Metropolitan-scale People Flow Prediction

Zipei FAN[1,2], Xuan SONG[2], Xiaojie YANG[2,3] and Ryosuke SHIBASAKI[1,3,4]

[1]R&D Division, LocationMind Inc.
(PMO Kandatsukasamatchi 4F, 2-8-1, Kandatsukasamachi, Chiyoda-ku, Tokyo, 101-0048, Japan)
fan@locationmind.com
[2]Center for Spatial Information Science, University of Tokyo
(5-1-5 Kashiwanoha, Kashiwa-shi, Chiba, 277-8568, Japan)
songxuan@csis.u-tokyo.ac.jp, xiaojie_yang@csis.u-tokyo.ac.jp
Interfaculty Initiative in Information Studies, University of Tokyo
[3] (Hongo 7-3-1 Bunkyo-ku Tokyo, 113-0033, Japan)
shiba@csis.u-tokyo.ac.jp
[4]Reitaku University
(2-1-1, Hikarigaoka, Kashiwa City, Chiba Prefecture, 277-8686, Japan)

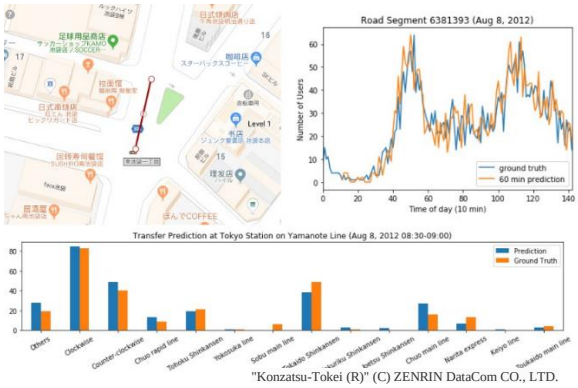
An efficient and accurate real-time people flow prediction at a metropolitan scale is the building block of various smart city applications. However, the various characteristics of different individuals, intrinsic complexity of the spatial layout and connectivity of the transportation network, and the time-varying of the city situation make predicting the flow of people flow a very difficult task. Bearing these in mind, we aim at addressing the main challenges of developing a practical people flow prediction system in the following three directions: 1) how to predict the people flow at the individual level rather than aggregated value; 2) how to predict a fine-grained people flow at a transportation network level rather than at a coarse level; 3) how to make the people flow prediction that adapts to different city situations (e.g. workdays, holidays, extreme events and disasters). In this talk, we will introduce our recent studies on real-time people flow prediction systems, and how our proposed prediction methods addressed these challenges.

Key Words : *Human Mobility Prediction, Spatiotemporal modeling, Deep Neural Network, GPS Trajectory Prediction*

1. Introduction

For crowd monitoring and traffic systems, fine-grained human movement predictions are crucial for decision-making and governmental responses. Many current studies offer broad predictions, like total population movement or next immediate location, rather than detailed paths aligned with transportation networks. These general forecasts aren't suitable for the advanced analysis needs of transportation agencies and companies.

Some examples of high-level human trajectory analysis from fine-grained human movement predictions are: 1) What will the traffic volume on a particular road/railway segment be in 1h, as shown in Figure 1; 2) How many people will get on/off and transfer in a specific station using a particular line (e.g., Tokyo station and Yamanote line), as shown in Figure 1; 3) How the users will move alternatively if the line stopped; 4) Where the future arrivals come from? And how they will get to the destination, as shown in Figure 1; 5) How many people will get on/off and transfer in a specific station using a particular line (e.g., Tokyo station and Yamanote line).



"Konzatsu-Tokei (R)" (C) ZENRIN DataCom CO., LTD.

Fig. 1 Traffic volume analysis for a particular road segment (upper), and transfer prediction at Tokyo station on the Yamanote line (lower).

For a thorough understanding of metropolitan scale human mobility prediction, we need to replicate and predict future agent movements more holistically in the virtual space based on current urban conditions

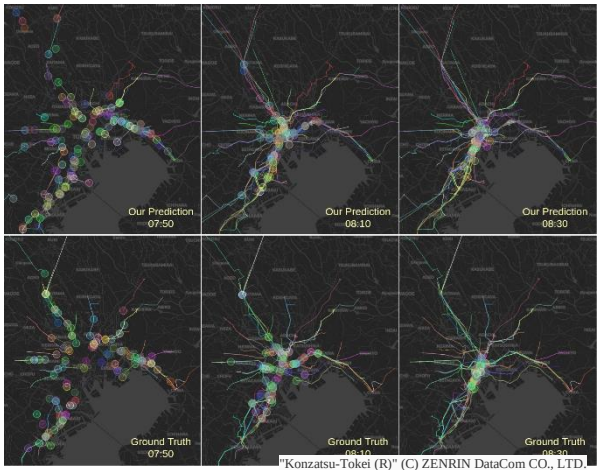


Fig. 2 Predicting the gathering pattern of Tokyo station at the morning rush hour.

or simulated scenarios. While advancements suggest the feasibility of real-time city mobility predictions, few studies have achieved fine-grained predictions for numerous agents efficiently. In our study, we use raw GPS data to refine our predictions on individual agents, aligning them with the transportation network and forecasting complete future trajectories. Our fine-grained predictions enable advanced analyses like anticipating traffic volume on specific roads an hour in advance or predicting transfer behavior at major transit hubs like Tokyo station. Challenges include:

1) Adapting predictors to daily variations in human mobility by considering both individual patterns and larger crowd contexts.
2) The complexity of urban multimodal transportation networks, with numerous road segments and types, makes human mobility predictions difficult.
3) Accurately generating long, fine-grained human trajectories is a significant challenge due to the vast potential routes and destinations.
4) Making fine-grained predictions from raw GPS data is computationally intensive, especially when considering uncertainties in location, possible travel modes, and varied road properties.

To tackle human mobility prediction challenges, we introduce a two-stage predictor that differentiates between coarse and fine-grained movements. In the initial stage, we automatically derive a city's current mobility state as a crowd context and adapt the coarse-level predictor via meta-learning and ensemble learning to forecast destination distribution at a broader scale. Here, the emphasis is on large-scale, long-duration, and inter-user relations without taking into account the local transport network, given that our local transport choices are mostly influenced by long-term destinations. In the subsequent stage, we employ a retrieval-based approach to craft predicted trajectories by sampling from a fine-grained historical trajectory database. The sampling weight is set by the coarse prediction from the first stage. This method moves intensive calculations to an offline stage, ensuring our system efficiently produces timely, fine-grained future trajectories. This paper is a summarized and extended version of our previous published work [1) 2)], we fused the models to address the extreme case of predicting the mobility in Tokyo on the day of 311 earthquake.

2. Our Approach

(1) Coarse Level Prediction

To capture long-distance and long-term patterns from user trajectories, we convert raw GPS data into cluster-level trajectories. This transformation is accomplished through a four-step process based on the raw GPS logs:

i) Adjust the trajectory using forwarding-fill to achieve a uniform-sampling trajectory. This method is preferable for online systems as it doesn't need future data.
ii) Assign a location index to the trajectory using Uber's Hexagonal Hierarchical Spatial Index (H3 index) at the 8th resolution, which equates to an average hexagon area of 0.74 km^2.
iii) Determine the visit frequency of hexagons using training data. Two hexagon sets are then

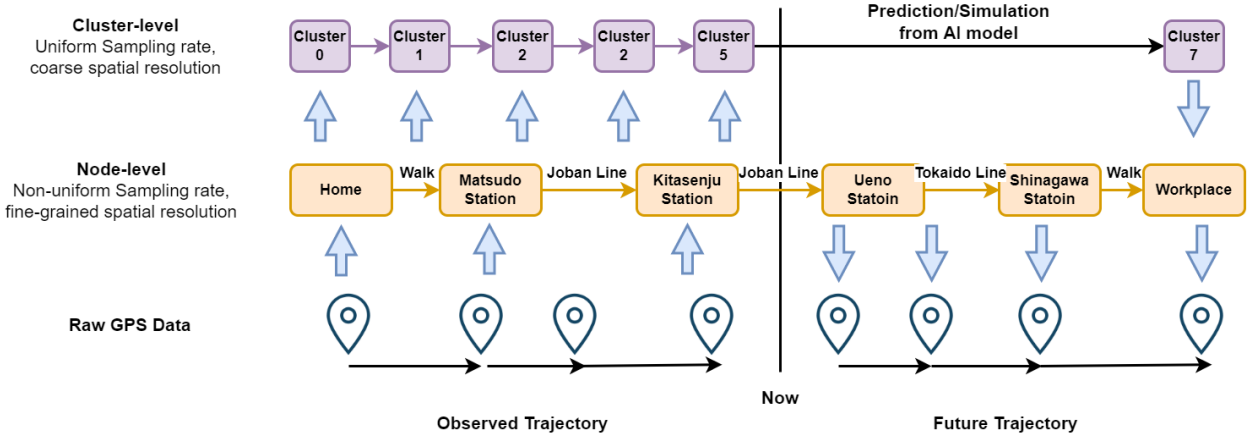


Fig. 3 Two staged mobility prediction framework

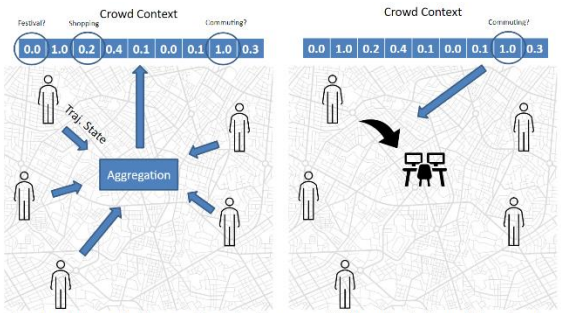


Fig. 4 Crowd context collects all the users' current trajectory state to recognize the current city state (left). Crowd context is conditioned on the prediction of each individual (right).

created: Type a) the top 1000 most visited hexagons and Type b) a sample of 2600 hexagons from the rest, chosen based on their visit frequency proportion.

iv) Simplify the H3-indexed trajectory by approximating the H3-indices that do not belong to Type a) using the nearest Type b) hexagon. Thus, we use 3600 indices to represent all the locations in the dataset.

Therefore, we regularize the raw GPS trajectory from both spatial and temporal aspects. Temporally, we convert the non-uniform trajectory into a uniformly sampled one, streamlining spatiotemporal features and speeding up processing. Spatially, we opt for discrete location representation over continuous coordinates due to the non-linear spatial dependency arising from intricate transport networks. Locations on highways or railways, for example, can be viewed as unique points difficult to represent with continuous functions.

To address the daily mobility pattern variations, we proposed a crowd-context predictor[1] that predicts the future movement of each individual user based on his/her most recent observations as well as the crowd context, which is aggregated from the recent observations of all the users in the city, as shown in the Figure 4. Figure 5 illustrates the crowd context's encoding of both regular and irregular aspects of human mobility. By examining two dimensions from the crowd context (4th and 58th), we observe their values' variations during the testing phase. The 4th dimension (upper figure) is highly sensitive to irregular mobility patterns, notably distinguishing the Obon festival week from typical weekdays. In contrast, the 58th dimension (lower figure) captures more regular mobility patterns, subtly differentiating between the Obon festival and a normal week, while being less affected by irregularities. Each dimension of our crowd context captures unique facets of large-scale human mobility, complementing each other. This

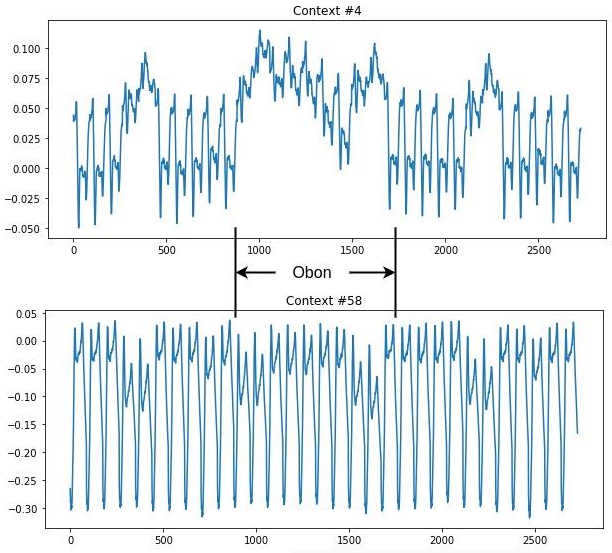


Fig. 5 Different dimension of crowd-context describes different dynamics of the city mobility.

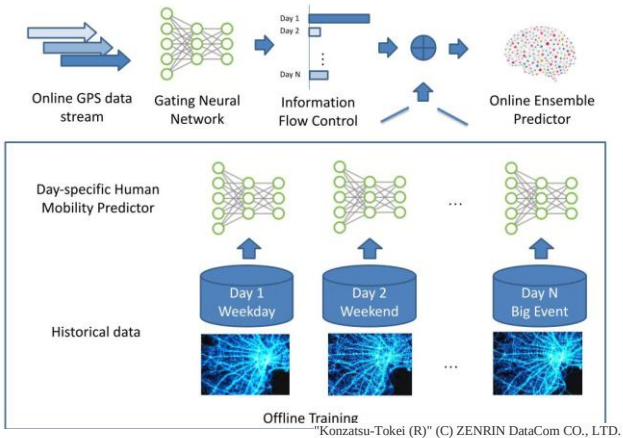
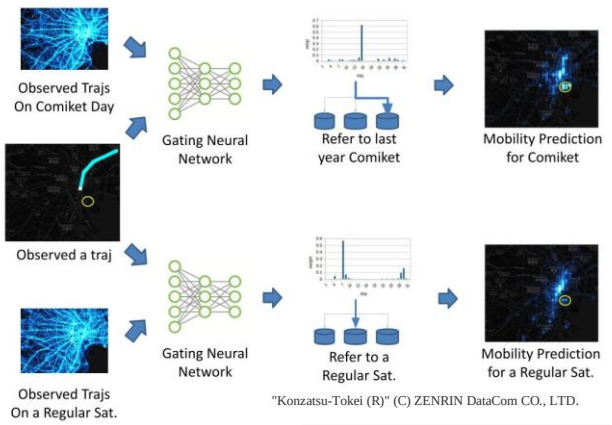


Fig. 6 Ensemble prediction framework

comprehensive view enables our cluster-level predictors to address both regular and irregular human mobility patterns.

To address the extreme irregular mobility cases such as during a big earthquake, we proposed an online adaptive ensemble predictor[2] that can automatically switch between the regular mode, which can be addressed by the crowd-context predictor, and irregular mode, which is designed for handling the irregular case specifically, as shown in Fig. 6. The irregularities in each day in the training data is encoded by the "expert" model that is trained offline, and selected in an online adaptive manner. As shown in Figure 7, we show how this framework can successfully address the irregular case of the mobility prediction during the comiket day. Our ensemble model can automatically recognize the irregular pattern of the mobility around the key stations access to the venue of the comiket, and adjust the ensemble predictor to switch to the "Comiket" mode that leverages the last year's



"Konzatsu-Tokei (R)" (C) ZENRIN DataCom CO., LTD.

Fig. 7 Ensemble prediction on Comiket day.

comiket day information, which generates much more reasonable predictions by raising the probability of the passengers come to the Comiket venue.

(2) Fine-grained level prediction

Trajectory interpolation and map matching are crucial for connecting GPS coordinates to transportation nodes and links. Yet, this process is labor-intensive and largely dependent on future data to clarify ambiguities, as illustrated in Figure 7. Determining whether a user is on the railway (blue) or road (black) is challenging without knowing their subsequent move (blue cross). Hence, real-time interpolation and map matching of raw GPS trajectories for online systems present difficulties. Therefore, we propose a retrieval-based prediction approach that offloads the heavy computation task to the offline phase.

Based on the processing pipeline in this study [3], we conduct a 3-step process to obtain the node-level trajectory as shown in Figure 7;

i) Trip Segmentation: From the GPS data's spatiotemporal features, stay points are determined (requiring a max movement distance of 300m and at least a 15-minute stay). Moving states are then identified as MOVE or STAY. The transportation mode (e.g., WALK, CAR, TRAIN) for each trip is discerned based on referenced methods.

ii) Interpolation Between Segmented Trips: Due to the sparse nature of raw GPS data, gaps often exist between trip segments. Using criteria like moving distance (max 200m) and gap duration (max 1h), along with states at gap ends, these gaps are filled:
   a) For substantial moving distances, a mid-term move state is added.
   b) If there's a state change, the transition's precise time is derived from moving state trajectories.
   c) Similar STAY or MOVE trips with identical transportation methods are merged.



Fig. 7 Illustration of using future information to disambiguate localization uncertainty (left); the key steps of trajectory interpolation and map-matching (right).

iii) Route Estimation: For greater accuracy, GPS points are matched to transportation nodes, and travel time is estimated. For every MOVE state trip, a route is chosen based on transportation type and road/railway category.

iv) However, this process is challenging for real-time online systems. This is because future data is essential for clearing ambiguities, especially with sparse GPS data, non-consistent sampling, and positioning errors. The route estimation step, being non-deterministic, demands significant computational power. Therefore, our focus shifts to acquiring high-quality, fine-grained historical trajectories offline. The paper will further detail how this offline data aids the online system within our retrieval-based prediction approach.

By obtaining the fine-grained node-level trajectories, we create the key-value store of historical trajectory database that queries with the feature vectors that concatenates i) the cosine and sine transformations of the current time-of-day and ii) the latitude and longitude of the first record of the trajectory in the time intervals.

In the online prediction phase, our cluster-level predictor firstly predict the 1h ahead desetination cluster of each individual in the city, and then the destination cluster along with the current last observation of the user and the current time is used as the query feature to find out a sample of possible historical trajectory as future prediction. In this way, we can successfully generate fine-grained future trajectories via recombine the samples from historical trajectory database with the help of the reweighting from cluster-level predictions. The entire pipeline of our proposed metropolitan-scale mobility prediction system is summarized in Figure 8.

3. Experiements

In this paper, we use a dataset "Konzatsu-Tokei (R)" data. “Konzatsu-Tokei (R)" Data refers to people flows data collected by individual location data sent

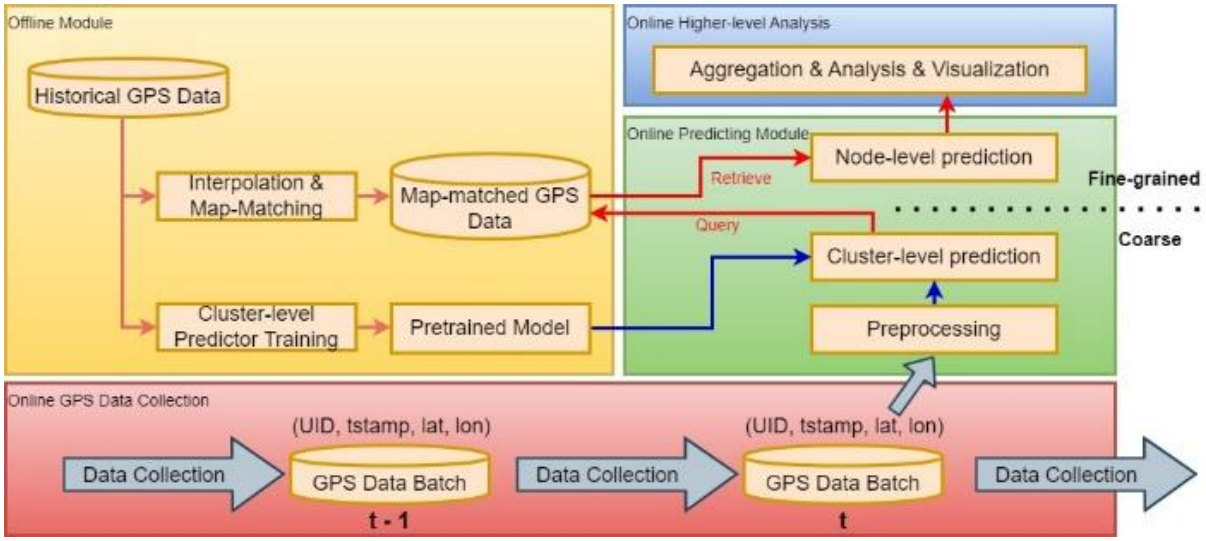


Fig. 8 The pipeline of our proposed metropolitan-scale mobility prediction system.

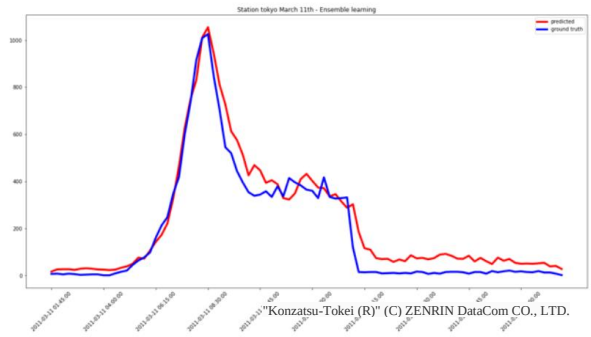


Fig. 9 Prediction results of the Tokyo station passenger using Ensemble prediction model (red) and the ground truth (blue)

from mobile phone under users' consent, through applications 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. Some applications such as “docomo map navi" service (map navi and local guide).

We use two months of data (from 2012.6.1 to 2012.07.31) for training/validating the cluster-level predictor and building the fine-grained historical trajectory database, and one month data (from 2012.08.01 to 2012.08.31) for evaluation. We cropped the data in Kanto area (covering Tokyo metropolitan area), with an average number of about 220K user IDs. For the sake of protecting user's privacy as well as efficiency, we do not store users' ID for long-observation and thus we are more focused on the movement distribution at the large scale while a long dependency of the user's mobility is not taken into consideration.

We deployed our algorithm on a deep learning workstation with Intel Xeon E5-2690v4, 2 x TitanX Pascal 12GB GDDR5X, 128GB and 1.2TB Intel® NVMe SSD DC P3600 Series. The algorithm was implemented in Python, except the trajectory interpolation and map-matching part were implemented in Java. We utilized the deep learning framework PyTorch 1.3.1 to construct the cluster-level predictor, and key-value storage library levelDB 1.20 for building and online querying the fine-grained history trajectory database.

We measure the time efficiency of the three steps in our online prediction phase. Fine-grained 1h-ahead trajectory prediction for about 220K users at a metropolitan scale takes about 2 min, which is a reasonable time latency for a practical system, without worrying about the data staleness and communication latency problems. Note that, except computing the crowd context, cluster-level and node-level prediction are conducted independently on each user. Thus a parallel acceleration strategy can be expected to achieve a desirable acceleration in our future work.

(1) Experiments on 311 Earthquake

Figure 9 shows our prediction results from ensemble predictor that can successfully recognize the earthquake mobility pattern and adapt the cluster-level predictor that also shows a sudden drop of the passenger flow due to the stoppage of the public transport after the earthquake.

(2) Experiments on Regular days

In Figure 10, we selected four types of aggregated analysis: 1) number of transfer, 2) traffic volume on a particular road segment, 3) number of passengers using the station on a particular line, and 4) the total number of passengers for all lines in an interchange. We compared our 1h-ahead prediction with the most widely used time-series prediction methods, namely ARIMA and prophet from 2012.08.01 to 2012.08.08. We can see from Fig. 10 that our fine-grained predic-

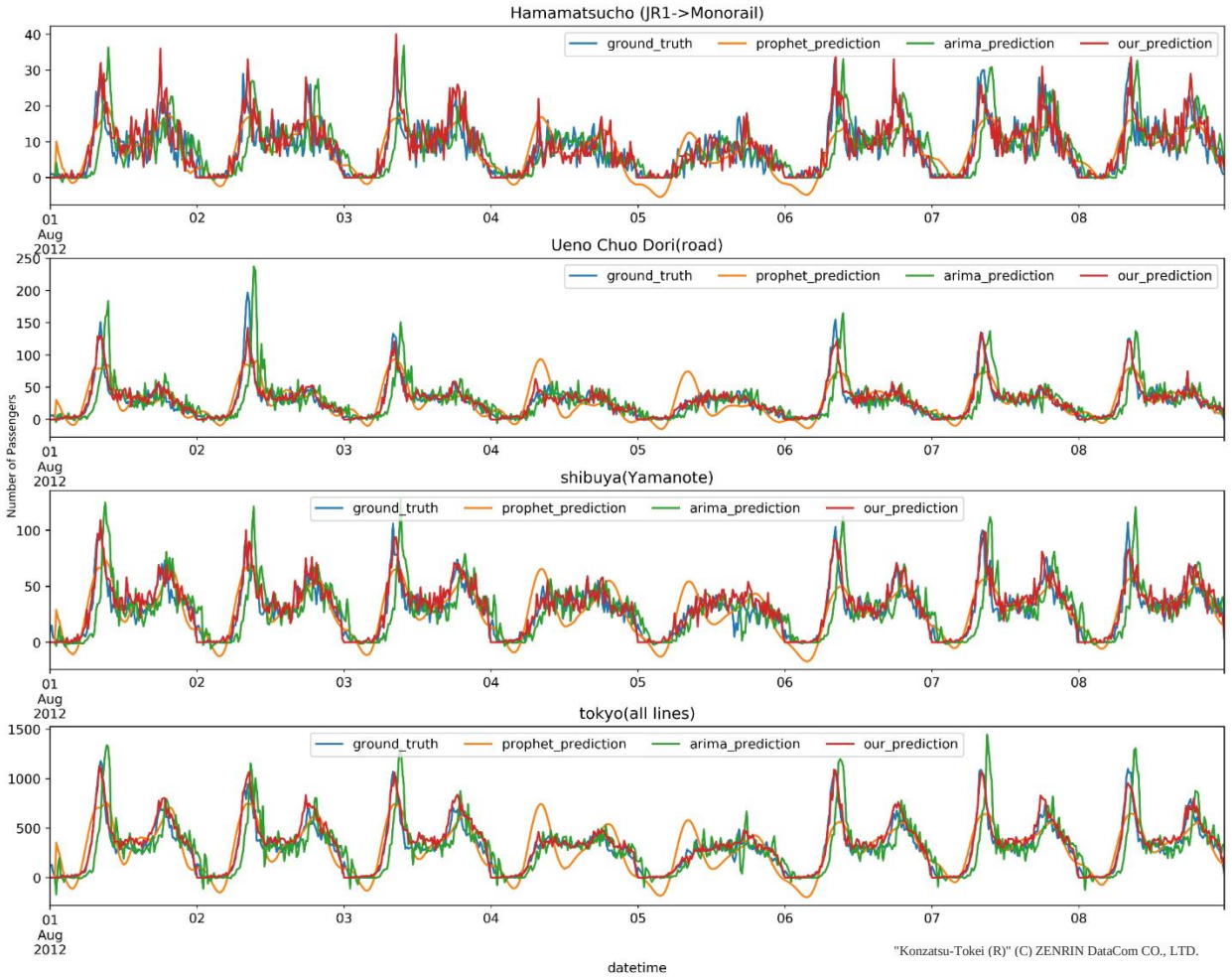


Fig. 10 Four types of aggregated analysis on our predictions on transport network.

tion can predict the sharp peaks of rush hours accurately and switch between the weekday and weekend state via crowd context automatically, which are difficult for traditional time series methods.

REFERENCES

1) Zipei Fan, Xiaojie Yang, Wei Yuan, Renhe Jiang, Quanjun Chen, Xuan Song, and Ryosuke Shibasaki. 2022. Online Trajectory Prediction for Metropolitan Scale Mobility Digital Twin. In Proceedings of the 30th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems.

2) Zipei Fan, Xuan Song, Tianqi Xia, Renhe Jiang, Ryosuke Shibasaki, and Ritsu Sakuramachi. 2018. Online deep ensemble learning for predicting citywide human mobility. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies 2, 3 (2018), 1--21..

3) Yu Zheng, Quannan Li, Yukun Chen, Xing Xie, and Wei-Ying Ma. 2008. Understanding Mobility Based on GPS Data. In Proceedings of the 10th International Conference on Ubiquitous Computing (Seoul, Korea) (UbiComp '08). ACM, New York, NY, USA, 312–321.