond to many candidate places in another).

Bearing these difficulties in mind, we propose a novel algorithm, CityCoupling, to establish probabilistic intercity spatial mapping that aligns the layouts of two cities, preserving the multi-modality of intercity layout alignment. In addition, to

avoid making unreliable generalization from rare events, we apply a spatial transformation to rare human mobility, while leaving all the intrinsic factors of human mobility unchanged.

Note that this intercity spatial mapping can be assumed to be invariant with respect to both routine and rare human mobility. Even if there is a dramatic change in regional functionalities when the rare event occurs (e.g. a transportation hub is struck by a powerful earthquake), the intercity spatial mapping will not be subject to change. Instead of altering the spatial mapping, such regional functionality changes are transferred to the other city. Due to this invariance property, we can exploit the entire dataset to make inferences about intercity spatial mapping. This mapping is applicable whenever new human mobility data are collected.

Our CityCoupling algorithm can be divided into two phases: 1) an expectation maximization (EM) framework for intercity spatial mapping estimation and 2) a hidden Markov model (HMM) that incorporates knowledge transferred from the source city, geographical continuity, and population density prior to generating simulated trajectories.

Our contribution can be summarized as follows:

- As far as we know, this work is the first to simulate human mobility during generic rare events without requiring any prior knowledge of the rare event.
- We propose the novel CityCoupling algorithm, which estimates intercity spatial mapping in an EM framework and generates the simulated trajectories using Gibbs sampling based on multiple hidden Markov model.
- We apply our algorithm to real-world GPS log dataset, and successfully transfer human mobility during the Great Eastern Earthquake from Tokyo to Osaka, and validate our algorithm using human mobility during New Year's celebrations in 2011.

OVERVIEW

In this section, we provide an overview of our CityCoupling algorithm. As shown in Fig. 2, our algorithm can be divided into two phases. In the first phase, we estimate intercity spatial mapping with an EM algorithm by designating intercity trajectory matching as latent variables and the intercity spatial mapping as the parameters. In the estimation step, we compute the similarity between the trajectories in the two cities by applying the **intercity spatial mapping** and finding the optimal **intercity trajectory matching**; in the maximization step, we update our estimation of the **intercity spatial mapping** based on the co-occurrence of trajectories pairs in the **intercity trajectory matching**.

In the second phase, we use the intercity spatial mapping estimation from the previous phase, along with an HMM, to enforce a geographical transition prior. We utilize the Viterbi algorithm to find the optimal simulated trajectories. As shown in the right panel, each GPS record in a trajectory in city A implies a geographical distribution in city B . Thereafter, this distribution, along with a population density prior that we will introduce later, constructs the HMM emission probability, and a gamma distribution, with the expectation that the distance traveled between two continuous records in the source trajectory defines a spatial transition prior probability distribution. In this framework, we reduce the multi-modality of intercity spatial mapping and make the simulated trajectory in city B more similar to the source trajectory in city A .

PRELIMINARY

We first define terms and essential concepts that are frequently used in the rest of this paper:

Definition 1 (GPS log data): The original dataset can be described by a set of 4-tuples:

$$X = \{(u, t, lat, lon)\}$$

where u , t , lat , and lon are the unique mobile phone user ID, time stamp, latitude and longitude of the GPS record respectively.

To simplify our model, in most cases we use discretized GPS log data:

$$\hat{X} = \{(u, \hat{t}, \hat{l})\}$$

where \hat{t} and \hat{l} are the time slot and grid map indices, respectively. For conciseness, in the following parts of this paper, we use discretized time and location information, denoted by X , by default.

Definition 2 (Intercity spatial mapping): We define an intercity spatial mapping from city A to city B as a matrix $P_{A \rightarrow B}$ of size (n_{l_B}, n_{l_A}) , where n_{l_*} is the number of grid cells in city $*$. In addition, to impose equality in the knowledge transfer of each trajectory in city A , we constrain the mapping matrix to be left stochastic (summing to 1 for each column).

Definition 3 (Intercity trajectory matching): Given two sets of unique IDs U_A and U_B of the same size ($|U_A| = |U_B|$) from city A and city B respectively, an intercity trajectory matching is a permutation matrix Z_{U_A, U_B} (a binary matrix that has exactly

one 1 in each row and each column and 0s elsewhere) of the same dimensions as $|U_A|$ (or $|U_B|$).

Definition 4 (Simulated GPS data): Given GPS log data X_A from city A , we generate simulated GPS data $X_{B|A, P_{A \rightarrow B}}$ in city B via intercity trajectory matching $P_{A \rightarrow B}$. The simulated GPS data have the same data format as the GPS log data, and can be visualized and compared with the GPS log data.

INTERCITY SPATIAL MAPPING ESTIMATION

In this section, we present our algorithm for finding optimal intercity spatial mappings. Given two long-term discrete GPS log datasets X_A from city A and X_B from city B with the same time span, we approach the problem from an EM framework, treating intercity trajectory matching as latent variables and intercity spatial mapping as maximum likelihood estimates of parameters.

Recall our definition of intercity trajectory matching; we require the number of trajectories to be equal. To meet this requirement, we formulate expectation in a stochastic manner:

$$\mathcal{Q}(P_{A \rightarrow B}, P_{A \rightarrow B}^{old}) = \sum_{(U_A, U_B) \in S_{A,B}(N)} E_f [\log(L(X_{U_B}, Z_{U_A, U_B} | X_{U_A}, P_{A \rightarrow B}))] \quad (1)$$

where $S_{A,B}(N)$ denotes the set of some random samples of (U_A, U_B) satisfying $|U_A| = N$ and $|U_B| = N$. X_{U_*} is the trajectory subset cropped from discrete GPS log data X_* by

$$X_{U_*} = \{(u, t, l) \in X_* | u \in U_*\}$$

and $E_f(\cdot)$ is the expectation of the expression \cdot with respect to the probability distribution $f = p(Z_{U_A, U_B} | X_{U_A}, X_{U_B}, P_{A \rightarrow B}^{old})$, which will be further explained in the next subsection. In the Intercity Spatial Mapping subsection, we provide a more detailed definition of the likelihood function L , which evaluates our estimation of $P_{A \rightarrow B}$.

Intercity Trajectory Matching

In this subsection, we discuss a method for finding the optimal intercity trajectory matching Z_{U_A, U_B} given GPS data X_{U_A} and X_{U_B} and previous intercity spatial mapping estimate $P_{A \rightarrow B}^{old}$.

In the first step, calculate the intercity trajectory similarity matrix M_{U_A, U_B} , where u_A -th row and the u_B -th column entry is the trajectory similarity between user ID u_A and u_B is defined as the probability of drawing the trajectory of u_B given the trajectory of u_A and the intercity spatial mapping $P_{A \rightarrow B}^{old}$:

$$m_{u_A, u_B} = \log(p(X_{u_A} | X_{u_B}, P_{A \rightarrow B}^{old})) \propto \text{sum}(Y_{u_B} \circledast \log(P_{A \rightarrow B}^{old} Y_{u_A})) \quad (2)$$

where \circledast is the element-wise multiplication operator between two matrices and "sum" denotes the function that computes the sum of all entries in a matrix. Y_{u_*} denotes the trajectory matrix of subject u_* ; each entry of this matrix is defined as:

$$Y_{u_*}(l, t) = \begin{cases} 1, & \text{if } (u_*, t, l) \in X_{u_*} \\ \varepsilon, & \text{otherwise} \end{cases} \quad (3)$$

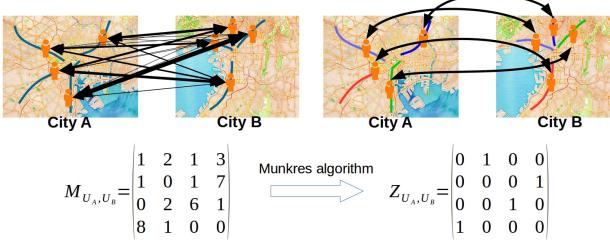


Figure 3. A simple example that illustrates intercity trajectory matching. The values in M_{U_A, U_B} indicate the similarity between trajectories, while the boolean value in Z_{U_A, U_B} encodes whether two trajectories are matched in our matching result.

where ε is a small positive constant. We normalize Y_{u_*} with respect to each column so that the t -th column can be regarded as the spatial distribution of u_* at time t .

Intuitively, $P_{A \rightarrow B}^{old} Y_{u_A}$ estimates a spatial distribution of u_A under the intercity spatial mapping, while an element-wise multiplication by Y_{u_B} picks out the posterior distributions in Equation 2 and sums them.

Obtaining the intercity trajectory similarity matrix M_{U_A, U_B} , we define the optimal intercity trajectory matching Z_{U_A, U_B} as:

$$\begin{aligned} Z_{U_A, U_B} &= \underset{Z}{\operatorname{argmax}} M_{U_A, U_B} \otimes Z \\ \text{s.t. } Z &\text{ is a permutation matrix} \end{aligned} \quad (4)$$

Therefore, as shown in Figure 3, we formulate intercity trajectory matching as a bipartite graph matching. In this study, we apply the widely used Munkres assignment algorithm [17] (we use the implementation from [7]) to find the optimal intercity trajectory matching.

Intercity Spatial Mapping

We initially guess at $P_{A \rightarrow B}$ by taking the population density variation over the entire time span of the training dataset into consideration. We obtain this guess by minimizing the error of reconstructing the citywide population density of city B from city A of intercity spatial mapping P :

$$\begin{aligned} P_{A \rightarrow B}^0 &= \underset{P}{\operatorname{argmin}} \|D_{X_B} - PD_{X_A}\|_2 + \lambda \|P\| \\ \text{s.t. } P &> 0, \text{ and } P^T \mathbf{1}_{n_{l_B}} = \mathbf{1}_{n_T} \end{aligned} \quad (5)$$

where D_{X_*} is the city $*$ population density variation matrix (of size $n_{l_*} \times T$) estimated from X_* . n_T is the number of time slots in the entire dataset. Each column in D_{X_*} presents the population density distribution over locations of city $*$ at time t ; therefore, D_{X_*} is a left stochastic matrix. $\mathbf{1}_n$ is the column vector of all 1s with length n .

Next, we complete the iteration of the EM algorithm. Recall that in the previous subsection, we found the optimal intercity trajectory matching Z_{U_A, U_B} in the expectation step. In the maximization step, we update our intercity spatial mapping estimation $P_{A \rightarrow B}$ by

$$P_{A \rightarrow B} = \left(P_{A \rightarrow B}^{old} + \gamma \Delta P \right) C \quad (6)$$

Input: GPS log data X_A, X_B
Output: Intercity Spatial Mapping $P_{A \rightarrow B}$

/* Initialization */

Initialize $P_{A \rightarrow B}$ using Eq. 5;

/* Expectation-maximization iteration */

for $i = 1$ to $ITERATION_{mapping}$ do

Randomly sample U_A, U_B ($|U_A| = |U_B| = N$);

Crop the data X_{U_A}, X_{U_B} ;

/* Expectation step */

Calculate M_{U_A, U_B} from Eq. 2;

Find Z_{U_A, U_B} using Munkres algorithm from Eq. 4;

/* Maximization step */

Calculate W from Eq. 8;

Calculate ΔP from Eq. 7;

Update $P_{A \rightarrow B}$ by Eq. 6;

end

return $P_{A \rightarrow B}$;

Algorithm 1: Intercity Spatial Mapping Estimation

where ΔP is a non-negative increment for the estimation of $P_{A \rightarrow B}$ with step size of γ and C is a diagonal matrix that performs a column normalization to maintain the constraints of the left stochastic matrix.

From a Bayesian perspective, each column of $P_{A \rightarrow B}$ can be regarded as a multinomial distribution over locations in city B , and for each maximization step, we want to estimate the posterior probability distribution given the new intercity trajectory matching observation. We draw an analogy between updating our estimation of each column of $P_{A \rightarrow B}$ and a Pólya Urn model [13]: for each pair of matched trajectories u_A and u_B , when u_A is observed traveling to l_A while u_B travels to l_B , we can treat it as a new ball labeled (l_A, l_B) drawn from the urn; therefore we increment our belief that l_A should have a higher probability of being mapped to l_B and hence increment the entry at the l_A -th column, l_B -th row of $P_{A \rightarrow B}$. Following this intuition, we define our increment for estimating $P_{A \rightarrow B}$ as

$$\Delta P = W \otimes \sum_{Z[u_A, u_B] = 1} Y_{u_B} Y_{u_A}^T \quad (7)$$

where W , as is defined below, is real matrix of values ranging $[0, 1]$ imposing a population density variation prior for the entire time span.

$$W = \frac{1}{2} (1 + \cos \langle D_{X_B} - P_{A \rightarrow B} D_{X_A}, D_{X_A} \rangle) \quad (8)$$

where $\cos \langle G, H \rangle$ calculates a cosine similarity matrix between each row vector from G and each column vector from H^T . Note that $D_{X_B} - P_{A \rightarrow B} D_{X_A}$ estimates the error of using the population density variation matrix D_{X_A} to reconstruct D_{X_B} via intercity spatial mapping $P_{A \rightarrow B}$. Therefore, $W[l_B, l_A]$ will have a higher value if the population density variation from location l_A compensates for the reconstruction error of l_B .

We have so far completed our estimation of the intercity spatial mapping. A pseudo-routine for the EM algorithm summarizing the essential equations in this section is given in Algorithm



Figure 4. Visualization of a typical trajectory in Tokyo that travels through Toukaido line (on the left) and the mapping results (on the right) from $P_{Tokyo \rightarrow Osaka}$ for four important stations on this railway: Kawasaki, Shinagawa, Shimbashi, and Tokyo stations.

1. In Figure 4, we visualize a typical trajectory in Tokyo that travels through the Toukaido line (left panel), which is the main railway connecting Tokyo and Yokohama, Tokyo's largest neighboring city. We map four important stations on the railway in Tokyo (Kawasaki, Shinagawa, Shimbashi, and Tokyo stations) and visualize the corresponding spatial distribution in Osaka (right panel) by extracting the corresponding column of the intercity spatial mapping $P_{Tokyo \rightarrow Osaka}$. We can see that these four spatial distributions roughly indicate a trajectory that travels from the north to the center of Osaka. It is interesting that Kyoto, which is the second largest city neighboring Osaka, is located north to the Osaka, and the spatial distribution sequence roughly mirrors the main railway (Kyoto line) that travels from Kyoto to Osaka.

SIMULATED TRAJECTORY GENERATION

To transfer rare event human mobility from city A to city B , we have shown how to estimate an intercity spatial mapping and generate a spatial distribution sequence in city B (Figure 4) from a trajectory in city A in the previous section. In this section, we present a spatio-temporal filtering method and other cues to generate more realistic trajectories in city B .

For each subject, HMM is a good choice to incorporate both the information from spatial mappings and geographical transitions. However, if we treat each subject independently, the Viterbi algorithm for HMM inference will disrupt the population density distribution at the citywide scale, because "maximum a posterior" characteristics will enforce a tendency of a higher density region so that lower density region will be underestimated.

To cope with this difficulty, we propose a Gibbs sampling-based multiple hidden Markov model (GSMHMM) to generate simulated trajectories that incorporate cues from intercity spatial mapping, population density, and hop distance. In Figure 5, we illustrate a simulated trajectory generation algorithm in a probabilistic graphical model, from which we can see that generating each trajectory in city B defines a hidden Markov chain, and all trajectories are dependent on the global variable of population density. Moreover, the transition probability is defined by gamma distributions [22] with shape parameters α , determined by the corresponding hop distances from the source trajectory in city A . From Figure 5, the joint generative probability distribution can be decomposed as

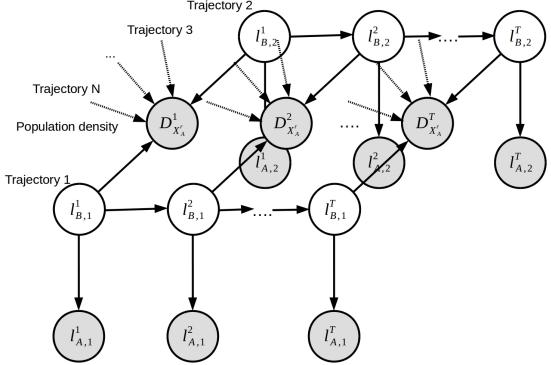


Figure 5. A graphical model that illustrates our Gibbs sampling-based multiple hidden Markov model for simulated trajectory generation. Shaded nodes denote the observed data and unshaded nodes denote latent variables.

$$\begin{aligned} p\left(X_A^r, X_{B|A}^r, D_{X_A^r}, P_{A \rightarrow B}\right) = \\ p\left(D_{X_A^r}^1 | l_{B,U_A}^1, P_{A \rightarrow B}\right) p\left(l_{A,U_A}^1 | l_{B,U_A}^1, P_{A \rightarrow B}\right) p\left(l_{B,U_A}^1\right) \\ \prod_{\tau=2}^T p\left(l_{B,U_A}^\tau | l_{B,U_A}^{\tau-1}\right) p\left(D_{X_A^r}^\tau | l_{B,U_A}^\tau, P_{A \rightarrow B}\right) p\left(l_{A,U_A}^\tau | l_{B,U_A}^\tau, P_{A \rightarrow B}\right) \quad (9) \end{aligned}$$

where X_A^r is the discrete GPS log data from city A (r indicates that X_A^r contains rare event and is utilized as the source data for human mobility during rare events), while $X_{B|A}^r$ is our simulated GPS data for city B based on X_A^r . In addition, U_A is the set of all unique IDs in X_A^r and l_{A,U_A}^τ is the observed locations of all trajectories in U_A at time τ from city A , while l_{B,U_A}^τ is the simulated location of trajectories at time τ in $X_{B|A}^r$. $D_{X_A^r}^\tau$ denotes the population density distribution calculated from X_A^r in city A . In the following subsections, we discuss the details of calculating each term in this equation.

Emission Probability

From Equation 9, the emission probability of each hidden Markov chain has two inputs: $p\left(l_{A,U_A}^\tau | l_{B,U_A}^\tau\right)$, the spatial distribution estimated by the trajectories from city A , and $p\left(D_{B|A,P_{A \rightarrow B}}^\tau | l_{B,U_A}^\tau\right)$, a posteriori population density.

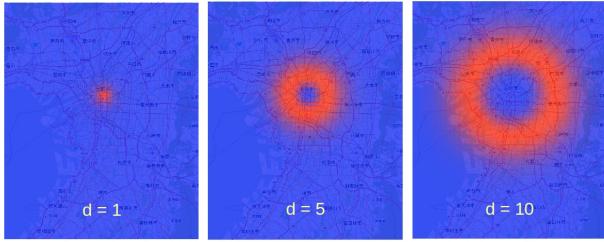


Figure 6. The gamma distribution under different shape parameters $\alpha = d + 0.4$ and $\beta = 1$, where d is the distance traversed from city A , with a grid size of approximately 900m.

For the first term, all trajectories are conditionally independent and can be easily calculated by drawing the location l_{B,u_A}^τ in the spatial distribution from the l_{A,u_A}^τ -th column of $P_{A \rightarrow B}$ for all trajectories $u_A \in U_A$:

$$p(l_{A,u_A}^\tau | l_{B,U_A}^\tau) \propto \prod_{u_A \in U_A} p(l_{B,u_A}^\tau | l_{A,u_A}^\tau) = \prod_{u_A \in U_A} P_{A \rightarrow B} [l_{B,u_A}^\tau, l_{A,u_A}^\tau] \quad (10)$$

However, such conditional independence is not satisfied in the second term because all trajectories are dependent on the observed population density variable $D_{X_A}^\tau$. We propose a Gibbs sampling-based algorithm that for each iteration, resamples only one trajectory while assume all others are observed.

$$p(D_{X_A}^\tau | l_{B,U_A}^\tau, P_{A \rightarrow B}) \propto p(l_{B,u_A}^\tau | D_{X_A}^\tau, l_{B,U_A}^{\tau-u_A}, P_{A \rightarrow B}) \quad (11)$$

where $U_A^{\tau-u_A}$ denotes the set U_A with u_A removed.

To compute the conditional probability of Equation 11, first we must calculate the estimated population density distribution $D_{B|X_A}^\tau$. Given population density $D_{X_A}^\tau$ in city A and the inter-city spatial mapping $P_{A \rightarrow B}$, the estimated population density distribution can thus be computed as:

$$D_{B|X_A}^\tau = P_{A \rightarrow B} D_{X_A}^\tau \quad (12)$$

Therefore, the estimated population density distribution, along with the simulated trajectories $U_A^{\tau-u_A}$, defines the conditional probability for generating a new trajectory as

$$\begin{aligned} p(l_{B,u_A}^\tau | D_{X_A}^\tau, l_{B,U_A}^{\tau-u_A}, P_{A \rightarrow B}) &\propto \\ \exp \left\{ \eta \left(\|D_{B|X_A}^\tau - D_{U_A}^{\tau-u_A}\|_2^2 - \|D_{B|X_A}^\tau - D_{U_A}^\tau\|_2^2 \right) \right\} \end{aligned} \quad (13)$$

where η is a constant controlling the weight of the discriminative power of this term, while $D_{U_A}^\tau$ and $D_{U_A}^{\tau-u_A}$ are the population densities generated from l_{B,u_A}^τ and $l_{B,U_A}^{\tau-u_A}$ respectively. Intuitively, if the simulated location l_{B,u_A}^τ sums the error of the estimation of population density, it will be designated as having a smaller weight in the conditional probability in Equation 13.

Transition Probability

The hop distance of the trajectory in city A can also be treated as important prior knowledge for simulating trajectories in city B . More specifically, we may expect our simulated trajectory

Input: GPS log data for training X_A, X_B
GPS log data for transferring X_A^r
Output: Simulated GPS data $X_{B|A}^r$

```

/* Intercity spatial mapping estimation */
Estimate  $P_{A \rightarrow B}$  from  $X_A, X_B$  using Algo. 1;

/* Simulated Trajectory Generation */
for  $i = 1$  to  $ITERATION_{GSMHMM}$  do
    /* Iterate  $U_A$  in a random order */
    for  $u_A$  in  $shuffle(U_A)$  do
        /* Calculate emission and transition probabilities */
        Compute  $p(l_{B,u_A}^\tau | D_{X_A}^\tau, l_{B,U_A}^{\tau-u_A}, P_{A \rightarrow B})$  using Eq. 13;
        for  $\tau = 1$  to  $T$  do
            Compute  $p(l_{A,u_A}^\tau | l_{B,U_A}^\tau)$  using Eq. 10;
            Compute  $p(l_{B,U_A}^\tau | l_{B,U_A}^{\tau-1})$  using Eq. 14;
        end
        /* Viterbi algorithm for latent sequence inference */
        Initialize Viterbi algorithm with the above
        probabilities;
        Find the optimal sequence  $l_{B,u_A}^{[1:T]}$ ;
        Update the trajectory of  $u_A$  in  $X_{B|A}^r$  by  $l_{B,u_A}^{[1:T]}$ ;
    end
return  $X_{B|A}^r$ ;

```

Algorithm 2: CityCoupling Algorithm

in city B to have a similar hop distance to the corresponding trajectory in city A . To this end, we model the transition probability $p(l_{B,u_A}^\tau | l_{B,U_A}^{\tau-1})$ as a gamma distribution with shape parameters determined by the hop distance in city A :

$$p(l_{B,u_A}^\tau | l_{B,U_A}^{\tau-1}) = \prod_{u_A \in U_A} \text{Gamma} \left(d(l_{A,u_A}^{\tau-1}, l_{A,u_A}^\tau) + \mu, \beta \right) \quad (14)$$

where μ is the bias induced by discretization and β is the predefined rate parameter of the gamma distribution. $d(l_1, l_2)$ is the function calculating the distance (in grid size) between l_1 and l_2 . In Figure 6, we have shown the power of the gamma distribution. We visualize the transition probability distribution from the Osaka station under three different hop distances in Tokyo.

Having obtained the emission and transition probabilities, we infer the optimal simulated location sequence $l_{B,u_A}^{[1:T]}$ by applying the Viterbi algorithm. Recall our estimation of the population density posterior from Equation 13. We utilize a Gibbs sampling framework that randomly samples and infers the optimal location sequence for each trajectory in X_A^r . To conclude this section, we give the entire pseudo-routine of the CityCoupling algorithm in Algorithm 2.

EXPERIMENTAL RESULTS

In this section, we present experimental results from two rare events: the Great Eastern Japan Earthquake that quenched

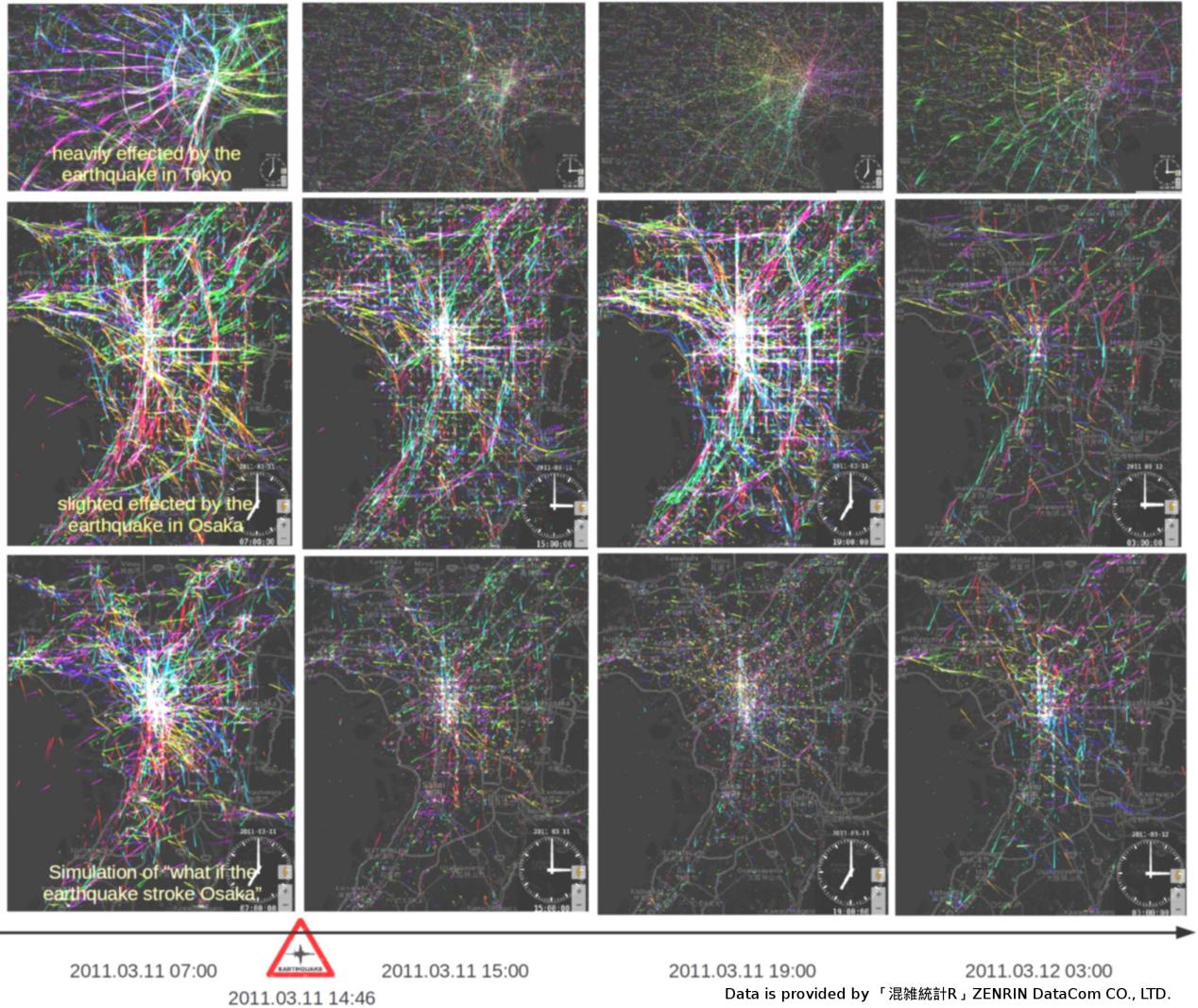


Figure 7. ¹A visualization transferring human mobility during the Great Eastern Japan Earthquake in Tokyo (first row), which was heavily affected, to Osaka (second row), which was only slightly affected. We apply our CityCoupling algorithm to generate a simulation of human mobility during the earthquake in Osaka (third row). We compare citywide human mobility before the earthquake (first column), immediately after the earthquake (second column), 4h after the earthquake (third column), and 12h after the earthquake (fourth column), at which point human mobility starts to recover from the earthquake. The hue of the trajectory is colored by direction.

citywide human mobility and a celebration New Year's eve celebration. In the rest of this section, we qualitatively compare our simulation in Osaka with routine human mobility in Osaka or human mobility during rare events in Tokyo. Meanwhile, we quantitatively evaluate our CityCoupling algorithm with and without GSMHMM smoothing on the New Year's Eve data.

Data

To address the problem of real-world human mobility and develop innovative applications, NTT DOCOMO INC collects an anonymous GPS log dataset¹, "Konzatsu-Tokei (R)" Data, from about 1.6 million real mobile-phone users (30 billion GPS records in total) in Japan over a three-year period (August 1, 2010 to July 31, 2013). "Konzatsu-Tokei (R)" Data refers to people flows data collected by individual location data

sent from mobile phone with enabled AUTO-GPS function under users' consent, through the "docomo map navi" service provided by NTT DOCOMO INC. Those data are processed collectively and statistically in order to conceal the private information. Original location data is GPS data (latitude, longitude) sent in about every a minimum period of 5 minutes and does not include the information to specify individual such as gender or age.

In our experiment, the proposed methodology is applied to raw GPS data by NTT DOCOMO INC. The records within the Tokyo ($35.5 < \text{latitude} < 35.796, 139.4 < \text{longitude} < 139.9$) and Osaka ($34.45 < \text{latitude} < 34.85, 135.3 < \text{longitude} < 135.7$) areas are cropped. A time span between June 1, 2012 and July 31, 2012, is used to estimate intercity spatial mapping, and the rest of the cropped dataset is utilized to transfer rare event human mobility between these two cities.

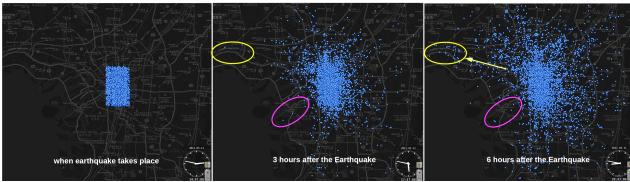


Figure 8. A simulation of the dispersion behavior of people in Osaka if this earthquake occurred there.

Experimental Setup

Here, we detail the parameter settings in the experiment. To obtain the discrete GPS dataset, we discretize the time by 20 min; the grid size is 0.008 in latitude and 0.010 in longitude (approximately 900m × 900m). In the intercity spatial mapping estimation phase, we set a random sample size $N = 200$ in Equation 1, $\lambda = 0.01$ in Equation 5, $\gamma = 0.005$ in Equation 6, and $ITERATION_{mapping} = 1000$. In the simulated trajectory generation phase, we set $\eta = 0.01$ in Equation 13, and the bias induced by discretization $\mu = 0.4$ (as the expectation hop distance is discretized into the same location), the shape parameter $\beta = 1$, and $ITERATION_{GSMHMM} = 10$.

Experiment on the Great Eastern Japan Earthquake

On Friday, March 11, 2011, a tremendous earthquake (a magnitude $M_w = 9.0$) with its epicenter approximately 70 km offshore struck eastern Japan. As shown in the top row in Figure 7, Tokyo was heavily affected and the earthquake reduced citywide human mobility due to severe disruptions of the transportation network. The earthquake occurred at 14:46 on a weekday; therefore, most people were trapped at their workplace, because the transportation network failed to recover until the next morning. However, about 400 km west of Tokyo, Osaka was only slightly affected by the earthquake. From the second row in Figure 7, we can see that in Osaka human mobility on that day did not differ significantly from a regular day.

With an intercity spatial mapping estimation from Tokyo to Osaka, we generate the trajectories simulating human mobility in response to the same disaster as in Tokyo. As shown in the third row, our simulated trajectories preserve different phases of the human response to the earthquake "regular → heavily affected by the disaster → gradual recovery". Compared with human mobility in Osaka on that day, we find that our simulation results in the first and fourth columns, which are from before and 12h after the earthquake (when the transportation network had mostly recovered), are considerably similar to the true human mobility at those times, while those in the second and the third column, which depict transportation network disruption, are substantially different from the second row. Putting the first and third rows together, we find a similar human mobility pattern regardless of city layout.

By transferring the knowledge of human mobility in Tokyo during the earthquake, this simulation result can help the Osaka government to preview and take precautions should the same disaster strike Osaka. The 311 Earthquake took place in the afternoon on a weekday, when most of people in Tokyo are working in the business and commercial areas. As a result, one

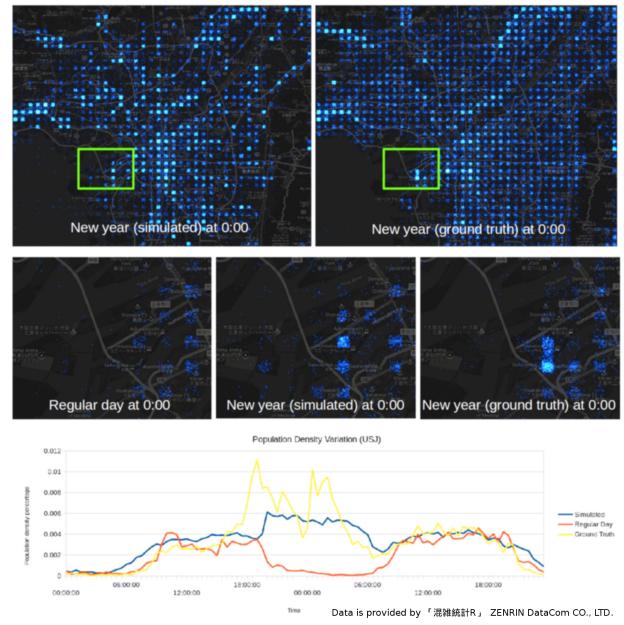


Figure 9. ¹A comparison of our simulated human mobility on New Year's Eve at 00:00 in Osaka with the ground truth (first row), specifically, near Universal Studios Japan (USJ) (within the green bounding box), which is normally closed at this time while still open for New Year Countdown celebration (second row). The population density variation (population density percentage of the whole city) at USJ over two days is illustrated in the third row.

of the biggest challenges of emergency management is how to guarantee that such a large number of victims can either return home or find suitable shelter to await for the resumption of the transportation network operation [21]. Figure 8 shows simulated dispersion behavior by tracking subjects in the central area in Osaka. There are two regions that are worth noting. The first, marked by a yellow oval, is a transportation hub connecting Osaka and Kobe, which is the biggest city that neighbors Osaka. Although there are few subjects who reach this region 3h after the earthquake (on a regular day, this distance can be covered within 1h using railway), a considerable number of subjects are heading to this region and 6h after the earthquake, we can see a dramatic increase in subjects, and a long queue of followers, heading to this region. Conversely, although the region within the magenta oval (where Osaka USJ and Osaka Bay are located), is quite close to the central area, few people would like to travel to this region after the earthquake. With our simulation, the Osaka government can have a clear view of what the citywide human mobility would be if the 311 earthquake struck Osaka.

Experiment on New Year's Eve

Note that it is hard to evaluate our simulation results on the Great Eastern Japan Earthquake, because there is no ground truth for the human mobility in Osaka during the earthquake. However, the New Year's Eve celebrations take place in both Tokyo and Osaka at the same time, hence can be used to evaluate the performance of our CityCoupling algorithm.

The first row of Figure 9 shows a comparison of our simulated human mobility on New Year's Eve at 00:00 in Osaka with

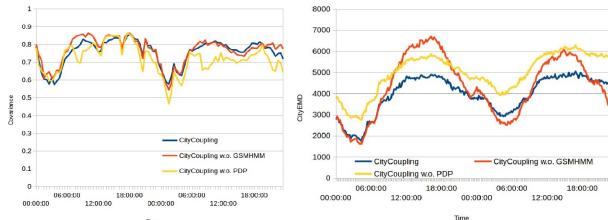


Figure 10. A comparison of the population variation covariance and the CityEMD for CityCoupling, CityCoupling without GSMHMM, and CityCoupling without PDP.

real human mobility. We can see that most of the crowded places have been correctly simulated by our algorithm. Taking USJ as an example, which is highlighted by the green bounding box and is enlarged in the second row, we compare human mobility on a regular day (June 1, 2012). The third row of Figure 9 illustrates the population density variation of our simulated results (blue), the ground truth (yellow), and regular days (red). Except for the New Year's Eve, USJ is closed at night; hence, we can see from the leftmost figure in the second row and the red curve in the third row that there are few people there at night on a regular day. What is interesting is that in our training data for the intercity spatial mapping estimation, there is little information regarding a nighttime gathering in USJ, nor do we have any axillary information indicating that there is a big event in USJ on New Year's Eve. Human mobility in USJ on New Year's Eve is simulated by establishing a connection to the most similar places in Tokyo, for example the Disney park, which is also a theme park that is closed at night on a regular day and holds an overnight party to celebrate the coming of the new year. When the connection is established through the intercity spatial mapping, the human mobility for a night gathering for New Year's Eve can be naturally transferred. Therefore, rather than a population density decrease of USJ at about 19:00, we successfully simulate an increase of population density without any previous knowledge of this New Year's Eve celebration in USJ.

Evaluation

To quantitatively evaluate our algorithm, we define three metrics: 1) the population variation covariance, which treats population density at each time as a random variable and calculates the covariance between our simulation and the ground truth for each value of time t ; 2) CityEMD defined in [11], which characterizes the sequential patterns of citywide human mobility; and 3) average speed. Using these metrics, we conduct a quantitative comparison of 1) the CityCoupling algorithm; 2) the CityCoupling without GSMHMM, which uses the intercity spatial mapping $P_{Tokyo \rightarrow Osaka}$ to draw the simulated locations from the corresponding column vector of $P_{Tokyo \rightarrow Osaka}$; and 3) CityCoupling without the population density posterior (PDP) by setting $\eta = 0$ in Equation 13.

As shown in the left panel of Figure 10, we compare the population variation covariance of the three algorithms. On average, the covariance for CityCoupling is 0.7549, for CityCoupling without GSMHMM it is 0.7671 and for CityCoupling without PDP it is 0.7171. We can see that CityCoupling without

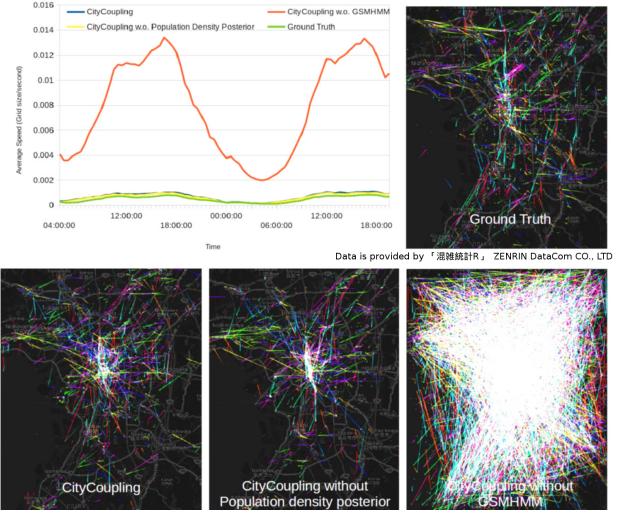


Figure 11. ¹A comparison of the average speed of the three algorithms (CityCoupling, CityCoupling without GSMHMM, and CityCoupling without PDP) with the ground truth. For a more intuitive view of the differences in average speeds, we visualize human mobility at 7 a.m. on 31 December, 2011 from the results of each algorithm.

GSMHMM has a slight performance advantage over CityCoupling, while it significantly outperforms CityCoupling without PDP. We conjecture that this is because the intercity spatial mapping guarantees robust population density estimation in Osaka, while the maximum a posteriori sense of Viterbi algorithm disrupts this estimation. As a result, the algorithm without PDP suffers from over-concentration at high probability locations. Meanwhile, the GSMHMM algorithm generates more realistic trajectories at the expense of a negligible loss in covariance. We further show that the result of CityCoupling is more realistic using the remaining two metrics.

From the right panel of Figure 11, we find that the average speed of CityCoupling with and without PDP are close to the ground truth, while the average speed dramatically drifts from the ground truth without GSMHMM. The visualization of simulated trajectories also makes this limitation apparent. At first glance, human mobility from the CityCoupling with and without PDP are similar to the ground truth. On the contrary, the result from the CityCoupling without GSMHMM looks substantially different from the ground truth. This is mainly because of the multi-modality underlying the intercity spatial mapping. Considering a location l_{Tokyo} , where the geographical distribution with respect to l_{Tokyo} has multiple "peaks" in Osaka, even for a subject staying in l_{Tokyo} statically, multi-modality would very likely lead to the simulated trajectory moving among the "peaks" in Osaka. This problem can be solved by spatio-temporal filtering, namely the GSMHMM algorithm in this study.

RELATED WORKS

Human Mobility Big Data & Urban Computing: In recent years, there has been an explosion of emerging big human mobility datasets available to researchers, such as GPS log data from cell phones [10, 20, 23], taxis [24], ships [4], as well as Call Records Details data [6, 5] and location-based Apps [19, 15]. Researchers from different backgrounds have shown the immeasurable potential for societal benefit from these big human mobility datasets through solving a wide range of ubiquitous computing problems that cannot be solved in traditional ways. For instance, these data can be used for an emergency management system that predicts human movement during a natural disaster [20] or a major event [11], a finer granularity of air quality estimation [27], optimizing station configuration for bike-sharing [3] and assessing the potential of car-sharing in a city [6]. Good surveys on these related topics can be found in [12] and [26].

Human Mobility Modeling: In general, we categorize human mobility modeling into two types: routine pattern modeling and rare behavior modeling. Routine pattern modeling is relatively simple and has been well-studied in a number of state-of-the-art works [10, 9, 15, 14], while rare behavior modeling is much more daunting and understudied because of the particularity of rare events and a variety of potential factors that may affect human mobility. Many previous studies either model a particular type of rare event (e.g. earthquakes [20]) heuristically, or only focus on the trend of variation of human mobility to make short-term predictions during a rare event [11]. In this paper, we present a novel perspective on transferring human mobility during a rare event from one city to another. Therefore, this simulated human mobility preserves the semantic information of human mobility during a rare event while the trajectories adapt to the layout of the target city.

Transfer Learning: This work can also be viewed from a Transfer Learning perspective. Using the terms in [18], we can reinterpret our algorithm as an example of domain adaption, in which the source domain is human mobility in the training time period (on regular days) in city *A*, while the target domain is human mobility in city *A* when a rare event takes place. Source domain "labels" are given by human mobility data during the training period in city *B*, while the target task is to predict human mobility in city *B* in the target domain. Contrary to most traditional transfer learning research [8, 16, 25, 28], we are handling citywide human mobility that is high-dimensional and has highly structured "domains" and "labels". The fact that every trajectory in the source domain is not explicitly labeled increases the complexity of the problem. Specifically, the complete set of all trajectories in city *A* is labeled with all trajectories in city *B*, without knowing their correspondence. To cope with this difficulty, we introduce intercity trajectory matching as latent variables and iteratively estimate their correspondence to label the source domain and obtain intercity spatial mapping that transfer knowledge to the target domain.

CONCLUSION

In this paper, we propose an EM framework to estimate intercity spatial mapping and a Gibbs sampling multiple hidden Markov model to generate simulated trajectories. We validate our algorithm with a big real-world GPS log dataset and show that we can successfully transfer human mobility during rare events such as the Great Eastern Japan Earthquake or the New Year's Eve celebrations, between cities.

We note that several limitations to our current work and there is still room for improvement in the current version of our algorithm. First, we do not use any map information in this version, though we believe that map information can be very helpful in implementing a more robust intercity spatial mapping, as well as more realistic trajectory generation. Second, some auxiliary information is not included in this framework. For example, if we have human mobility from Tokyo during 2020 Olympic Games, and we want to indicate the Olympic venues in Osaka in order to simulate the human mobility during a hypothetical Olympic Games held there.

ACKNOWLEDGEMENT

This work was partially supported by, JST, Strategic International Collaborative Research Program (SICORP); Grant-in-Aid for Young Scientists (26730113) of Japan's Ministry of Education, Culture, Sports, Science, and Technology (MEXT); US National Science Foundation under grant CNS-1461926; Microsoft Research collaborative research (CORE) program; and DIAS/GRENE project of MEXT. We specially thank ZENRIN DataCom CO., LTD for their supporting.

REFERENCES

1. Japanese National Police Agency. 2012a. Damage Situation and Police Countermeasures associated with 2011 Tohoku district-off the Pacific Ocean Earthquake. (2012). <https://tokyo2020.jp/en/>.
2. Japanese National Police Agency. 2012b. Damage Situation and Police Countermeasures associated with 2011 Tohoku district-off the Pacific Ocean Earthquake. (2012). <http://www.npa.go.jp/archive>.
3. Longbiao Chen, Daqing Zhang, Gang Pan, Xiaojuan Ma, Dingqi Yang, Kostadin Kushlev, Wangsheng Zhang, and Shijian Li. 2015. Bike Sharing Station Placement Leveraging Heterogeneous Urban Open Data. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '15)*. ACM, New York, NY, USA, 571–575.
4. Longbiao Chen, Daqing Zhang, Gang Pan, Leye Wang, Xiaojuan Ma, Chao Chen, and Shijian Li. 2014. Container Throughput Estimation Leveraging Ship GPS Traces and Open Data. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '14)*. ACM, New York, NY, USA, 847–851.
5. Blerim Cici, Minas Gjoka, Athina Markopoulou, and Carter T. Butts. 2015. On the Decomposition of Cell Phone Activity Patterns and Their Connection with Urban Ecology. In *Proceedings of the 16th ACM*

- International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc '15).* ACM, New York, NY, USA, 317–326.
6. Blerim Cici, Athina Markopoulou, Enrique Frias-Martinez, and Nikolaos Laouraris. 2014. Assessing the Potential of Ride-sharing Using Mobile and Social Data: A Tale of Four Cities. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '14)*. ACM, New York, NY, USA, 201–211.
 7. Brian Clapper. 2014. munkres algorithm for the Assignment Problem. (2014). <https://pypi.python.org/pypi/munkres/>.
 8. Wenyuan Dai, Gui-Rong Xue, Qiang Yang, and Yong Yu. 2007. Co-clustering Based Classification for Out-of-domain Documents. In *Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '07)*. ACM, New York, NY, USA, 210–219.
 9. Nathan Eagle and Alex Pentland. 2006. *Eigenbehaviors: Identifying Structure in Routine*. Technical Report. IN PROC. OF UBICOMP 06.
 10. Zipei Fan, Xuan Song, and Ryosuke Shibasaki. 2014. CitySpectrum: A Non-negative Tensor Factorization Approach. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '14)*. ACM, New York, NY, USA, 213–223.
 11. Zipei Fan, Xuan Song, Ryosuke Shibasaki, and Ryutaro Adachi. 2015. CityMomentum: An Online Approach for Crowd Behavior Prediction at a Citywide Level. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '15)*. ACM, New York, NY, USA, 559–569.
 12. Bin Guo, Zhu Wang, Zhiwen Yu, Yu Wang, Neil Y. Yen, Runhe Huang, and Xingshe Zhou. 2015. Mobile Crowd Sensing and Computing: The Review of an Emerging Human-Powered Sensing Paradigm. *ACM Comput. Surv.* 48, 1 (2015), 7.
 13. Fred M. Hoppe. 1984. Polya-like urns and the Ewens' sampling formula. *Journal of Mathematical Biology* 20, 1 (1984), 91–94.
 14. Sibren Isaacman, Richard Becker, Ramón Cáceres, Margaret Martonosi, James Rowland, Alexander Varshavsky, and Walter Willinger. 2012. Human Mobility Modeling at Metropolitan Scales. In *Proceedings of the 10th International Conference on Mobile Systems, Applications, and Services (MobiSys '12)*. ACM, New York, NY, USA, 239–252.
 15. Takeshi Kurashima, Tomoharu Iwata, Takahide Hoshide, Noriko Takaya, and Ko Fujimura. 2013. Geo Topic Model: Joint Modeling of User's Activity Area and Interests for Location Recommendation. In *Proceedings of the Sixth ACM International Conference on Web Search and Data Mining (WSDM '13)*. ACM, New York, NY, USA, 375–384.
 16. Lilyana Mihalkova, Tuyen Huynh, and Raymond J. Mooney. 2007. Mapping and Revising Markov Logic Networks for Transfer Learning. In *Proceedings of the 22Nd National Conference on Artificial Intelligence - Volume 1 (AAAI'07)*. AAAI Press, 608–614.
 17. James R. Munkres. 1957. Algorithms for the Assignment and Transportation Problems. *J. Soc. Indust. Appl. Math.* 5, 1 (March 1957), 32–38.
 18. Sinno Jialin Pan and Qiang Yang. 2010. A Survey on Transfer Learning. *IEEE Trans. on Knowl. and Data Eng.* 22, 10 (Oct. 2010), 1345–1359.
 19. Masamichi Shimosaka, Keisuke Maeda, Takeshi Tsukiji, and Kota Tsubouchi. 2015. Forecasting Urban Dynamics with Mobility Logs by Bilinear Poisson Regression. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '15)*. ACM, New York, NY, USA, 535–546.
 20. Xuan Song, Quanshi Zhang, Yoshihide Sekimoto, and Ryosuke Shibasaki. 2014. Prediction of Human Emergency Behavior and Their Mobility Following Large-scale Disaster. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '14)*. ACM, New York, NY, USA, 5–14.
 21. The Japan Times. 2012. Your memories of March 11, 2011. (2012). <http://www.japantimes.co.jp/news/2012/03/11/national/your-memories-of-march-11-2011>.
 22. Eric W Weisstein. 2012. "Gamma Distribution." From MathWorld—A Wolfram Web Resource. (2012). <http://mathworld.wolfram.com/GammaDistribution.html>.
 23. Zhiwen Yu, Hui Wang, Bin Guo, Tao Gu, and Tao Mei. 2015. Supporting Serendipitous Social Interaction Using Human Mobility Prediction. *IEEE Trans. Human-Machine Systems* 45, 6 (2015), 811–818.
 24. Fuzheng Zhang, Nicholas Jing Yuan, David Wilkie, Yu Zheng, and Xing Xie. 2015. Sensing the Pulse of Urban Refueling Behavior: A Perspective from Taxi Mobility. *ACM Trans. Intell. Syst. Technol.* 6, 3, Article 37 (April 2015), 23 pages.
 25. Vincent Wenchen Zheng, Evan Wei Xiang, Qiang Yang, and Dou Shen. 2008. Transferring Localization Models over Time. In *Proceedings of the Twenty-Third AAAI Conference on Artificial Intelligence, AAAI 2008, Chicago, Illinois, USA, July 13-17, 2008*, 1421–1426.
 26. Yu Zheng, Licia Capra, Ouri Wolfson, and Hai Yang. 2014. Urban Computing: Concepts, Methodologies, and Applications. *ACM Transaction on Intelligent Systems and Technology* (October 2014).
 27. Yu Zheng, Furui Liu, and Hsun-Ping Hsieh. 2013. U-Air: When Urban Air Quality Inference Meets Big Data. In *KDD 2013*. ACM.
 28. Hankui Zhuo, Qiang Yang, and Lei Li. 2009. *Transfer Learning Action Models by Measuring the Similarity of Different Domains*. Springer Berlin Heidelberg, Berlin, Heidelberg, 697–704.