Real-Time Human Mobility Modeling with Multi-View Learning

DESHENG ZHANG, Rutgers University
TIAN HE, University of Minnesota
FAN ZHANG, Shenzhen Institutes of Advanced Technology, China

Real-time human mobility modeling is essential to various urban applications. To model such human mobility, numerous data-driven techniques have been proposed. However, existing techniques are mostly driven by data from a single view, for example, a transportation view or a cellphone view, which leads to overfitting of these single-view models. To address this issue, we propose a human mobility modeling technique based on a generic multi-view learning framework called coMobile. In coMobile, we first improve the performance of single-view models based on tensor decomposition with correlated contexts, and then we integrate these improved single-view models together for multi-view learning to iteratively obtain mutually reinforced knowledge for real-time human mobility at urban scale. We implement coMobile based on an extremely large dataset in the Chinese city Shenzhen, including data about taxi, bus, and subway passengers along with cellphone users, capturing more than 27 thousand vehicles and 10 million urban residents. The evaluation results show that our approach outperforms a single-view model by 51% on average. More importantly, we design a novel application where urban taxis are dispatched based on unaccounted mobility demand inferred by coMobile.

Categories and Subject Descriptors: H.4 [**Information System Application**]: Miscellaneous

General Terms: Algorithms, Model, Experimentation, Application

Additional Key Words and Phrases: Smart cities, mobility model, model integration

ACM Reference format:
Desheng Zhang, Tian He, and Fan Zhang. 2017. Real-Time Human Mobility Modeling with Multi-View Learning. *ACM Trans. Intell. Syst. Technol.* 9, 3, Article 22 (December 2017), 25 pages.
https://doi.org/10.1145/3092692

1 INTRODUCTION

Nowadays, we are in a rapid process of urbanization where more than half of the people in the world has moved to urban areas (Zheng et al. 2014; Yuan et al. 2011a, 2012; Zhang et al. 2015a; Zheng et al. 2011b). Such quick urbanization leads to a new interdisciplinary research area called *urban computing* (Zheng et al. 2014), where various data generated from urban areas are used to improve urban efficiency (Aslam et al. 2012; Balan et al. 2011; Cho et al. 2011; Liu et al. 2011; Shang

This work was supported in part by the US NSF Grants CNS-1544887, CNS-1446640 and China 973 Program No. 2015CB352400. A preliminary result of this work was published at Zhang et al. (2015b).
Authors' addresses: D. Zhang, Department of Computer Science, Rutgers University, 110 Frelinghuysen Rd, Piscataway Township, NJ 08854; email: desheng.zhang@cs.rutgers.edu; T. He, Computer Science Department, University of Minnesota, 200 Union St SE, Minneapolis, MN 55455; email: tianhe@umn.edu; F. Zhang, Dean, Shenzhen Institute of Advanced Technology, Chinese Academy of Science, China, 1068 Xueyuan Ave, Shenzhen, Guangdong, China; email: zhangfan@siat.ac.cn.

https://doi.org/10.1145/3092692

et al. 2014; Wei et al. 2012; Wang et al. 2014; Zhang et al. 2012; Isaacman et al. 2012; Lathia and Capra 2011; Rhee et al. 2008; Bhattacharya et al. 2013). In urban computing, how to capture human mobility at urban scale is one of the fundamental challenges we need to address. Such human mobility has many real-world urban applications, for example, urban planning, transportation, social networking, and location based services (Zheng 2015). To capture generic human mobility patterns, several theoretical models have been proposed, for example, the gravity model and the radiation model (Simini et al. 2014). However, a key drawback of these theoretical models is that they cannot capture human mobility at fine spatiotemporal granularity, for example, mobility at small region levels in *real time*.

Recently, thanks to upgrades of urban infrastructures, many real-time location-tracking devices have become available, for example, cellphones, onboard GPS devices, and smartcards. These devices generate massive real-time location data, which hold the key potential to revolutionize real-time human mobility modeling. Based on these real-time data, several data-driven models have been proposed, for example, driven by data from cellphones (Jiang et al. 2013), smartcards (Sun et al. 2012), taxis (Ganti et al. 2013), buses (Bhattacharya et al. 2013), or subways (Giannotti et al. 2011). However, a common feature of these models is that they capture mobility only from one view, for example, a cellphone view or a transportation view. These single-view models are sufficient if single-view data are complete, but in reality this is not the case. From the cellphone view, the models driven by cellphone data cannot capture residents without cellphone data, for example, residents who do not have cellphones and residents who have cellphones but do not use their cellphones during our modeling time; similarly, from the transportation view, the models driven by one kind of transportation data, for example, taxi, cannot capture the passengers who use other transportation modes, for example, bus and subway, and further there is no urban infrastructure that can capture private vehicles at urban scale. To our knowledge, no data-driven urban human mobility models are based on a complete view so far. As a result, these single-view human mobility models essentially use residents captured by these single views as a sample to study all residents, which inevitably leads to a bias and thus over-fitting of their models, as shown in Section 2.

To address this issue, we aim to combine different views for multi-view modeling. Each view is incomplete to capture mobility by itself, but one view is often complementary to others, for example, the cellphone view can capture some private-vehicle passengers, whereas the transportation view can capture some inactive cellphone users. But a view's ability to capture human mobility is unknown *a priori* and is highly dynamic based on spatiotemporal contexts. As a result, such dynamic view completeness makes multi-view human mobility modeling extremely challenging.

In this work, we propose coMobile, a generic framework to capture human mobility with a multiple-view learning technique. In coMobile, we first design a single-view learning technique based on context-based tensor decomposition to improve completeness of single-view models. Then, we integrate those improved single-view models together by formulating a convex optimization to obtain the ground truth of urban mobility. Mostly importantly, we implement coMobile based on extremely large datasets in the Chinese city Shenzhen with cellphone data and transportation data including taxis, buses, and subways. In particular, the key contributions of the article are as follows:

—We propose the first multi-view learning framework for human mobility to integrate incomplete yet complementary knowledge from individual views. To our knowledge, the proposed model is the only human mobility model driven by more than one view, which aims to address over-fitting of single view models. It is challenging to apply multi-view learning in human mobility modeling, because data-driven views are mostly incomplete to urban-scale mobility.

—We design a single-view learning technique based on context-aware tensor decomposition with both real-time and historical data to improve completeness of single-view models. This technique addresses data sparsity challenges of particular views to improve their completeness. In particular, we use a cellphone-view model as an example to show how we extract three contexts, that is, cellphone user density, calling location patterns, and calling time patterns, based on historical data for joint tensor decomposition.
—Based on improved single-view models, we formulate a multi-view modeling problem by designing a joint optimization, which minimizes overall weighted deviation from observed mobility to the ground truth. To solve this optimization, we propose an iterative learning process to alternatively update ground truth and view completeness until no further improvement can be made for the objective function. We prove the convexity of the optimization and the convergence of our iterative learning.
—We implement our multi-view human-mobility model based on two datasets in the Chinese city Shenzhen, with 10 million cellphone users and 16 million smartcard users involved. To our knowledge, this is one of the largest human mobility models driven by real-world datasets. We evaluate our model by comparing it to single-view models driven by cellphone and transportation data, and results show that we reduce error rates by 51% and 58% on average.
—To validate the usefulness of our coMobile, we design, implement, and evaluate a novel application where we dispatch taxi supply according to unaccounted mobility demand inferred by coMobile. The result shows that coMobile-based dispatching outperforms a dispatching based on a state-of-the-art model by 15% on average.

2 MOTIVATION

In this section, we first show the drawback of single-view models and the opportunity of multi-view models for human mobility modeling at urban scale in real time.

2.1 Drawback of Single-View Models

We give two comparisons: (i) models driven by two cellphone views and (ii) models driven by a cellphone view and a transportation view. As in Figure 1, we first compare models driven by two 1-day CDR (call detail records) data from two carriers in Shenzhen.

This kind of models driven by single-carrier data is mostly used for human mobility modeling (Isaacman et al. 2012). A point indicates a spatial unit covered by a cell tower, and an edge linking two points together indicates the mobility between them. We only show the major mobility for the clarity of the figure. As shown by the circles, we found that each model can capture some unique mobility that cannot be captured by the other Further, to provide quantitative comparison, we calculate the user density difference of two cellular networks among 496 urban regions for the morning and evening rush hours as shown by Figures 2 and 3.

We found that the difference of user densities captured two cellular networks could be as high as 7,000 users in some regions, which indicates each carrier has different business concentration in different regions, which lead to a statistical bias again some less concentrated regions.

We combine the CDR data from different carriers and obtain a model driven by combined CDR data, that is, a model driven by the cellphone view. Similarly, we combine data from different urban transportation, that is, taxi, bus, and subway, together, and then obtain a model driven by the transportation data, that is, a model driven by the transportation view. Due to different spatial granularity (for example, cellphone data give a location at cell tower levels and bus data give a location at bus station levels), we use an urban-region-based model to show captured mobility in the morning rush hour.

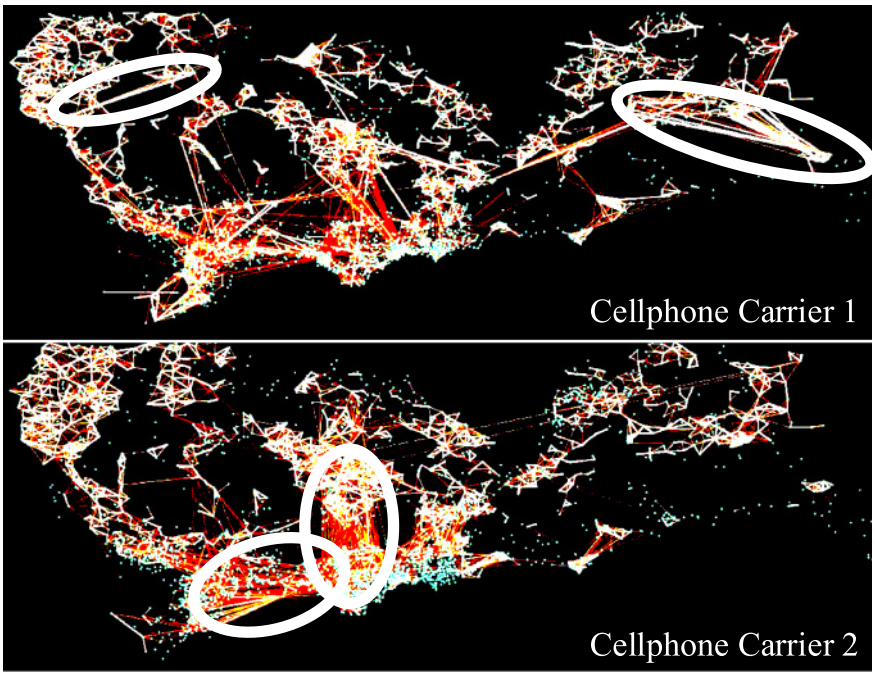

Fig. 1. Models driven by two carrier's CDR data.

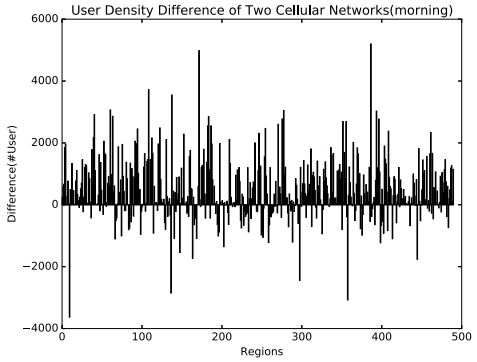

Fig. 2. Difference in morning.

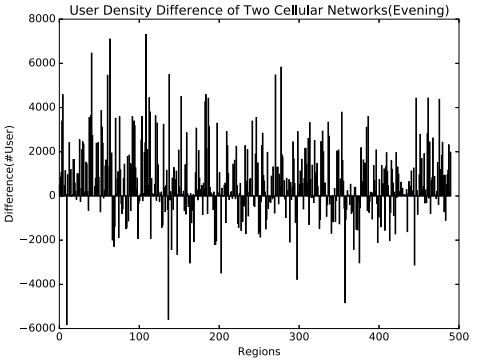

Fig. 3. Difference in evening.

As in Figure 4, every point indicates a region in the Shenzhen urban area; every edge linking two regions together indicates the mobility volume between them. The size of a vertex indicates associated mobility, and the different color indicates urban districts. As shown by the circles, we also found that each model can capture some mobility that cannot be captured by the other. In short, we conclude that the single-view human mobility models introduce a bias against the residents who are not involved in particular views. In this article, a model is considered as a biased model against the certain residents if no data about these residents are fed into this model as an input.

2.2 Opportunity of Multi-View Models

Due to the limitation of the single-view models, we are motivated to combine two separate views together to design a multi-view model for human mobility.

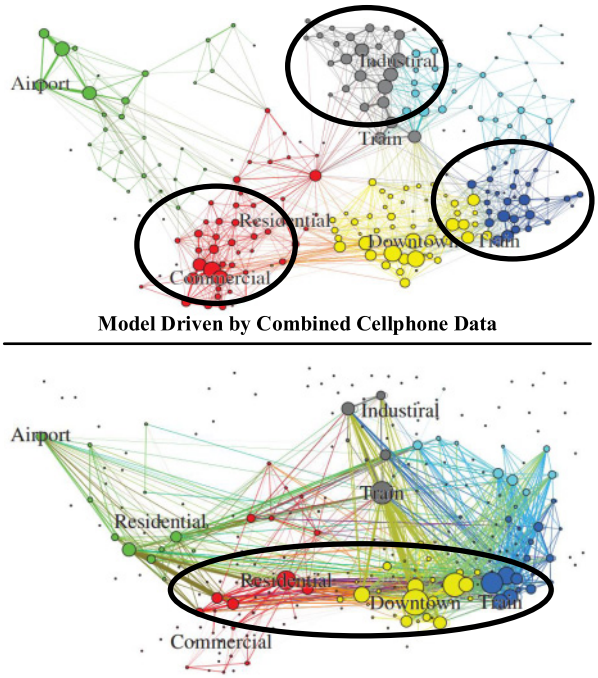

Fig. 4. Models driven by cellphone and transit data.

As shown by Figure 5, from the transportation view, we aim to combine four independent models (i.e., four triangles) driven by data from taxis, buses, subways, and private vehicles for human mobility modeling. But currently there is no urban infrastructure that can capture private transportation in real time at an urban scale. Some efforts have been made by the research community to install GPS devices in the private vehicles to study human mobility (Zheng and Xie 2011), but only limited private vehicles are involved, and using them introduces a bias that cannot be quantified.

Alternatively, we can design a model driven by cellphone CDR data as in Figure 5 from cellphone users, assuming that every resident has a cellphone. But there are two challenges.

—Some cellphone users would not use their cellphones (i.e., being inactive) during the time we perform modeling. To address this issue, we design a technique based on tensor decomposition with correlated contexts to infer locations of inactive cellphone users in Section 4.
—Some urban residents who opt out of allowing their CDR data used for other purposes or do not have cellphones at all.

Therefore, for these residents, we cannot capture their mobility. Some methods have been proposed to infer the activity of non-cellphone users based on correlated cellphone users (Jiang et al. 2013), but such a correlation cannot be obtained at an urban scale in real time.

As a result, neither the transportation view nor the cellphone view is complete by itself, but one view is often complementary to another. For example, the model driven by cellphone data can provide some mobility about residents using private transportation; whereas the model driven by transportation data can provide some mobility about residents without cellphone CDR data. It motivates us to design an effective modeling technique to combine these two views for better mobility modeling.

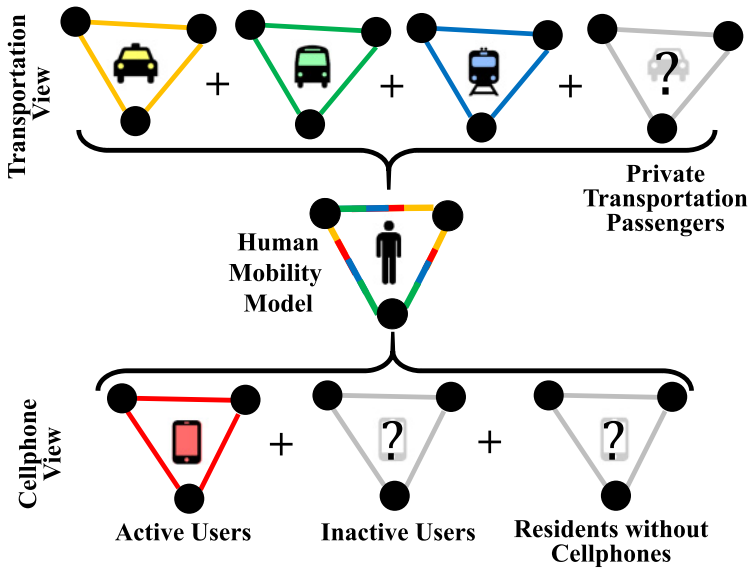

Fig. 5. Multi-view modeling.

3 PRELIMINARY

In this section, we first introduce the data we collected for multi-view modeling, and then we present a concept called mobility graph to capture the real-time human mobility, and, finally, we give the architecture of coMobile.

3.1 Multi-View Data

We have been working with several service providers and the Shenzhen Transport Committee (hereafter STC) for data access of urban infrastructures. We consider two kinds of data, that is, cellphone data and transportation data, as two individual views to model human mobility. A summary of these data is given by Figure 6. The heat map of their spatial granularity is given by Figure 7 with an area of 14×5 km^2.

Cellphone View: Cellphone CDR (call detail records) data are used to infer cellphone users' locations at cell tower levels. We utilize CDR data through two major operators in Shenzhen with more than 10 million users. The CDR data give 220 million locations per day.

Transportation View: Data from three kinds of transportation modes, that is, taxi, subway, and bus, are used to detect transportation passengers' locations. We study transportation data through STC to which taxicab, bus, and smartcard companies upload their data in real time.

> ***Taxi Data.*** They are used to infer taxi passengers' origins and destinations based on status (i.e., pickups and dropoffs) at GPS location levels. They account for 14 thousand taxis, each of which generates 2 records/min.
> ***Smartcard Data.*** They are used to infer origins and destinations of residents with smartcards used to pay bus and subway fares, which capture more than 10 million rides and 6 million passengers per day. In particular, there are two kinds of smartcard readers: (i) a total of 14,270 onboard mobile readers in 13 thousand buses capturing 168 thousand bus passengers per hour, and (ii) a total of 2,570 fixed readers in 127 subway stations capturing

Cellphone Dataset		Taxicab Dataset	
Beginning	10/1/2013	Beginning	1/1/2012
# of Users	10,432,246	# of Taxis	14,453
Size	1 TB	Size	1.7 TB
# of Records	19 billion	# of Records	22 billion
Format		Format	
SIM ID	Date&Time	Plate ID	Date&Time
Cell Tower ID	Activities	Status	GPS&Speed
Bus Dataset		**Smartcard Dataset**	
Beginning	1/1/2013	Beginning	7/1/2011
# of Buses	13,032	# of Cards	16,000,000
Size	720 GB	Size	600 GB
# of Records	9 billion	# of Records	6 billion
Format		Format	
Plate ID	Date&Time	Card ID	Date&Time
Stop ID	GPS&Speed	Device ID	Station ID

Fig. 6. Cellphone and transportation data.

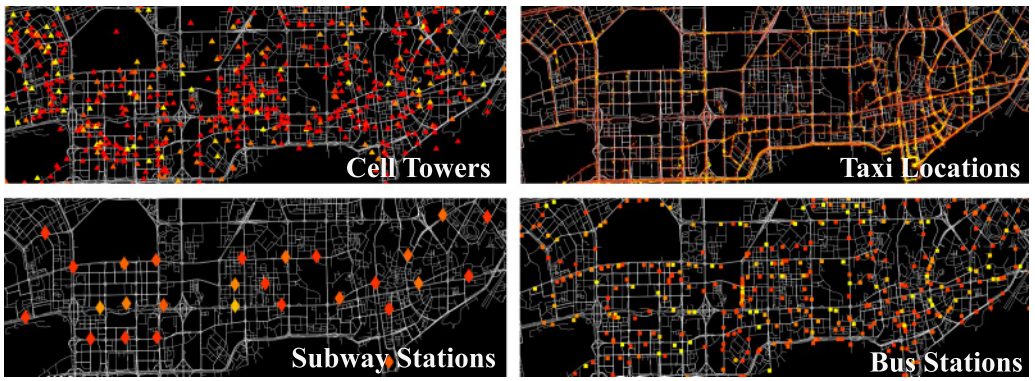

Fig. 7. Data spatial granularity.

60 thousand subway passengers per hour. Smartcard data and subway map data are used together to detect subway passengers' origins and destinations at subway station levels.

Bus Data. They are used to infer bus passengers' origins and destinations along with smartcard data (showing that a passenger uses a smartcard at a bus station) at 4,849 bus station levels. They account for all 13 thousand buses, each of which generates 2 records/min.

Our endeavor of consolidating the above data enables extremely large-scale real-time urban phenomenon rendering, for example, human mobility, which is unprecedented in terms of both quantity and quality.

In this article, we use the concept view to model the mobility from a particular urban domain, for example, a cellphone network can be used as a view for modeling because a cellphone network provides potential complete coverage for all urban residents. But a transportation mode, for example, a taxi, cannot be used as a view in the setting of this article because a taxi conceptually cannot

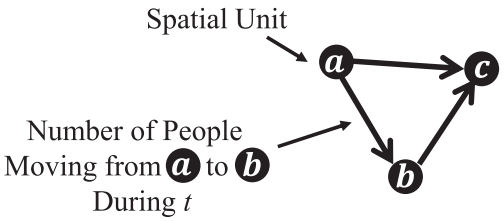

Fig. 8. Mobility graph.

cover all residents mobility as many residents would use bus and subway systems. That is why we add taxi, bus, and subway together as a single view for transportation instead of using each one transportation model as a view.

3.2 Mobility Graph

In this work, we use Mobility Graph to capture human mobility in real time at urban scale, which is a time-varying graph where a vertex indicates a spatial unit (e.g., a urban region or a street block) and a weight of an edge linking two vertices indicates the mobility volume between them. Due to its time-varying nature, a mobility graph G_t is associated with a time period t (e.g., 4–5PM), which shows the mobility during this particular time period. Figure 8 gives a simplified example of a mobility graph with only three vertices. The number of people moving between different spatial units, that is, weights of edges, should include people associated with a particular view, for example, the cellphone view or the transportation view. In this work, our main objective is to obtain mobility graphs based on single-view modeling and then to combine them together by multi-view modeling for a comprehensive human mobility graph.

3.3 coMobile Framework

We introduce our coMobile Framework by Figure 9. From the bottom, we have urban data generated by urban infrastructures, for example, cellphone data and transportation data, which are introduced in the previous subsection. Based these two kinds of data, we design two single-view models capturing mobility patterns of cellphone users and urban transportation users by two mobility graphs, which are introduced in Section 4. Then, we present our multi-view learning to integrate single-view models for more complete human mobility modeling, which is introduced in Section 5 and evaluated in Section 6. Finally, in Section 7, the obtained model is used in a real-world application where we dispatch taxis to meet unaccounted human mobility demand inferred by coMobile.

Note that we only consider two specific views in coMobile but it can be generalized to more views if more data are available. In coMobile, we first generate single-view models and then combine them together at model levels, instead of raw data levels (e.g., using multi-source raw data to directly design a multi-view model). This is because in many applications due to privacy issues, raw data are not available, and only high-level single models can be used as input. Our coMobile is still applicable to this situation.

4 SINGLE-VIEW MOBILITY MODELING

We introduce how to model urban mobility based on two single views, that is, a cellphone view and a transportation view.

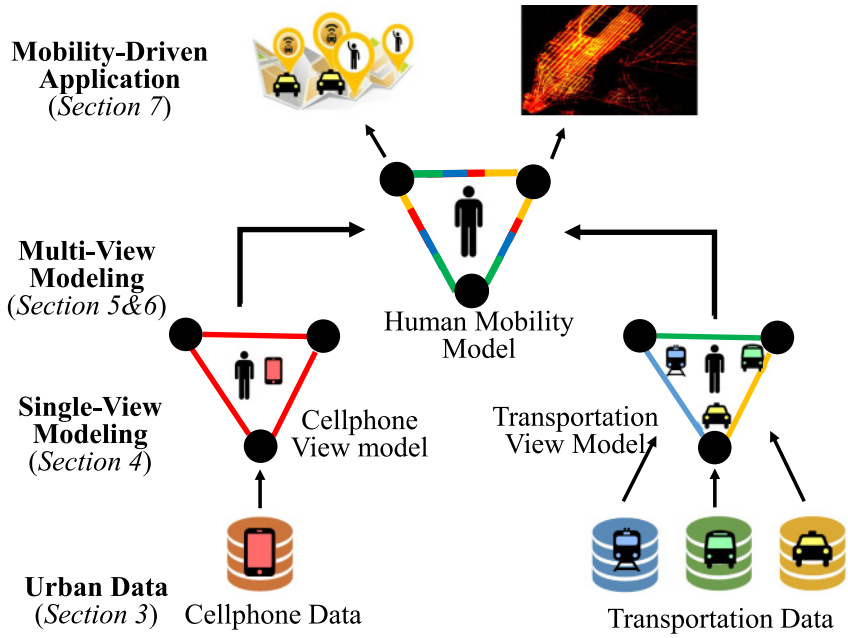

Fig. 9. coMobile framework.

4.1 Cellphone-View Modeling

As introduced earlier, the key challenge to model human mobility based on cellphone data is that inactive cellphone users or residents without cellphones do not generate any CDR data. As a result, we cannot model their mobility to obtain mobility graph. For residents without cellphones, the solution is limited although the model based on transportation can capture some of them. In this subsection, we focus on inactive cellphone users to infer their mobility by an observation that inactive users who did not use their cellphones today may use their cellphones before during the similar trips (Gonzalez et al. 2008). Thus, we formulate a tensor decomposition problem to infer mobility of active and inactive users with real-time and historical data.

4.1.1 Tensor Construction. We infer locations of cellphone users for specific time slots by a three dimensional tensor $\mathcal{A} \in \mathbb{R}^{N \times K \times M}$.

- —A cellphone user dimension indicates individual cellphone users differentiated by SIM IDs: $[u_1, \ldots, u_N]$.
- —A time slot dimension indicates specific time windows (e.g., 1h window from 5PM to 6PM): $[t_1, \ldots, t_K]$.
- —A spatial unit dimension indicates specific spatial units (e.g., a urban region): $[s_1, \ldots, s_M]$.
- —An entry $\mathcal{A}(n, k, m)$ indicates the number of CDR records a user n has in a spatial unit m during a slot k.

With our cellphone data, we fill this tensor $\mathcal{A}$, and then obtain all cellphone users' locations with a specific spatiotemporal partition. However, a key challenge is that the tensor $\mathcal{A}$ is sparse because for inactive cellphone users, their corresponding entries are empty due to lacking CDR data.

A common approach to address this issue is to use tensor decomposition. As in Figure 10, we have a tensor with three dimensions indicating cellphone users, spatial units, and time slots. An

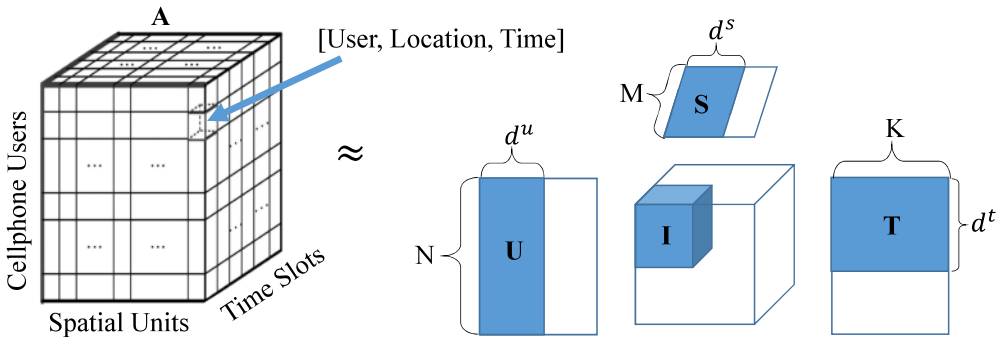

Fig. 10. Tensor decomposition.

entry denotes a tuple [user, location, time]. But this tensor is sparse due to inactive cellphone users. Based on the classic Tucker decomposition model (Kolda and Bader 2009), we decompose $\mathcal{A}$ into a core tensor $\mathcal{I}$ along with three matrices, $\mathcal{U} \in \mathbb{R}^{N \times d^u}$, $\mathcal{S} \in \mathbb{R}^{M \times d^s}$, and $\mathcal{T} \in \mathbb{R}^{K \times d^t}$. $\mathcal{U}$, $\mathcal{S}$, and $\mathcal{T}$ infer correlations between different cellphone users, different spatial units, and different time slots, respectively. d^u, d^s, and d^t are the number of latent factors and very small.

The following objective function is used to optimize the decomposition.

$$||\mathcal{A} - \mathcal{I} \times \mathcal{U} \times \mathcal{S} \times \mathcal{T}||^2 + \lambda(||\mathcal{I}||^2 + ||\mathcal{U}||^2 + ||\mathcal{S}||^2 + ||\mathcal{T}||^2),$$

where the first term is to measure the error of decomposition and the second term is a regularization function to avoid over-fitting. $||\cdot||^2$ denotes the l_2 norm and λ is the parameter to control the contribution of the regularization function. By minimizing this objective function, we obtain the optimized $\mathcal{I}$, $\mathcal{U}$, $\mathcal{S}$, and $\mathcal{T}$ by the sparse tensor $\mathcal{A}$, which is given by cellphone data. As a result, we use $\mathcal{I} \times \mathcal{U} \times \mathcal{S} \times \mathcal{T} = \mathcal{A}$' to approximate $\mathcal{A}$ where $\times$ is the tensor-matrix multiplication.

However, a key challenge for this decomposition is that $\mathcal{A}$ is over-sparse especially under fine spatiotemporal partition, which leads to poor performance of decomposition. To address this issue, in this work, we propose a technique to use historical cellphone data to establish correlated contexts that improve the performance of the decomposition.

4.1.2 Context Extraction. To provide additional information for the decomposition, we use the historical cellphone data to extract three contexts, that is, cellphone user density, calling location pattern, and calling time pattern. We use three matrices to denote these three contexts as in Figure 11.

- —Cellphone User Densities are given by a matrix $\mathcal{B}$ where a row denotes a spatial unit; a column denotes a time slot; an entry denotes the average CDR record count in this spatial unit for this time slot over a period of historical time.
- —Calling Location Patterns are given by a matrix $\mathcal{C}$ where a row denotes a spatial unit; a column denotes a cellphone user; an entry denotes a cellphone user's CDR record count in this spatial unit given a period of historical time.
- —Calling Time Patterns are given by a matrix $\mathcal{D}$ where a row denotes a time slot; a column denotes a cellphone user; an entry denotes a cellphone user's CDR record count in this time slot given a period of historical time.

All the matrices $\mathcal{B}$, $\mathcal{C}$, and $\mathcal{D}$ can be obtained by a set of historical cellphone data.

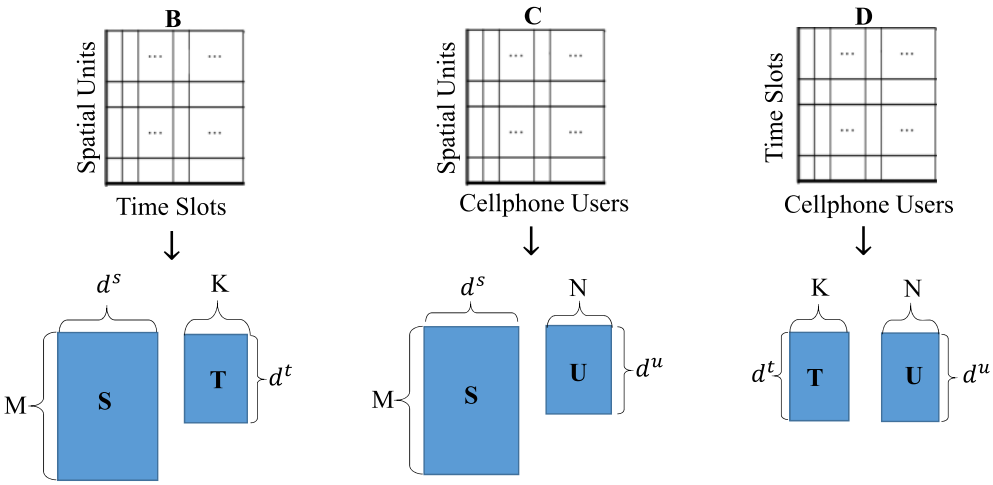

Fig. 11. Context matrix factorization.

4.1.3 Context-based Tensor Decomposition. We present a joint tensor decomposition based on the three extracted context matrices. In particular, we design the objective function as follows:

$$\begin{aligned} \min_{\mathcal{I},\mathcal{U},\mathcal{S},\mathcal{T}} \mathbf{L}(\mathcal{I},\mathcal{U},\mathcal{S},\mathcal{T}) &= ||\mathcal{A} - \mathcal{I} \times \mathcal{U} \times \mathcal{S} \times \mathcal{T}||^2 \\ +\lambda_1||\mathcal{B} - \mathcal{S} \times \mathcal{T}||^2 + \lambda_2||\mathcal{C} - \mathcal{S} \times \mathcal{U}||^2 + \lambda_3||\mathcal{D} - \mathcal{T}^T \times \mathcal{U}||^2 \\ +\lambda_4(||\mathcal{I}||^2 + ||\mathcal{U}||^2 + ||\mathcal{S}||^2 + ||\mathcal{T}||^2), \end{aligned} \tag{1}$$

where the first term is to measure the error of decomposing $\mathcal{A}$; the second, third, and fourth terms are to measure the error of factorizing matrix $\mathcal{B}$, $\mathcal{C}$, and $\mathcal{D}$, respectively; the last term is to avoid over-fitting. In our setting, $d^u = d^s = d^t$. λ_1, λ_2, λ_3, and λ_4 are preset parameters to indicate term weights. We normalized all values to [0, 1] for the decomposition.

In this objective function, $\mathcal{A}$ and $\mathcal{B}$ share $\mathcal{S}$ and $\mathcal{T}$; $\mathcal{A}$ and $\mathcal{C}$ share $\mathcal{S}$ and $\mathcal{U}$; $\mathcal{A}$ and $\mathcal{D}$ share $\mathcal{U}$ and $\mathcal{T}$. Since $\mathcal{B}$, $\mathcal{C}$, and $\mathcal{D}$ are not sparse, they lead to accurate $\mathcal{S}$, $\mathcal{T}$, and $\mathcal{U}$, which increases the performance of decomposing $\mathcal{A}$. As a result, the historical cellphone user calling patterns are transferred into the decomposition of $\mathcal{A}$, which leads to an accurate tensor decomposition.

Because this objective function does not have a closed-form solution to find the global optimal $\mathcal{I}$, $\mathcal{U}$, $\mathcal{S}$, and $\mathcal{T}$, we use an element-wise optimization algorithm as a numeric method (Karatzoglou et al. 2010) to obtain a local optimal solution. Finally, after we obtain $\mathcal{I}$, $\mathcal{U}$, $\mathcal{S}$, and $\mathcal{T}$, we use $\mathcal{I} \times \mathcal{U} \times \mathcal{S} \times \mathcal{T} = \mathcal{A}'$ to obtain cellphone mobility graph G^C of all cellphone users.

4.2 Transportation-View Modeling

Based on our transportation data, we model human mobility by three transportation modes, that is, taxi, bus, and subway. Given attributes of our transportation data, we directly obtain origins and destinations of taxi, bus, and subway passengers at GPS, bus station, and subway station levels. In this work, we use a space alignment technique where we assign taxi GPS locations, bus stations, and subway stations into corresponding spatial units based on a specific spatial partition of urban areas. Thus, for a pair of spatial units, for example, from an airport to a train station, we aggregate all the above passengers who traveled between these two spatial units to obtain a mobility volume during a particular time period, because these three kinds of transportation modes are independent

of each other. Thus, from the transportation view, obtaining transportation mobility graph G^T is straightforward.

Our context-aware tensor decomposition can also be used to improve completeness of the transportation-view model since we have missing data issues (e.g., GPS records) as well. The process is conceptually similar to the tensor decomposition for the cellphone-view model, which is omitted due to the space limitation.

Further, we did not consider private vehicles in our transportation view due to lack of private vehicle data. However, some urban residents using private transportation would be captured by multi-view learning, which is introduced as follows.

5 MULTI-VIEW MOBILITY MODELING

In this section, based on single-view modeling, we introduce multi-view modeling in coMobile. Even though our data can only form two views to obtain two mobility graphs, that is, the cellphone mobility graph G^C and the transportation mobility graph G^T, we aim to tackle a more generic problem, that is, multi-view modeling, and thus double-view modeling is a concrete example of multi-view modeling. We first formulate a joint optimization problem for multi-view human mobility modeling, and then we develop an iterative learning processing to solve this problem, and, finally, we theoretically analyze the performance of modeling in terms of convexity and convergence.

5.1 Joint Optimization

The key notations we used in this section based on graph G with different spatiotemporal effects, for example, from a spatial unit a (e.g., an airport) to b (e.g., a train station) during a time period t (e.g., 4–5PM). The main objective of our multi-view modeling is to obtain a comprehensive human mobility graph G^H for a given time period based on several single-view mobility graphs, for example, G^C and G^T. Because we have the same spatial partition for different mobility graphs, they have the same number of edges and vertices, and the key difference is edge weights. Since different edges are independent in a human mobility graph, we use one edge ab in a human mobility graph G^H as an example to show how we obtain the human mobility from one spatial unit a to another spatial unit b by our multi-view technique and combine different edge weights together to obtain a complete human mobility graph G^H.

For a specific edge ab in G^H, the volume of passengers traveling from a spatial unit a (e.g., an airport) to b (e.g., a train station) during a time period t (e.g., 4–5PM) is $x^*_{ab\cdot t}$, which is the unknown ground truth we want to infer. Assuming we have V different views, this leads to V different mobility graphs that are incomplete by themselves yet complementary to each other. For a specific view $v \in [1, V]$, we use $x^v_{ab\cdot t}$ to indicate the weight of the edge ab during t in the mobility graph G^v; for a specific view $v \in [1, V]$, we use $w^v_{ab\cdot t}$ to indicate the completeness degree of this view during a time period t from this edge ab of G^v. The completeness degree of a view quantifies its capability to capture human mobility. The stronger the capability, the higher the degree. Under different spatiotemporal contexts, the completeness degree of the same view is different. We use a vector $\mathcal{W}_{ab\cdot t} = \{w^1_{ab\cdot t}, \ldots, w^v_{ab\cdot t}, \ldots, w^V_{ab\cdot t}\}$ to indicate completeness degrees for all V views.

In coMobile, based on the above definitions, V and $x^v_{ab\cdot t}$ are given in advance by the datasets, whereas $x^*_{ab\cdot t}$ and $\mathcal{W}_{ab\cdot t}$ are unknown. Therefore, we present a joint optimization to obtain optimal $x^*_{ab\cdot t}$ and $\mathcal{W}_{ab\cdot t}$ together. The basic idea behind our multi-view learning is that a view with a higher completeness degree provides more comprehensive information, so the ground truth should be close to mobility observed by a view with a higher completeness degree. As a result, we should minimize the deviation from mobility observed by a view v to the ground truth $x^*_{ab\cdot t}$ (unknown), proportionally to its completeness degree $w^v_{ab\cdot t}$ (also unknown). Therefore, we

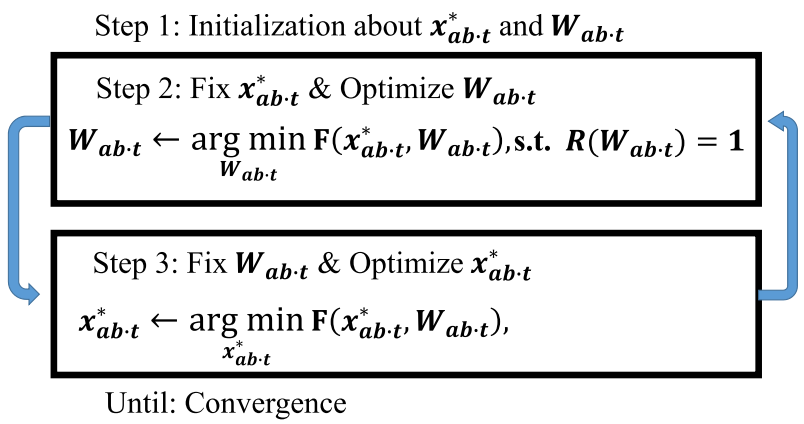

Fig. 12. Iterative multi-view learning.

develop the following objective function for multi-view learning:

$$\min_{x^*_{ab\cdot t}, \mathcal{W}_{ab\cdot t}} \mathbf{F}(x^*_{ab\cdot t}, \mathcal{W}_{ab\cdot t}) = \sum_{v=1}^{V} [w^v_{ab\cdot t} \cdot \mathbf{D}(x^*_{ab\cdot t}, x^v_{ab\cdot t})], \quad \text{s.t.}, \mathbf{R}(\mathcal{W}_{ab\cdot t}) = 1. \tag{2}$$

$\mathbf{D}(x^*_{ab\cdot t}, x^v_{ab\cdot t})$ is a distance function that describes the distance between $x^*_{ab\cdot t}$ and $x^v_{ab\cdot t}$. Therefore, the term $\sum_{v=1}^{V}[w^v_{ab\cdot t} \cdot \mathbf{D}(x^*_{ab\cdot t}, x^v_{ab\cdot t})]$ indicates the overall weighted distance between the observed mobility and the ground truth. We aim to find the optimal $x^*_{ab\cdot t}$ and $\mathcal{W}_{ab\cdot t}$ that minimize this overall weighted distance under a constraint.

$\mathbf{R}(\mathcal{W}_{ab\cdot t})$ is a constraint function, which gives the distribution of view completeness. Without this constraint, the optimization problem is unbounded. For the sake of simplicity, we set $\mathbf{R}(\mathcal{W}_{ab\cdot t}) = 1$. Other constraint functions can also be used since we can divide $\mathbf{R}(\mathcal{W}_{ab\cdot t})$ by a constant.

The rationale behind this function is that for a more-complete view, we have a high penalty if the mobility observed from this view has a longer distance to ground truth. In contrast, for a less-complete view, we have a low penalty if the mobility observed from this view has a longer distance to ground truth. Thus to minimize the objective function, ground truth relies on the more complete views.

5.2 Iterative Learning

We develop an iterative learning technique based on the block coordinate descent (Bertsekas 1999) to solve this optimization. Since in our objective function we have two sets of variables, that is, both the ground truth $x^*_{ab\cdot t}$ and the view completeness degree $\mathcal{W}_{ab\cdot t}$, we aim to iteratively yet alternatively optimize these two sets of variables until the result converges. In particular, we optimize the value of one set to minimize the objective function while keeping the value of the other set fixed, and then we swap the fixed variable and the optimized variable to continue this process until the result converges. Figure 12 gives the description of our iterative technique. In the above algorithm, we have three key steps:

—In Step 1, we first initialize $x^*_{ab\cdot t}$ and $\mathcal{W}_{ab\cdot t}$ based on the average value of $x^*_{ab\cdot t}$, because the initialization does not affect the final results based on the property of the block coordinate descent (Bertsekas 1999).
—In Step 2, we first fix the initialized $x^*_{ab\cdot t}$, and then find the optimal $\mathcal{W}_{ab\cdot t}$ that minimizes the objective function.
—In Step 3, with this optimized $\mathcal{W}_{ab\cdot t}$, we fix it and then find the optimal $x^*_{ab\cdot t}$ that minimizes the objective function again.
—Then, with this optimized $x^*_{ab\cdot t}$, we go back to Step 2 to fix $x^*_{ab\cdot t}$ again, and then to further optimize $\mathcal{W}_{ab\cdot t}$.

This is an iterative process to alternatively optimize $x^*_{ab\cdot t}$ and $W_{ab\cdot t}$ until the result converges.

Based on the property of the block coordinate descent (Bertsekas 1999), the convergence of the above iterative process is based on the distance function and constraint function used. As follows, we theoretically analyze the performance of our technique in terms of convexity and convergence.

5.3 Theoretical Analyses

We use Negative Log Function as our constraint function:

$$\mathbf{R}(\mathcal{W}_{ab\cdot t} = \{w^1_{ab\cdot t}, \ldots, w^v_{ab\cdot t}, ldots, w^V_{ab\cdot t}\}) = \sum_{v=1}^{V} \exp(-w^v_{ab\cdot t}).$$

This negative log function maps a number between 0 and 1 to a number from 0 to ∞, which enlarges the difference between different view completeness degrees for better modeling.

Further, we use normalized squared loss function as our distance function given as

$$\mathbf{D}(x^*_{ab\cdot t}, x^v_{ab\cdot t}) = \frac{(x^*_{ab\cdot t} - x^v_{ab\cdot t})^2}{\mathrm{STD}(x^1_{ab\cdot t}, \ldots, x^v_{ab\cdot t}, \ldots, x^V_{ab\cdot t})}.$$

This normalized squared loss is an effective method to measure the distance between two variables and consider the distribution of $x^v_{ab\cdot t}$ at the same time.

As follows, we prove the convexity and convergence of our iterative learning with the above two functions.

Theorem. *If the negative log function and the normalized squared loss function are used, then convergence of our iterative process in Figure 12 is guaranteed.*

Proof. Based on the convergence proposition on the block coordinate descent (Bertsekas 1999), the iterative process converges to a stationary point, if the optimizations in Steps 2 and 3 are convex. Thus, the rest of our proof has 2 steps: (i) in Step 2, if $x^*_{ab\cdot t}$ is fixed, then the optimization for $\mathcal{W}_{ab\cdot t}$ is convex; (ii) in Step 3, if $\mathcal{W}_{ab\cdot t}$ is fixed, then the optimization for $x^*_{ab\cdot t}$ is convex.

To prove the convexity of Step 2, we use another variable $y_v = \exp(-w^v)$. Therefore, the optimization problem becomes a new function with only one variable of y_v,

$$\min_{y_1,\ldots,y_v,\ldots,y_V} \mathbf{F}(y_1, \ldots, y_v, \ldots, y_V) = \sum_{v=1}^{V} [-\log(y_v) \cdot \mathbf{D}(x^*_{ab\cdot t}, x^v_{ab\cdot t})],$$

$$\text{s.t.,} \sum_{v=1}^{V} y_v = 1.$$

With this new variable y_v, we have a linear constraint function and a linear objective function (i.e., a linear combination of negative logarithm functions). Therefore, both the constraint function and

objective function are convex, which leads to the fact that any local optimal solution is also the global optimal solution for Step 2.

To prove the convexity of Step 3, we treat the objective function as an unconstrained optimization with only one variable. In Step 3, since the normalized squared loss function is convex, the objective function is a linear combination of convex functions, which makes it convex. □

Note that other constraint and distance functions can also be used in our iterative process but may not lead to the convexity of the optimization problem, and thus the convergence of the iterative process cannot be guaranteed.

6 IMPLEMENTATION AND EVALUATION

In this section, we first introduce our implementation of coMobile based on data from the Chinese city Shenzhen. Then, we present our evaluation by comparing coMobile and a single-view model to the ground truth.

6.1 coMobile Implementation

We implement coMobile based on cellphone and transportation data in Shenzhen introduced in Section 3. Since our article concentrates on modeling, we briefly introduce our data-related issues during our implementation. We establish a secure and reliable transmission mechanism, which feeds our server the data collected by STC and service providers with a wired connection. As shown in Section 3, we have been storing a large amount of data, requiring significant efforts for the daily management. We utilize a 34TB Hadoop Distributed File System (HDFS) on a cluster consisting of 11 nodes, and each of them is equipped with 32 cores and 32GB RAM. For daily management, we use the MapReduce based Pig and Hive. Right now, we are migrating to the Spark platform due to its high performance compared to Hadoop. Because of the extremely large size of our data, we have been finding several kinds of errant data, for example, duplicated data, missing data, and data with logical errors. To address these issues, we conduct a detailed cleaning process to filter out errant data. We implement our tensor decomposition technique by HaTen2 (Jeon et al. 2015), a scalable distributed suite of tensor decomposition algorithms running on the MapReduce platform. By carefully reordering the operations, and exploiting the sparsity of formulated tensors, we can dramatically reduce the intermediate data and the number of jobs for real-world implementation.

For real-world implementation, we have to decide the spatiotemporal partition for the mobility graph, which decides the spatiotemporal granularity of our model. For example, we have more than 110 thousand road segments, 496 urban regions, and 10 urban districts in Shenzhen, and we can capture the mobility with one of those three spatial partitions for every 15min, 30min, 60min, or even longer. Due to the spatial resolutions of our data (especially for bus, subway, and cellphones), we use a urban-region partition proposed by Shenzhen government as our spatial partition, which is given by Figure 13. Different colors indicate different population density. Based on this partition, we implement our multi-view mobility modeling technique coMobile based on two views. A human mobility graph obtained by coMobile for major urban areas during the evening rush hour at region levels is given by the left of Figure 14. We protect the privacy of residents by anonymizing all data and presenting models in aggregation, and only processing information related to mobility.

6.2 coMobile Evaluation

In this subsection, we introduce our evaluation about coMobile in terms of methodology and results.

6.2.1 Evaluation Methodology. Based on our implementation, we compare coMobile with a single-view human mobility model called Work and Home Extracted REgions (WHERE).

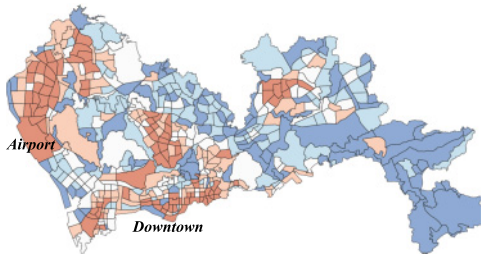

Fig. 13. Shenzhen urban partition.

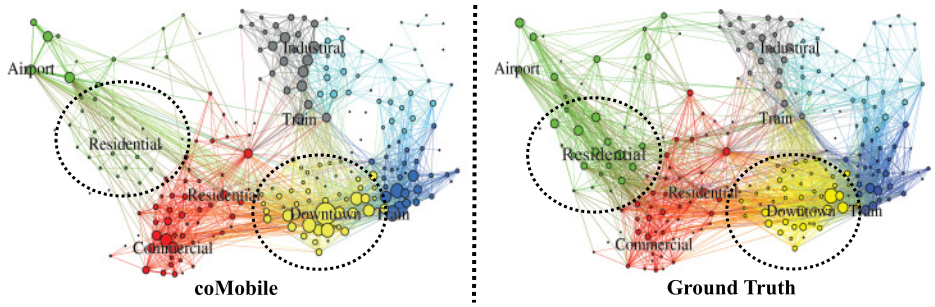

Fig. 14. Mobility graphs in Shenzhen.

WHERE (Isaacman et al. 2012) is a model driven by cellphone data, and it is based on spatial and temporal probability distributions of human mobility and produces synthetic cellphone records as the inferred mobility. We also implement a single-view model based on transportation data alone called Transit where we only feed transportation data to our coMobile. To test our tensor decomposition of single-view model improvement, we implement a model WHERE+ by inferring the missing data in the WHERE model. We compare these four models with the inferred ground truth. In this project, to infer the ground truth, we introduce another new cellphone related dataset for the evaluation. Different from regular CDR data, this dataset logs locations of all cellphone users at cell tower levels for every 15min even without activities. We use the mobility graph obtained from this dataset as the ground truth, which is given in Figure 14 given a urban partition. In this article, we use an urban region partition proposal by the Shenzhen government, and then map all the data from cellphone network and transportation systems into regions defined by this partition and finally obtain the aggregated mobility between regions based on our model to get the evaluation results. By a visual comparison, we found that we underestimate the mobility at residential areas and overestimate the mobility at downtown areas.

We utilize 3 months of data to evaluate these two models. We use Mean Average Percent Error (MAPE) in a time slot as a metric to test those two models MAPE $= \frac{100}{n}\sum_{i=1}^{n}\frac{|\bar{\mathbf{T}}_i - \mathbf{T}_i|}{\bar{\mathbf{T}}_i}$, where $n = 496 \times 496 = 246016$ is the total number of region pairs, that is, the total number of edges in a mobility graph; $\mathbf{T}_i$ is the inferred mobility between a region pair i; $\bar{\mathbf{T}}_i$ is the ground truth of the mobility between a region pair i. An accurate model yields a small MAPE and vice versa. We use 90 days of data, leading to 90 experiments. The average results were reported.

We investigate the impact of different contexts by adjusting three model parameters, that is, λ_1, λ_2, and λ_3, which control contributions of different contexts in our tensor decomposition with Equation (1). The default setting is $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = \frac{1}{4}$, where we consider all contexts and the

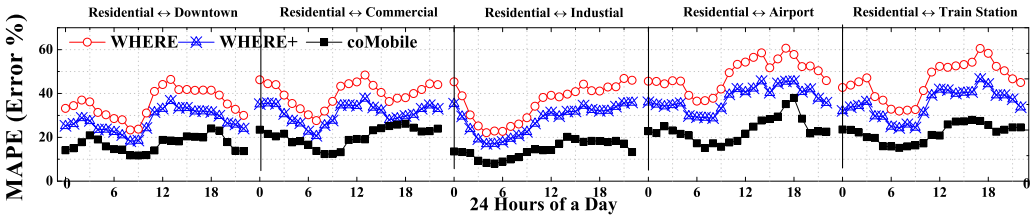

Fig. 15. MAPE under 1h slot for 24 hours of a day.

regularization term equally. Further, we investigate the impact of historical cellphone data on the model performance in terms of extracting correlated contexts.

6.2.2 Evaluation Results. We compare three models' inferring accuracy in terms of MAPE values by (i) a low-level comparison on five particular region pairs, (ii) a high-level comparison on all 246,016 region pairs, (iii) different lengths of slots, and (iv) different amount of historical data.

Figure 15 plots the MAPE under 1h slots with the two-way mobility between a residential region and five other regions. We did not implement Transit in this setting, because these are regions limited public transportation, for example, lacking subway stations. We found that coMobile outperforms WHERE and WHERE+ in general. This is because WHERE only uses the cellphone data to model the human mobility from the cellphone view alone, whereas coMobile uses two views to model the human mobility, which leads to better performance. We apply our technique to obtain a updated model WHERE+ for WHERE, which improves the performance of WHERE by 25%. But coMobile still performs the best among these two.

We also found that the performance gain between coMobile and WHERE is lower during rush hour. One of the possible explanations is that the repeatable mobility patterns are higher during rush hour, so all models have better performance. Comparing the five region pairs, we found that for the commuting region pairs (e.g., between the residential region and the industrial, commercial or downtown regions), all models have better performance than the region pairs on which the residents go for travel (i.e., between the residential region to the airport or train station regions). This is due to the fact that repeatable pattern for travel is limited.

Figure 16 gives the MAPE on all region pairs under one hour slots. We found that all three models have higher MAPE than the MAPE we observed in Figure 15. This is because the urban mobility is highly dynamic between various regions pairs, and many region pairs have very limited mobility, which leads to high MAPE. But we also found that relative performance between these three models is the same as in Figure 15. coMobile is better than WHERE and Transit, which shows the advantage of using multi-view learning to model the human mobility. coMobile outperforms WHERE and Transit by 51% and 58% in terms of MAPE, resulting from its multi-view learning from both cellphone data and transportation data. But WHERE is better than Transit because of large-scale coverage of cellphone users.

Figure 17 plots the MAPE of coMobile, Transit, and WHERE with different slot lengths from 15min to 12h. Basically, the MAPE of all three models reduces with the increase of the modeling lengths. This is because the mobility in a longer time slot is much more stable. coMobile significantly outperforms WHERE and Transit when the slot length is short. This is because the transportation data can capture lots of mobility during a short time period. We notice that the slot length becomes longer than 6h, both all three models have the similar performance, because in a long time slot, the cellphone data alone are capable of inferring mobility.

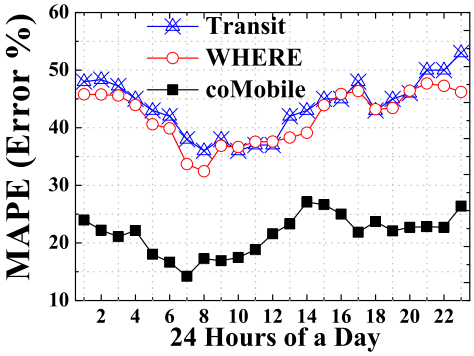

Fig. 16. Hourly MAPE.

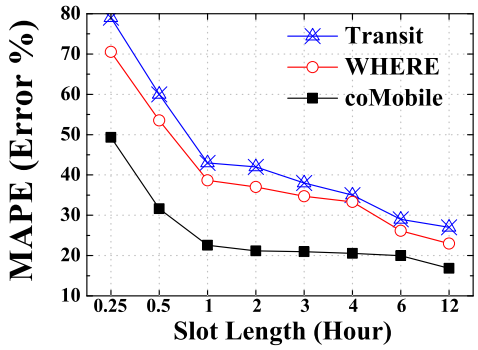

Fig. 17. Effects of lengths.

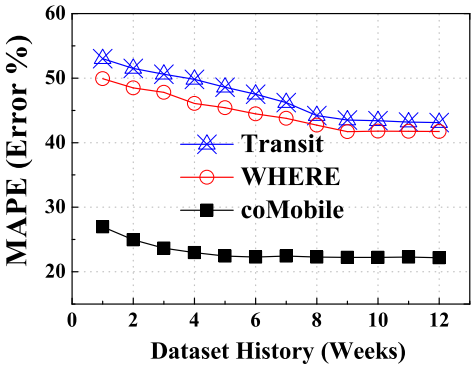

Fig. 18. Historical data.

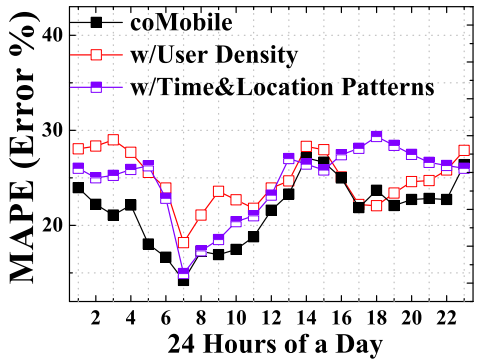

Fig. 19. Contexts.

Figure 18 shows how much historical information is necessary for all three models. As expected, the longer the time, the lower the MAPE for both models, the better the performance. But a too-long history does not help much, especially for coMobile, whose MAPE became stable when the historical data are longer than 4 weeks. It shows that coMobile does not reply on long-term historical cellphone data, thanks to the transportation view. But WHERE and Transit need a longer historical period of data, that is, 9 weeks, before its MAPE becomes stable.

Figure 19 shows the impact of two contexts, that is, cellphone user densities and calling time&location patterns as introduced in Section 4.1.2. In particular, we set $\lambda_2 = \lambda_3 = 0$ and $\lambda_1 = \lambda_4 = \frac{1}{2}$ to obtain a model called coMobile w/ User Density, which only considers the cellphone user density as a context. Similarly, we set $\lambda_1 = 0$ and $\lambda_2 = \lambda_3 = \lambda_4 = \frac{1}{3}$ to obtain a model called coMobile w/ Time&Location Patterns, which only considers time&location patterns as contexts. We compare them with coMobile, which considers all contexts. In generally, coMobile outperforms the other two models. We found that for the early morning, considering time&location patterns is better than considering user density; while for the late night, considering user density is better than considering time&location patterns. Also, during some slots in the afternoon or evening, for example, 14:00, 15:00, and 18:00, it leads to better performance if we do not consider certain contexts.

In short, we have the following observations. (i) As in Figure 15, the accuracy of human mobility modeling is highly depended on both locations and time of day. Our tensor decomposition based

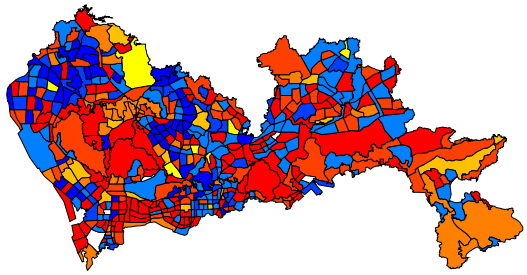

Fig. 20. Mobility demand vs. taxi supply.

single-view improvement, that is, WHERE+, works well but still cannot outperform coMobile because of its complementary data sources. (ii) As in Figure 16, all models have better performance in the morning rush hour in general due to the predicability of morning commutes, and coMobile outperforms WHERE and Transit during all times. (iii) As in Figure 17, the length of slots has significant impacts on performance of all models. (iv) As in Figure 18, how much historical data to be used by coMobile does not significantly affect the performance of coMobile and Transit. (v) As in Figure 19, the same contexts have different effects according to the time of day, but considering them together leads to better average performance.

7 COMOBILE APPLICATION: TAXI DISPATCHING

Built on coMobile, we introduce an mobile application where we dispatch vacant taxis among urban regions with high human mobility demand yet low transit supply in rea time for better urban transportation services.

7.1 Application Overview

Our overall dispatching objective is to achieve an equilibrium between unaccounted mobility demand and taxi supply at urban region level. The unaccounted mobility demand can be obtained by general mobility inferred by coMobile minus urban public transit demand. Thus, the unaccounted mobility demand (including demand for all other mode of transport except bus and subway) indicates the potential demand for taxis. The taxi supply, that is, the number of vacant taxis at different urban regions, can be obtained straightforwardly based on real-time taxi GPS data. Our taxi dispatching strategy aims to balance the relationship between unaccounted mobility demand and taxi supply to potentially improve the mobility efficiency at urban region levels. As shown in by Figure 20, we show a visualization of unaccounted mobility demand and taxi supply where the warmer the color, the higher the mobility demand, the lower the taxi supply.

7.2 Dispatching Strategy

In our dispatching strategy, at the end of a slot t_i, we first use coMobile to infer unaccounted mobility demand $\mathbb{D}^{r_x}_{t_{i+1}}$ for the next time slot t_{i+1} in a predefined region r_x. In particular, we first use coMobile to obtain general human mobility at urban region levels, then aggregate subway and bus passenger demand to obtain public transit demand, and, finally, we use the general human mobility minus public transit demand as unaccounted mobility demand $\mathbb{D}^{r_x}_{t_{i+1}}$. Next, we employ real-time taxi GPS data as introduced in Section 3 to aggregate vacant taxicabs to obtain "dispatchable" vacant taxi supply in region r_x for the next slot t_{i+1}, indicated by $\mathbb{S}^{r_x}_{t_{i+1}}$. Similarly, we shall have all $\mathbb{D}^{r_x}_{t_{i+1}}$ and $\mathbb{S}^{r_x}_{t_{i+1}}$ for all urban regions where $1 \leq x \leq 495$. Finally, we use a straightforward scheme to dispatch vacant taxicab supply $\sum_{1 \leq x \leq 495} \mathbb{S}^{r_x}_{t_{i+1}}$ among 495 regions. The final objective is to achieve

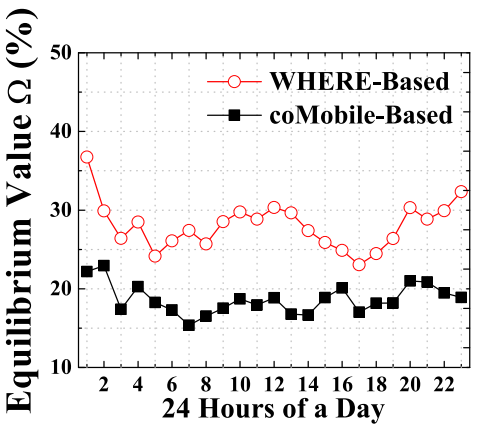

Fig. 21. Ω in 1h slots.

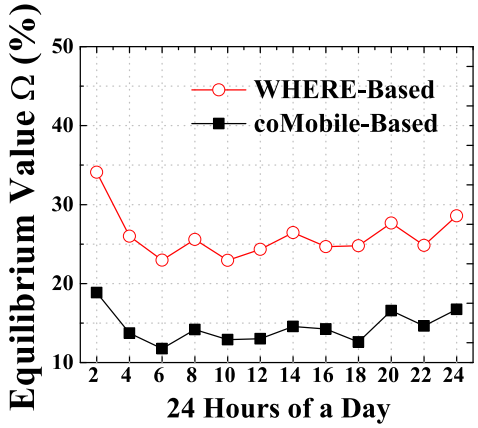

Fig. 22. Ω in 2h slots.

an equilibrium where the taxicab supply $\hat{\mathbb{S}}^{r_x}_{t_{i+1}}$ is proportional to inferred unaccounted mobility demand $\mathbb{D}^{r_x}_{t_{i+1}}$ in region r_x for a further time slot t_{i+1}.

7.3 Dispatching Evaluation

The evaluation is based on the ground truth of general mobility demand $\bar{\mathbb{D}}^{r_x}_{t_{i+1}}$ of slot t_{i+1}, and the dispatched vacant taxicab supply $\hat{\mathbb{S}}^{r_x}_{t_{i+1}}$ in each region at hourly slots. To evaluate effectiveness of dispatching, we propose a metric, that is, normalized equilibrium value $0 \leq \Omega \leq 1$, where

$$\Omega_{t_{i+1}} = \text{avg}_{1 \leq x \leq 495} \frac{|\bar{\mathbb{D}}^{r_x}_{t_{i+1}} - \hat{\mathbb{S}}^{r_x}_{t_{i+1}}|}{\bar{\mathbb{D}}^{r_x}_{t_{i+1}} + \hat{\mathbb{S}}^{r_x}_{t_{i+1}}}.$$

If the unaccounted mobility demand inferred by an inference method, for example, coMobile, is similar to the ground truth, then the corresponding dispatch will lead to a smaller $\Omega_{t_{i+1}}$, indicating an higher equilibrium between unaccounted mobility demand and taxicab supply; otherwise, it leads to a large $\Omega_{t_{i+1}}$, that is, disequilibrium. We compared two dispatching strategies: coMobile-based dispatching and WHERE-based dispatching. coMobile-based dispatching strategy uses coMobile as the human mobility model to infer unaccounted mobility demand, whereas WHERE-based dispatching strategy uses WHERE (Isaacman et al. 2012) as a human mobility model to infer unaccounted mobility demand as introduced in Section 6. Note that dispatching taxicabs would skew historical taxicabs' GPS dataset. To eliminate dispatching effects, we only used dispatched supply to calculate Ω and did not manipulate taxis' traces, and we start over at the end of the next slot. Since we did not introduce additional effects on the taxi GPS datasets, dispatching at every hour is similar to the first dispatching, which leads to accurate evaluations.

Figure 21 gives the equilibrium value Ω for different times of a day under 1h time windows. Although we found that there are some fluctuation for equilibrium values in both of dispatching strategies, coMobile-based dispatching has a lower equilibrium value at every slot compared to WHERE-based dispatching. Figure 22 gives the average equilibrium value Ω under 2h slots. In general, we found under 2h slots that the equilibrium values are lower than the equilibrium values under 1h slots for both coMobile-based dispatching and WHERE-based dispatching. There results verify our previous observation that inferring based on 2h slots is better than inferring based on 1h slots in terms of accuracy. But the relative performance in terms of equilibrium values between two

model-based dispatching strategies shown in Figure 22 is similar to the one in Figure 21. coMobile-based dispatching outperforms WHERE-based dispatching by 15% on average, because of accurate mobility inferring by coMobile.

8 RELATED WORK

Modeling the human mobility in urban scales is crucial for mobile applications, urban planning, and social networks (Zheng et al. 2014). Recently, due to the ubiquity of GPS devices and urban infrastructure upgrades, analyzing human mobility based on empirical mobility data has received significant attention (Hess et al. 2015; Isaacman et al. 2012; Gonzalez et al. 2008; Song et al. 2010; Girardin et al. 2008, 2009, 2007). Various data from urban systems can be used to infer human mobility, for example, data from smartcards (Sun et al. 2012), check-ins (Cho et al. 2011), and transport systems including subways (Lathia and Capra 2011) and taxicabs (Ge et al. 2010; Huang and Powell 2012; Yuan et al. 2011b; Zhang et al. 2011). However, almost all existing models are driven by single views, for example, the cellphone view (Isaacman et al. 2012) or the transportation view (Ganti et al. 2013). We made the first attempt to model human mobility with multi-source data (Zhang et al. 2014), but our previous work was to use transportation data to adjust the modeling process based on cellphone data and did not treat these two kinds of data equally as two views. As follows, we summarize the related work by different views.

- **Cellphone View:** Modeling from the cellphone view based on call detail records (CDR) is the most common method, for example, modeling how residents move around cities (Isaacman et al. 2012); estimating cellphone users' travel range (Isaacman et al. 2011); predicting where cellphone users will travel next (Dufková et al. 2009); identifying cellphone users' important locations, for example, work or home (Isaacman et al. 2010); and specific models for different cities, for example, Los Angles or New York City (Girardin et al. 2009), or Rome (Girardin et al. 2008). However, the models from cellphone views are mostly biased against a certain group of residents, leading to inaccurate analyses. To our knowledge, we are the first to combine data from more than one carrier to model human mobility and correlate cellphone data with transit data to address the bias issue to increase analysis accuracy.
- **Transportation View:** Transportation data are another important data source for human mobility, for example, bus data (Liu et al. 2017; Bhattacharya et al. 2013), subway data (Lathia and Capra 2011), taxicab data (Zheng et al. 2011a; Ganti et al. 2013), and private vehicle data (Giannotti et al. 2011). However, the models driven by data from one kind of transportation are mostly biased against the passengers using other transportation. To our knowledge, there is no model driven by more than one transportation mode, and we are the first to combine data from three kinds of transportation for mobility modeling. In contrast, our method is based on data from the entire set of urban transit networks correlated with data from cellular networks, instead of sampling residents using a specific transit mode.
- **Other Views:** Other data generated by urban residents have also been used to study human mobility, for example, social networks or mobile *ad hoc* networks, that is, with check-in data (Cho et al. 2011) and proximity data (Backstrom et al. 2010). However, the number of residents captured by these views is often limited compared to the cellphone data and transportation data, which leads to a bias that cannot be quantified.

In summary, almost all human mobility modeling is based on single views, which are often incomplete in terms of capturing human mobility at an urban scale in real time. Such a shortcoming motivates us to take a multi-view approach, which uses incomplete yet complementary views to model human mobility. We implement this idea by our design coMobile, which combines both the cellphone view and transportation view together to model the human mobility and outperforms

a single-view model. Conceptually, our work related to multi-view learning (Xu et al. 2013). But our key difference to traditional multi-view learning techniques, for example, co-training (Blum and Mitchell 1998), is in our tensor decomposition-based single-view improvements and the iterative optimization technique for multiple view merging introduced in Section 4 and Section 5, respectively.

Based on mobility modeling, many novel applications are proposed to improve urban mobility, for example, assisting urban residents to make transit decisions, such as taking a taxicab or not (Wu et al. 2012), inferring urban-scale maps based on vehicle GPS (Biagioni and Eriksson 2012), predicting parking locations for urban residents (Nandugudi et al. 2014), inferring arrival times within bus systems (Biagioni et al. 2011), inferring the potential passenger volume for taxi business (Ge et al. 2010), modeling urban transportation networks (Zheng et al. 2010), recommending optimal routes to pick up potential passengers (Ge et al. 2011), inferring potential locations with higher profits for taxi drivers (Powell et al. 2011), assisting new drivers with GPS data from expedited drivers (Yuan et al. 2011a), detecting anomalies for taxi business (Sen and Balan 2013), and understanding region functions in a particular city (Yuan et al. 2012). However, the above research has not focused on taxi dispatching based on general unaccounted urban mobility and typically utilizes data from one particular view instead of the multi-view as we did in coMobile.

9 DISCUSSION

We provide some discussions about coMobile as follows.

Privacy Protections. While the data for the human mobility study have the potential for great social benefits, we have to protect the privacy of the residents involved for wider applications. We took two active steps for privacy protection. (i) Anonymization: All data analyzed are anonymized by the service providers, who were not involved in this project, and all identifiable IDs, such as SIM card IDs, are replaced by a serial identifier during the analyses. (ii) Aggregation: the mobility patterns obtained by coMobile are given at aggregated results with a mobility graph in a large spatiotemporal partition. We do not focus on individual residents during the analyses.

Public Data Access. Accessing empirical datasets is vital to geographic information system research, but such datasets are usually not available to fellow researchers due to the privacy issues. As an initial step, the partial aggregated data used in this work have been made for public access in the website of Transport Committee of Shenzhen Municipality (Transport Committee of Shenzhen 2014). Most importantly, we released the first big urban dataset (Zhang et al. 2014), which includes the large-scale Shenzhen data including taxi, bus, subway, smartcard, and cellphone data. This is the first time that such comprehensive urban data have been released for the benefit of the research community.

10 CONCLUSION

In this work, we design, implement, and evaluate a human mobile modeling technique called coMobile based on context-aware tensor decomposition and iterative multi-view learning. It is the first human mobility model based on both a cellphone view and a transportation view. Our endeavors offer a few valuable insights, which may help fellow researchers to model other urban phenomena: (i) human mobility modeling based on single-view data introduces biases, which can be partially addressed by using historical data; (ii) to model human mobility, every view itself is incomplete but they are often complementary to each other, and thus it is essential to model the completeness degree of a view before inferring the mobility; and (iii) multi-view learning for human mobility requires an iterative optimization process to improve the accuracy of modeling and thus how to select an objective function and constraint function to ensure the convergence is essential for real-time applications.

REFERENCES

Javed Aslam, Sejoon Lim, Xinghao Pan, and Daniela Rus. 2012. City-scale traffic estimation from a roving sensor network. In *Proceedings of the 10th ACM Conference on Embedded Network Sensor Systems (SenSys'12)*.

Lars Backstrom, Eric Sun, and Cameron Marlow. 2010. Find me if you can: Improving geographical prediction with social and spatial proximity. In *Proceedings of the International World Wide Web Conference (WWW 10)*.

Rajesh Krishna Balan, Khoa Xuan Nguyen, and Lingxiao Jiang. 2011. Real-time trip information service for a large taxi fleet. In *Proceedings of the International Conference on Mobile Systems, Applications, and Services (MobiSys'11)*.

Dimitri P. Bertsekas. 1999. Non-linear programming. In *Athena Scientific*.

Sourav Bhattacharya, Santi Phithakkitnukoon, Petteri Nurmi, Arto Klami, Marco Veloso, and Carlos Bento. 2013. Gaussian process-based predictive modeling for bus ridership. In *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp'13)*.

James Biagioni and Jakob Eriksson. 2012. Map inference in the face of noise and disparity. In *Proceedings of the International Conference on Advances in Geographic Information Systems (SIGSPATIAL'12)*.

James Biagioni, Tomas Gerlich, Timothy Merrifield, and Jakob Eriksson. 2011. EasyTracker: Automatic transit tracking, mapping, and arrival time prediction using smartphones. In *Proceedings of the ACM Conference on Embedded Networked Sensor Systems (SenSys'11)*.

Avrim Blum and Tom Mitchell. 1998. Combining labeled and unlabeled data with co-training. In *Proceedings of the 11th Annual Conference on Computational Learning Theory (COLT'98)*. ACM, New York, NY, 92–100. DOI: http://dx.doi.org/10.1145/279943.279962

Eunjoon Cho, Seth A. Myers, and Jure Leskovec. 2011. Friendship and mobility: User movement in location-based social networks. In *Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD'11)*.

Kateřina Dufková, Jean-Yves Le Boudec, Lukáš Kencl, and Milan Bjelica. 2009. Predicting user-cell association in cellular networks. In *Second International Workshop on Mobile Entity Localization and Tracking in GPS-less Environnments (MELT'09)*.

Raghu Ganti, Mudhakar Srivatsa, Anand Ranganathan, and Jiawei Han. 2013. Inferring human mobility patterns from taxicab traces. In *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp'13)*.

Yong Ge, Chuanren Liu, Hui Xiong, and Jian Chen. 2011. A taxi business intelligence system. In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'11)*.

Yong Ge, Hui Xiong, Alexander Tuzhilin, Keli Xiao, Marco Gruteser, and Michael Pazzani. 2010. An energy-efficient mobile recommender system. In *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'10)*.

Fosca Giannotti, Mirco Nanni, Dino Pedreschi, Fabio Pinelli, Chiara Renso, Salvatore Rinzivillo, and Roberto Trasarti. 2011. Unveiling the complexity of human mobility by querying and mining massive trajectory data. *The VLDB Journal* 20, 5 (2011), 695–719.

F. Girardin, F. Calabrese, F. Dal Fiorre, A. Biderman, C. Ratti, and J. Blat. 2008. Digital footprinting: Uncovering tourists with user-generated content. *IEEE Pervasive Computing* 7, 4 (2008), 36–43.

F. Girardin, F. Dal Fiore, J. Blat, and C. Ratti. 2007. Understanding of tourist dynamics from explicitly disclosed location information. In *Proceedings of the Symposium on LBS and Telecartography*.

F. Girardin, A. Vaccari, A. Gerber, A. Biderman, and C. Ratti. 2009. Towards estimating the presence of visitors from the aggregate mobile phone network activity they generate. In *Proceedings of the Internation Conference on Computers in Urban Planning and Urban Management*.

Marta C. González, César A. Hidalgo, and Albert-László Barabási. 2008. Understanding individual human mobility patterns. *Nature* 453 (2008), 779–782.

Andrea Hess, Karin Anna Hummel, Wilfried N. Gansterer, and Günter Haring. 2015. Data-driven human mobility modeling: A survey and engineering guidance for mobile networking. *ACM Comput. Surv.* 48, 3, Article 38 (Dec. 2015), 39 pages. DOI: http://dx.doi.org/10.1145/2840722

Yan Huang and Jason W. Powell. 2012. Detecting regions of disequilibrium in taxi services under uncertainty. In *Proceedings of the 20th International Conference on Advances in Geographic Information Systems (SIGSPATIAL'12)*.

Sibren Isaacman, Richard Becker, Ramón Cáceres, Stephen Kobourov, James Rowland, and Alexander Varshavsky. 2010. A tale of two cities. In *Proceedings of the International Workshop on Mobile Computing Systems and Applications (HotMobile'10)*.

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 International Conference on Mobile Systems, Applications, and Services (MobiSys'12)*.

Sibren Isaacman, Richard A. Becker, Ramón Cáceres, Stephen G. Kobourov, Margaret Martonosi, James Rowland, and Alexander Varshavsky. 2011. Ranges of human mobility in Los Angeles and New York. In *Proceedings of the IEEE International Conference on Pervasive Computing and Communications (PerCom Workshops'11)*.

I. Jeon, E. E. Papalexakis, U. Kang, and C. Faloutsos. 2015. HaTen2: Billion-scale tensor decompositions. In *Proceedings of the 2015 IEEE 31st International Conference on Data Engineering*. 1047–1058. DOI: http://dx.doi.org/10.1109/ICDE.2015.7113355

Shan Jiang, Gaston A. Fiore, Yingxiang Yang, Joseph Ferreira, Jr., Emilio Frazzoli, and Marta C. González. 2013. A review of urban computing for mobile phone traces: Current methods, challenges and opportunities. In *Proceedings of the International Workshop on Urban Computing (UrbComp'13)*.

Alexandros Karatzoglou, Xavier Amatriain, Linas Baltrunas, and Nuria Oliver. 2010. Multiverse recommendation: N-dimensional tensor factorization for context-aware collaborative filtering. In *Proceedings of the 4th ACM Conference on Recommender Systems (RecSys'10)*. ACM, New York, NY, 79–86. DOI: http://dx.doi.org/10.1145/1864708.1864727

Tamara G. Kolda and Brett W. Bader. 2009. Tensor decompositions and applications. In *SIAM Review* 51, 3 (2009), 455–500.

Neal Lathia and Licia Capra. 2011. How smart is your smartcard?: Measuring travel behaviours, perceptions, and incentives. In *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp'11)*.

Wei Liu, Yu Zheng, Sanjay Chawla, Jing Yuan, and Xie Xing. 2011. Discovering spatio-temporal causal interactions in traffic data streams. In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'11)*.

Yanchi Liu, Chuanren Liu, Nicholas Jing Yuan, Lian Duan, Yanjie Fu, Hui Xiong, Songhua Xu, and Junjie Wu. 2017. Intelligent bus routing with heterogeneous human mobility patterns. *Knowledge and Information Systems* 50, 2 (2017), 383–415. DOI: http://dx.doi.org/10.1007/s10115-016-0948-6

Anandatirtha Nandugudi, Taeyeon Ki, Carl Nuessle, and Geoffrey Challen. 2014. PocketParker: Pocketsourcing parking lot availability. In *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing(UbiComp'14)*.

J. Powell, Y. Huang, F. Bastani, and M. Ji. 2011. Towards reducing taxicab cruising time using spatio-temporal profitability maps. In *Proceedings of the 12th International Symposium on Advances in Spatial and Temporal Databases*.

I. Rhee, M. Shin, S. Hong, K. Lee, and S. Chong. 2008. On the levywalk nature of human mobility. In *International Conference on Advanced Communications and Computation (INFOCOM'08)*.

Rijurekha Sen and Rajesh Krishna Balan. 2013. Challenges and opportunities in taxi fleet anomaly detection (*SENSEMINE'13*).

Jingbo Shang, Yu Zheng, Wenzhu Tong, Eric Chang, and Yong Yu. 2014. Inferring gas consumption and pollution emission of vehicles throughout a city. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'14)*. ACM, New York, NY, 1027–1036. DOI: http://dx.doi.org/10.1145/2623330.2623653

Filippo Simini, Marta C. González, Amos Maritan, and Albert-László Barabási. 2014. A universal model for mobility and migration patterns. *Nature* 484 (2014), 96–100.

Chaoming Song, Zehui Qu, Nicholas Blumm, and Albert-László Barabási. 2010. Limits of predictability in human mobility. *Science* 327, 5968 (2010), 1018–1021.

Lijun Sun, Der-Horng Lee, Alex Erath, and Xianfeng Huang. 2012. Using smart card data to extract passenger's spatio-temporal density and train's trajectory of MRT system. In *Proceedings of the International Workshop on Urban Computing (UrbComp'12)*.

Transport Committee of Shenzhen. 2014. Statistical data for transportation in Shenzhen. Retrieved from http://www.sz.gov.cn/jtj/tjsj/zxtjxx/.

Yilun Wang, Yu Zheng, and Yexiang Xue. 2014. Travel time estimation of a path using sparse trajectories. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'14)*. ACM, New York, NY, 25–34. DOI: http://dx.doi.org/10.1145/2623330.2623656

Ling-Yin Wei, Yu Zheng, and Wen-Chih Peng. 2012. Constructing popular routes from uncertain trajectories. In *Proceedings of the International Conference on Knowledge Discovery and Data Mining (KDD'12)*.

Wei Wu, Wee Siong Ng, Shonali Krishnaswamy, and Abhijat Sinha. 2012. To taxi or not to taxi? - Enabling personalised and real-time transportation decisions for mobile users. In *Proceedings of the 2012 IEEE 13th International Conference on Mobile Data Management (MDM'12)*.

Chang Xu, Dacheng Tao, and Chao Xu. 2013. A survey on multi-view learning. *CoRR* abs/1304.5634. http://arxiv.org/abs/1304.5634.

Jing Yuan, Yu Zheng, and Xing Xie. 2012. Discovering regions of different functions in a city using human mobility and POIs. In *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'12)*.

Jing Yuan, Yu Zheng, Xing Xie, and Guangzhong Sun. 2011a. Driving with knowledge from the physical world. In *Proceedings of the International Conference on Knowledge Discovery and Data Mining (KDD'11)*.

Jing Yuan, Yu Zheng, Liuhang Zhang, XIng Xie, and Guangzhong Sun. 2011b. Where to find my next passenger. In *Proceedings of the 13th International Conference on Ubiquitous Computing (UbiComp'11)*.

Desheng Zhang, Jun Huang, Ye Li, Fan Zhang, Chengzhong Xu, and Tian He. 2014. Exploring human mobility with multi-source data at extremely large metropolitan scales (*MobiCom'14*). 201–212. http://doi.acm.org/10.1145/2639108.2639116

Daqing Zhang, Nan Li, Zhi-Hua Zhou, Chao Chen, Lin Sun, and Shijian Li. 2011. iBAT: Detecting anomalous taxi trajectories from GPS traces. In *Proceedings of the 13th Conference on Ubiquitous Computing (UbiComp'11)*.

Desheng Zhang, Juanjuan Zhao, Fan Zhang, and Tian He. 2015b. coMobile: Real-time human mobility modeling at urban scale by multi-view learning. In *Proceedings of the ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (SIGSPATIAL'15)*.

Fuzheng Zhang, Nicholas Jing Yuan, David Wilkie, Yu Zheng, and Xing Xie. 2015a. 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. DOI: http://dx.doi.org/10.1145/2644828

Wangsheng Zhang, Shijian Li, and Gang Pan. 2012. Mining the semantics of origin-destination flows using taxi traces. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing (UbiComp'12)*.

Yu Zheng. 2015. Trajectory data mining: An overview. *ACM Trans. Intell. Syst. Technol.* 6, 3, Article 29 (May 2015), 41 pages. DOI: http://dx.doi.org/10.1145/2743025

Yu Zheng, Licia Capra, Ouri Wolfson, and Hai Yang. 2014. Urban computing: Concepts, methodologies, and applications. *ACM Trans. Intell. Syst. Technol.* 5, 3, Article 38 (Sept. 2014), 55 pages. DOI: http://dx.doi.org/10.1145/2629592

Yu Zheng, Yukun Chen, Quannan Li, Xing Xie, and Wei-Ying Ma. 2010. Understanding transportation modes based on GPS data for web applications. *ACM Trans. Web* 4, 1 (Jan. 2010).

Yu Zheng, Yanchi Liu, Jing Yuan, and Xing Xie. 2011a. Urban computing with taxicabs. *Proceedings of the 13th Conference on Ubiquitous Computing (UbiComp'11)*.

Yu Zheng and Xing Xie. 2011. Learning travel recommendations from user-generated GPS traces. *ACM Trans. Intell. Syst. Technol.* 2, 1 (2011) 2:1–2:29.

Yu Zheng, Lizhu Zhang, Zhengxin Ma, Xing Xie, and Wei-Ying Ma. 2011b. Recommending friends and locations based on individual location history. *ACM Trans. Web* 5, 1, Article 5 (Feb. 2011), 44 pages. DOI: http://dx.doi.org/10.1145/1921591.1921596

Received January 2017; revised April 2017; accepted May 2017