ZEST: a Hybrid Model on Predicting Passenger Demand for Chauffeured Car Service

Hua Wei[†], Yuandong Wang[†], Tianyu Wo[†], Yaxiao Liu[‡], Jie Xu[†,§]
[†]State Key Laboratory of Software Development Environment, Beihang University, China;
[‡]Department of Computer Science and Technology, Tsinghua University, China;
[§]School of Computing, University of Leeds, UK
{weihua, wangyd, woty}@act.buaa.edu.cn; liuyaxiao12@mails.tsinghua.edu.cn;
J.Xu@leeds.ac.uk

ABSTRACT

Chauffeured car service based on mobile applications like Uber or Didi suffers from supply-demand disequilibrium, which can be alleviated by proper prediction on the distribution of passenger demand. In this paper, we propose a **Z**ero-Grid **E**nsemble **S**patio**T**emporal model (*ZEST*) to predict passenger demand with four predictors: a temporal predictor and a spatial predictor to model the influences of local and spatial factors separately, an ensemble predictor to combine the results of former two predictors comprehensively and a Zero-Grid predictor to predict zero demand areas specifically since any cruising within these areas costs extra waste on energy and time of driver. We demonstrate the performance of ZEST on actual operational data from ride-hailing applications with more than 6 million order records and 500 million GPS points. Experimental results indicate our model outperforms 5 other baseline models by over 10% both in MAE and sMAPE on the three-month datasets.

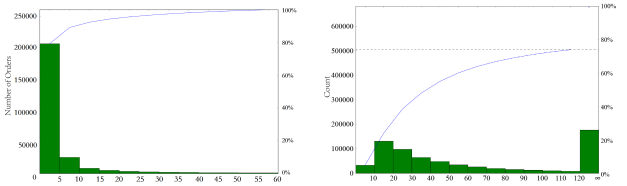
Keywords

Spatiotemporal Data Mining; Demand Prediction; Chauffeured Car Service

1. INTRODUCTION

Chauffeured car service based on ride-hailing applications like Uber and Didi serves as an important transportation to provide passengers with convenient and professional services. However, in metropolises like Beijing or New York, it is common to see passengers seeking for taxicabs roadside either in downtown or in suburbs while some taxi drivers are cruising idly on the street.

This contradiction reveals the supply-demand disequilibrium with the following two scenarios: *Scenario 1*, demand exceeds supply, where passengers' needs would not be met in a timely response. As is shown in the Figure 1(a), 90% passengers who withdrew their requests canceled their orders in the first 10 minutes and specifically, 80% canceled in their first 5 minutes. *Scenario 2*, supply exceeds demand, where drivers would spend overly long time in seeking for passengers. Figure 1(b) indicates that over 20% chauffeured drivers spend over 2 hours seeking for their next customers while only 40% drivers can pick up their next passengers within 30 minutes. In particular, areas that are highly possible of having zero demand in the next time period should be more concerned because any cruising within may be a waste in energy and time. Besides, after knowing some areas are very likely to be zero-demand area, the drivers can avoid driving to those areas, while dispatching centers of car companies can dispatch their cars and set their prices ahead of time dynamically, which can successively alleviates the disequilibrium of both Scenario-1 and Scenario-2 areas.



(a) Waiting time (min) (b) Cruising time (min)

Figure 1: Distribution of passengers' waiting time before canceling order and drivers' cruising time in Beijing during September 2015

To solve the problem of disequilibrium, an overall prediction for passenger demand provides a global distribution of passengers, upon which we can dispatch in supply and adjust prices dynamically in advance. Although there are plenty of methods for prediction like Linear Regression and Multiple Additive Regression Tree, particularly, the pattern of demand for road transportation can be modeled as Poisson Process [5]. Li et al. [4] and Moreira-Matias et al. [6] both took advantage of Auto-Regression Moving Average method in time series analysis, which predicted the number of passengers in certain region based on historical pick-up data. However, aforementioned studies have assumed that the quantity of demanding passengers is equal to the pick-up counts recorded by taxi. But in countries where roadside taxi hailing is permissible, there are still plenty of people who manage but fail to take taxis.

Thus, there is a bias between the pick-up records and passenger demand. Zhang et al. [7] noticed this bias and ob-

CIKM'16 , October 24-28, 2016, Indianapolis, IN, USA

DOI: http://dx.doi.org/10.1145/2983323.2983667

served online hidden context to infer the passenger demand. Instead of inferring the unpicked-up passengers, Jiang et al. [3] utilized the order data generated by demanding passengers through ride-hailing applications, which properly settled the bias since passengers in demand can only make requests through mobile applications. Therefore, the demand of passengers can be considered equal to the amount of orders recorded by applications. But the model they utilized was based on Linear Regression, which oversimplified the temporal and spatial influences. Moreover, as their model is based on regression, they cannot provide information on the possibility of areas having zero demand in the next time period.

In this manner, this paper focuses on the prediction of passengers in demand of chauffeured car service and how possible one region will have zero demand in certain periods. The major contributions of this paper can be summarized as follows: a) Based on the local data, a simple yet novel parameter named *FluctuationRate* is proposed in the temporal predictor to filter out irrelevant historical records; b) the influences of neighborhoods are modeled using different kinds of datasets, while most of the existing researches are based on onefold data; c) a specified model named *Zero-Grid Predictor* is presented to predicted areas with zero demand, along with an ensemble predictor combining the results of temporal and spatial model. The integration keeps the advantages of temporal and spatial model and provides the possibility of zero demand area, which is not covered by many state-of-art prediction models on passenger demand. Evaluations on real operational datasets with over 6 million order records and 500 million GPS points show that compared to ground truth, our model achieves 0.68 in MAE and 12.6% in sMAPE on average, which outperforms other five baseline models.

2. METHODOLOGY

In this section, we clarify some definitions and present the prediction model, including the framework and detailed design of each component.

2.1 Preliminary

Definition 1: *Time Period, Hour of Day and Day of Week.* We divide one day into several time periods while in experiment, the length of one period is 60 minutes and one day is divided into 24 periods. For example, when the length of each time period is one hour, "2015-06-01 16:03:22" at Monday's hour of day (HOD) and day of week (DOW) will be 16 and 0 accordingly.

Definition 2: *Grid, Row, Col.* We divide one city into mutually disjoint grids, where each point in the records can be mapped into one grid by locating the index of row number *row* and column number *col*. In experiments, the distance from one grid to another is approximately equal to the driving distance within 10 minutes in the city.

Definition 3: *Order.* Given the time period t and grid r, the amount of order records requested from passengers within t and r are denoted as $Order_t^r$.

Definition 4: *Zero-Grid.* When a grid r has zero order in time period t, i.e $Order_t^r = 0$, we call r is a Zero-Grid. From the perspective of saving energy and time, Zero-Grids should be given close attention in dispatching to meet the need of demanded areas.

2.2 Framework of Model

As is shown in Figure 2, the construction of our model is mainly based on the offline learning on historical data, online data preprocessing and online prediction. For order data and GPS data stream, we firstly partition and aggregate them into certain grids and time periods as *Order*. Then in the offline learning process, preprocessed data is separately fed into predictors to re-train our models every week. Upon the model, online predictions can be generated with preprocessed input data.

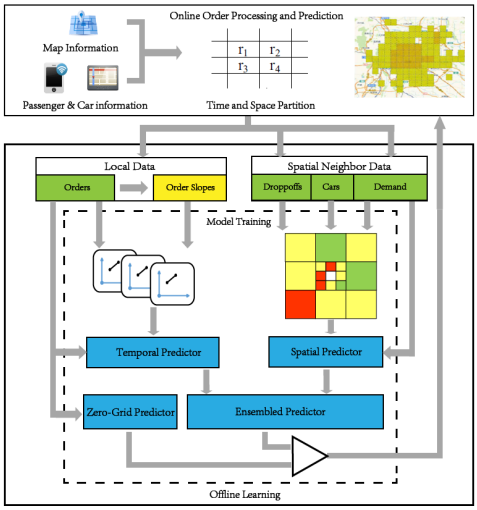


Figure 2: Framework of the prediction model

As is shown in Figure 2, our prediction model includes a temporal predictor (TP), a spatial predictor (SP), an ensemble predictor (EP) that combines former two predictors and a Zero-Grid predictor (ZP) which provides the possibilities of a grid r being a Zero Grid. When a grid has a high possibility of having zero demand in the next time period, the final prediction will be zero.

2.3 Temporal Predictor

Intuitively, given a region, its demand at certain period could be roughly deduced from demand at comparable periods, like same time of the day or same day of the week. However, the passenger demand is sometimes irregularly variant at same time of the day and may lead to bias if all historical periods are taken into prediction.

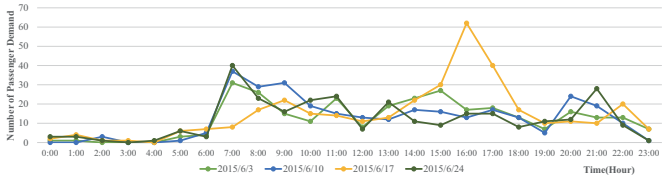
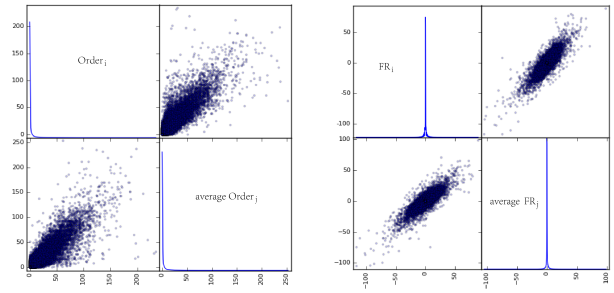


Figure 3: Passenger demand near Renmin University of Wednesdays in June 2015

Figure 3 gives an example of the passenger demand in Renmin Univerisity area at Wednesdays in July where the passenger demand did fluctuate in different days. When we were predicting for the passenger demand in 16:00 of



(a) $Order_i$ of average $Order_j$ (b) FR_i of average FR_j

Figure 4: Correlations of *Order*s and *FR*s

July 24th, the prediction value was $\frac{17+13+67}{3} \approx 30$ using all historical data. If we found out that patterns in July 3rd and 10th are more similar with 24th than 17th based on the change of demands, we would get the inference of $\frac{17+13}{2} = 15$ using more relevant data, which is closer to the actual value (16 in this case).

Therefore, we define a parameter *FluctuationRate* to reflect the similarities between demand fluctuation in different time periods. Given grid r at time i, *FluctuationRate* is defined as Equation 1:

$$FluctuationRate_i^r = \frac{Order_i^r - Order_{i-1}^r}{\Delta t} \quad (1)$$

where Δt is the length of time slots. In the following sections of this paper, we abbreviate $FluctuationRate_i^r$ as FR_i^r. As is shown in Figure 4, the correlations between FR_i and the average value of FR_j is more apparent than the correlations between $Order_i$ and the average $Order_j$ for every j, where $i - 4 \leq j < i$.

In temporal predictor, FR is employed to filter out less related period in past three months. The inputs of the TP are : a) the order number of grid r of last time slot t_{cur}, $Order_{t_{cur}}^r$; b) the set of historical $Order_i^r$ in last three months; c) the set of historical FR_i^r in last three months; and d) the similarity error ε where we say $FR_{t_m}^r$ and $FR_{t_n}^r$ are similar when $|FR_{t_m}^r - FR_{t_n}^r| < \varepsilon$.

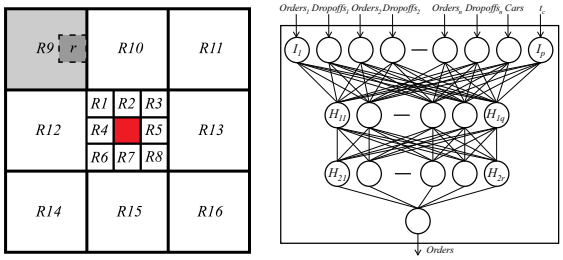
The details of the algorithm are presented as follows: Firstly, we calculate the parameter $FR_{t_{cur}}^r$ for current time t_{cur}. Secondly, we select all the time slot t_i whose FR is similar with $FR_{t_{cur}}^r$. Then we select t_i's next time slot and its corresponding FR, upon which we can predict the order number of grid r of next time slot t_{cur+1} through the average of FRs and $Order_{t_{cur}}$. Note that unless the data after filtering is null, a traditional time-varying poisson model will not be used.

2.4 Spatial Predictor

Spatial predictor makes use of target grid's neighborhood data to predict its future demand, including passenger demand and drop-off data, along with the vehicle data of the target grid.

Intuitively, passenger demand is related to each other in different regions as the passengers and vehicles are moving around the city. Passengers may take public transportation to the nearest subway or bus station and then orders chauffeured service to their destination. Therefore, it's inferred that passenger demand $Order_t^r$ in grid r at time t is related with the drop-off passengers in r's neighborhood before t, $Dropoff_{t_{before}}^{r_{neighbor}}$. Besides, it's also speculated that since passengers who had left grid r just before t was unlikely to show up again in r at t, passenger demand $Order_t^r$ is related with the demand in r before t, $Order_{t_{before}}^r$. What's more, drivers also have their own experience in pick-up points, which could be referred by the number of vehicle ($Cars$) in target region.

For a target grid to predict, we divide and aggregate its neighborhood into uneven grids as is shown in Figure 5(a) considering range-based influence, feature quality and dimension. For grids with larger size, the final value of each feature will be an average value of corresponding features regarding all included grids. Note that the overall neighborhood area in Figure 5(a) contains 79 grids, which would increase the feature dimension dramatically if there was no aggregation.



(a) Dividing scheme (b) Features for SP

Figure 5: Framework of spatial predictor

As is shown in Figure 5(b), for a target grid to predict, we use four types of features: time t_c to predict; the number of passenger demand in grid r_i, $Orders^{r_i}$; the number of drop-off passengers, $Dropoffs^{r_i}$; the number of vehicles, $Cars^i$. Firstly, passenger demand features are generated from the demand data of target's surrounding three-circle areas (containing 24 grids) in past three time periods (1 hour ago, 2 to 3 hours ago and 4 to 6 hours ago), one average value per period and grid. Secondly, passenger drop-off features are generated from the drop-off data of target's surrounding first circle with 8 grids in past three time periods (1 hour ago, 2 to 3 hours ago and 4 to 6 hours ago). Thirdly, the vehicle feature are generated from the quantity of vehicles within target grid in past one hour.

After generating features, an artificial neural network is trained to model the correlations in demands at different grids where the number of layers and the number of nodes in hidden layers are related with the input data.

2.5 Ensemble Predictor

In ensemble predictor, we aggregate the result of temporal predictor and spatial predictor. Although the idea of modeling temporally and spatially is intuitive yet widely used in spatiotemporal prediction [1] [8] [9], it is still flexible when confronted with different objectives. Note that the temporal predictor makes use of local data and spatial predictor mainly make use of its neighborhood data, which are non-repetitive and can provide not only local but relatively overall information in predicting target grid's passenger demand.

In practice, we train Gradient Boosting Decision Trees (GBDT) to combine the results of different predictors. To train such trees, we transform the time into two features:

HOD and DOW, along with the location of grids and other two features: the results of temporal and spatial predictor.

2.6 Zero-Grid Predictor

Based on historical records, a zero-grid predictor can provide possibilities of zero-demand areas for drivers and chauffeured car services. Before training ZP, we select features correlated to the zero demand including location, time period and this grid's historical records. For a given grid r at time t to predict, features are generated from r's average demand of past one and two hours and past demand at the same DOW and HOD of past one and two weeks. Besides, we use r's relative position (row and col) in a city, along with the time t's DOW and HOD. Specifically, considering that the location features row and col are non-linearly related, a Gradient Boosting Classifier is trained to predict the possibility of Zero Grid. Upon a zero-grid predictor, we utilize the ensemble predictor to infer regions where the demand has a low possibility of being zero.

3. EXPERIMENT

In this section, we introduce the datasets, baselines and results of our experiment.

3.1 Datasets

In this paper, we generate three kind of data, including the demand and drop-off data and the vehicle data through two benchmark datasets: the order dataset collected from mobile phone applications and GPS dataset collected from onboard GPS devices in real operating cars. These real operational datasets are sampled and provided by one of China's largest car rental and car-hailing service providers. The details are summarized in Table 1.

Table 1: Summary of datasets

Data Source	Domain	Entities
Order	Orders with estimated on-board location	6,792,756 orders
	Orders with actual drop-off location	4,206,832 drop-offs
GPS	GPS points of cars	512,658,480 GPS points
Urban Size	Area in 6th Ring Road of Beijing, 64×67 km	30×24 grids

3.2 Baselines

In experiments, we compare our model with 5 different baselines methods in total. In each experiment, the comparison of model performance is under the same temporal and spatial conditions for fair. The first three methods are classical and widely used techniques for prediction while the next two methods have been applied to passenger demand prediction in most recent studies:

Classical Poisson Process (TVP_C) was firstly proposed in [2] and the model assumes that the demand of passengers obeys a Non-homogeneous Poisson process with a time-dependent arriving rate function $\lambda(t)$.

Weighted Time-Varying Poisson Process (TVP_W) is a model modified from TVP_C, which adds weights for latest demand pattern [6].

Multiple Additive Regression Tree ($MART\_C$) is a widely-used machine learning technique for regression problems, which produces a prediction model and improves the quality of each base learner.

Classical Linear Regression (LR_C) is an approach for modeling the relationship between target variable and independent variables, which has been applied in studies in demand prediction [3].

Auto-Regression Moving Average ($ARMA$) provides a description of a stationary stochastic process in terms of two polynomials in the statistical analysis of time series, one for the auto-regression and the second for the moving average [6].

3.3 Ground Truth and Metrics

Since the actual value of orders (including the canceled ones) in grid r at time t, a_t^r can directly reflect the passenger demands, we take it as **ground truth**. In the predicting process, we compare the result of prediction p_t^r with the actual value a_t^r in the following metrics:

Regression Metrics. In the evaluation of regression problem, the MAE and sMAPE are utilized to reflect the accuracy of prediction on the number of passenger demand.

$$MAE = \frac{1}{n}\sum |p_t^r - a_t^r|, sMAPE = \frac{1}{n}\sum \frac{|p_t^r - a_t^r|}{\delta_t^r} \quad (2)$$

$$\delta_t^r = \begin{cases} c & p_t^r = a_t^r = 0, c \neq 0 \\ p_t^r + a_t^r & otherwise \end{cases} \quad (3)$$

Classification Metrics. The measures precision ($prec$) and recall (rec) defined as Equation 4 are popular metrics used to evaluate the quality of a classification system:

$$prec = \frac{tp}{tp + fn}, rec = \frac{tp}{tp + fp} \quad (4)$$

where tp is the number of true positive predictions, fn is the number of false negative predictions and fp is the number of false positive predictions.

3.4 Experimental Results

3.4.1 Results of Temporal Predictor

For temporal predictor, we reveal the effectiveness of using $FluctuationRate$ as a filter on the selection of historical data. All the methods in comparison are conducted on the dataset of Beijing in July, August and September, 2015. And the results of MAE and sMAPE are an overall average for every month.

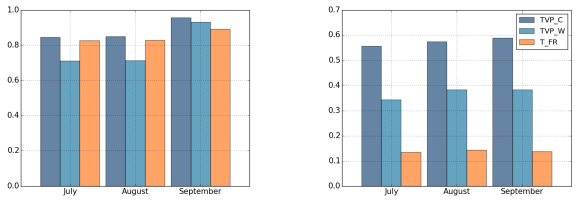
As is shown in Figure 6, while the MAEs of the three models are about the same, after adding FluctuationRate (T_FR), the performance on the metrics of sMAPE is greatly improved than traditional Poisson Process (TVP_C) and Weighted Time-Varying Poisson Process Model (TVP_W).

3.4.2 Results of Spatial Predictor

As we mentioned in Section 2.4, feature dimensions would be much higher without aggregation and might have great impact on the training of model. Thus, we firstly verify

Table 2: Results of spatial predictors with different features

	Feature Types			Orders' Time Range			Orders' Distance Range			Dropoffs' Distance Range		
	F_t+F_o	F_t+F_o +F_d	F_t+F_o +F_d+F_c	1h	3h	6h	1c	2c	3c	1c	2c	3c
MAE	3.1637	1.3664	**1.0339**	1.2309	1.0972	**1.0861**	1.1827	1.1031	**1.0331**	1.9564	1.8140	**1.0339**
sMAPE	0.2334	0.2329	**0.2316**	0.2428	0.2312	**0.2308**	0.2611	0.2371	**0.2313**	0.2394	0.2343	**0.2316**

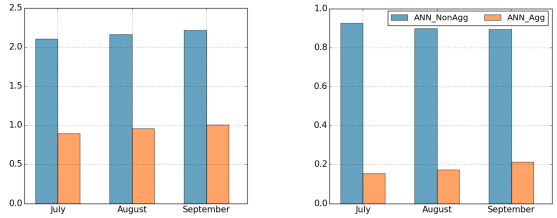


(a) MAE (b) sMAPE

Figure 6: Results of temporal predictor

the significance of division and aggregation, upon which we clarify the effectiveness of the features that we use in the spatial predictor in every month.

Effects of Division and Aggregation In this experiment, features are extracted from the order and GPS datasets of Beijing in July, August and September, 2015. Then we feed the original and aggregated features into *ANN_NonAgg* and *ANN_Agg* respectively. Results in the Figure 7 show that the process of division and aggregation has remarkably improved the performance of learning by over 50% in both MAE and sMAPE for all three months.



(a) MAE (b) sMAPE

Figure 7: Results of models with and without aggregation

Effects of Different Features. In this experiment, stochastic continuous data in two weeks between June and September 2015 are chosen as training data, and the models are validated on the data of the next week accordingly.

Comparison between feature types. Results in Table 2 illustrate that the increase in relevant feature types we select leads to the improvements on both MAP and sMAPE, where F_t stands for *time*, F_o for *Orders*, F_d for $dropoffs$ and F_c for *Cars*.

Range selection for $Orders_{t_j}^{r_i}$. As is shown in Table 2, by extending the time period of passenger demand from 1 hour before t_i (*1h*) to 3 hour (*3h*) and 6 hour (*6h*) step by step, there is a reduction in both MAE and sMAPE. So does with the expanding of grids inside the first circle (*1c*), the second (*2c*) and the third circle (*3c*).

Range selection for $Dropoffs_{t_j}^{r_i}$. The results in Table 2 indicate that adding time periods into features increases the performances of the prediction step by step (*1c, 3c, 6c*). Hence, it's helpful to feed the drop-offs in past 1 hour in surrounding three circle regions into the spatial predictor.

3.4.3 Results of Ensemble Predictor

Table 3 presents that the performance is reached in prediction when combining TP and SP into EP (called EST). In Table 3, the ensemble predictor not only outperforms the individual TP and SP by over 15% in MAE and over 6% in sMAPE, but also has a lower MAP and sMAPE than feeding all the features into baseline methods (LR_C and $MART\_C$). This is mainly because ensemble predictor keeps the advantages of temporal predictor and spatial predictior which utilizes both local and global information through linear and non-linear models.

Table 3: Results of ensemble predictors

Model	All Samples		Zero Samples	
	MAE	sMAPE	MAE	sMAPE
TP	0.8591	0.1407	0.1601	0.6082
SP	1.0339	0.2316	0.1687	**0.2611**
LR_C	1.2927	0.1445	6.3438	1.0000
$MART\_C$	0.8081	0.1394	8.9117	1.0000
EST	**0.7037**	**0.1321**	**0.0012**	1.0000

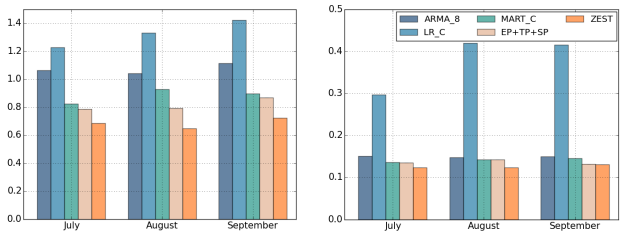
Specifically, the sMAPEs for samples whose actual demand is zero provided by LR_C, $MART\_C$ and $TP + SP$ are all 1.0000, which indicates the ineffectiveness of these predictors for zero samples. Therefore, a Zero-Grid Predictor is necessary for performance improvements.

3.4.4 Results of Zero-Grid Predictor

As mentioned in Section 2.6, we train a Zero-Grid Predictor (ZP) to provide information for driver chauffeured car service. The performances of ZP on different months when possibility threshold of positive is 0.5, where *prec* and *rec* of prediction are nearly 93% and 75% on average and the AUC of ZP reaches 95.8%.

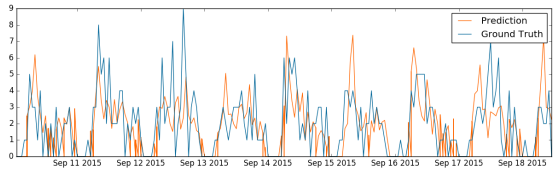
3.4.5 Overall Results

The comparison between different methods are shown in Figure 8 adding $ARMA\_8$ and the combination of our **Z**ero-Grid **E**nsemble **S**patio**T**emporal model ($ZEST$). $ARMA\_8$ is an ARMA model taking past eight hours as a moving average. The results in Figure 8 show that $ZEST$ outperforms $ARMA\_8$, LR_C and $MART\_C$ by over 10% in the prediction for every month. What's more, the integration of ZP also improves the overall performances of EP+TP+SP.

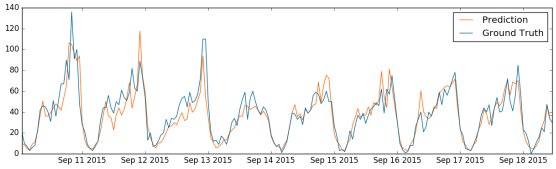


(a) MAE (b) sMAPE

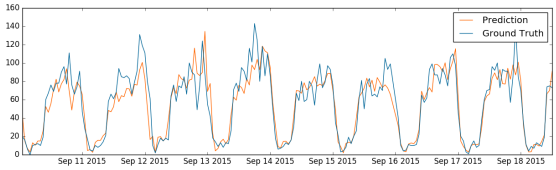
Figure 8: Overall results of predictors



(a) Prediction of Yizhuang area



(b) Prediction of Sanlitun area



(c) Prediction of Beijing Capital International Airport

Figure 9: Results of *ZEST* against ground truth

Specifically, several areas are displayed as examples of our model's results against the ground truth from September 10 to September 18 in Beijing in Figure 9. Figure 9(a) illustrates the prediction in Yizhuang Area, one of Beijing's biggest residential areas, and Figure 9(b) shows the prediction in Sanlitun Area, Beijing's business and entertainment district. It can be conveyed that Yizhuang suffers from morning rush-hour in a small extent while Sanlitun suffers from evening and late-night peaks, where in general, our model catches these peaks and deals with off-peaks properly, especially the time periods with zero demand. Besides, Figure 9(c) shows the performances of *ZEST* in Beijing's traffic zones, where our model can accurately forecast the demand in these districts.

4. CONCLUSION

In this paper, we develop a hybrid model in predicting the passenger demand which makes use of different kinds of datasets from chauffeured cars. The effectiveness of *FluctuationRate* is confirmed to select similar historical records in temporal prediction. Then we model the influences of neighborhood regions to the target region through features like demands, drop-offs and vehicle numbers in a spatial predictor and then combine the temporal and spatial predictor all together. What's more, to improve the working efficiency for chauffeured car drivers and alleviate supply-demand imbalance, we train a Zero-Grid predictor to predict the possibility of an area being a zero demand area. The integration of Zero-Grid predictor, ensemble predictor, spatial predictor and temporal predictor outperforms five baseline methods.

Acknowledgment

This work is supported by the 863 Program (2015AA01A202), China Key R&D Program (2016YFB1000103), NSFC (61421003) and Beijing Municipal Science and Technology Commission (Z151100003615002). Tianyu Wo is the corresponding author of this paper.

5. REFERENCES

[1] T. Cheng, J. Haworth, B. Anbaroglu, G. Tanaksaranond, and J. Wang. Spatiotemporal data mining. In *Handbook of Regional Science*, pages 1173–1193. Springer, 2014.

[2] A. Ihler, J. Hutchins, and P. Smyth. Adaptive event detection with time-varying poisson processes. In *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 207–216. ACM, 2006.

[3] W. Jiang, T. Wo, M. Zhang, R. Yang, and J. Xu. *A Method for Private Car Transportation Dispatching Based on a Passenger Demand Model*. Springer International Publishing, 2015.

[4] X. Li, G. Pan, Z. Wu, G. Qi, S. Li, D. Zhang, W. Zhang, and Z. Wang. Prediction of urban human mobility using large-scale taxi traces and its applications. *Frontiers of Computer Science*, 6(1):111–121, 2012.

[5] L. Matias, J. Gama, J. Mendes-Moreira, and J. Freire, de Sousa. Validation of both number and coverage of bus schedules using avl data. In *International IEEE Conference on Intelligent Transportation Systems*, pages 131–136, 2010.

[6] L. Moreira-Matias, J. Gama, M. Ferreira, J. Mendes-Moreira, and L. Damas. Predicting taxi–passenger demand using streaming data. *Intelligent Transportation Systems, IEEE Transactions on*, 14(3):1393–1402, 2013.

[7] D. Zhang, T. He, S. Lin, S. Munir, J. Stankovic, et al. Dmodel: Online taxicab demand model from big sensor data in a roving sensor network. In *Big Data (BigData Congress), 2014 IEEE International Congress on*, pages 152–159. IEEE, 2014.

[8] Y. Zheng, F. Liu, and H. P. Hsieh. U-air: when urban air quality inference meets big data. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1436–1444, 2013.

[9] Y. Zheng, X. Yi, M. Li, R. Li, Z. Shan, E. Chang, and T. Li. Forecasting fine-grained air quality based on big data. In *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 2267–2276, 2015.