Multi-modal Sequence to Sequence Learning with Content Attention for Hotspot Traffic Speed Prediction

Binbing Liao[1], Siliang Tang[1], Shengwen Yang[2], Wenwu Zhu[3], and Fei Wu[1(✉)]

[1] College of Computer Science and Technology, Zhejiang University, Hangzhou, China
{bbliao,siliang,wufei}@zju.edu.cn
[2] Baidu Inc., Beijing, China
yangshengwen@baidu.com
[3] Department of Computer Science and Technology, Tsinghua University, Beijing, China
wwzhu@tsinghua.edu.cn

Abstract. Traffic speed prediction is a crucial and fundamental task of the intelligent transportation systems (ITS). Due to the dynamic and non-linear nature of the traffic, this task is difficult. Nonetheless, the collection of crowd map queries data brings new ways to solve this problem. Generally speaking, in a short period of time, a large amount of crowd map queries aiming at the same destination may lead to traffic congestion. For instance, large queries for Family Restaurant during the dinner time lead to traffic jams around it. However, traffic speed prediction with crowd map queries is challenging due to the complexity and scale of the map queries, as well as their modalities. To bridge the gap, we propose Multi-Seq2Seq-Att for hotspot traffic speed prediction. Multi-Seq2Seq-Att is a multi-modal sequence learning model that deals with two sequences in different modalities, namely, the query sequence and the traffic speed sequence. The main idea of Multi-Seq2Seq-Att is to learn to fuse the multi-modal sequence with content attention. With this method, Multi-Seq2Seq-Att addresses the modality gap between queries and the traffic speed. Experiments on real-world datasets from Baidu Map demonstrates a 24% relative boost over other state-of-the-art methods.

Keywords: Traffic speed prediction · Map query · Content attention

1 Introduction

With the ever-increasing urbanization process, traffic congestion has become a common urban problem around the world. As a crucial task of the ITS, accurate and real-time traffic prediction is particularly useful for many applications like route guidance, traffic network planning, and congestion avoidance [17]. However,

R. Hong et al. (Eds.): PCM 2018, LNCS 11164, pp. 212–222, 2018.
https://doi.org/10.1007/978-3-030-00776-8_20

this task is challenging due to the complexity and dynamic property of the traffic environment in urban cities.

Previous methods for traffic prediction can be classified into traffic prediction with unimodal and multi-modal data. On one hand, autoregressive integrated moving average (ARIMA) [1] and its family [15,16] are widely used for traffic prediction. However, ARIMA and its family require high computational resources which make them not suitable for large-scale problems. Taking the stochastic and nonlinear nature of traffic into account, some non-parametric methods are proposed for traffic prediction, such as k-NN [3], RF [9] and SVR [8]. Recently, some deep learning methods are proposed for traffic prediction such as deep belief network [7] and stacked autoencoders (SAEs) [13]. Most of the aforementioned works consider the traffic prediction for highways, whose traffic condition are relatively stable and simple. While in urban cities, the traffic is highly dynamic and varies greatly due to diverse and complicated factors (e.g., crowd activities). Thus methods which ignore the multi-modal factors may no longer work.

On the other hand, a few researchers attempted to predict the traffic with related multi-modal data. [5,11] proposed an optimization framework to extract traffic indicators based on location-based social media and incorporate them into traffic prediction via linear regression. However, using a linear regression may be insufficient when the traffic is a nonlinear system. [14] proposed an LSTM neural network for traffic prediction using microwave detector data. Nonetheless, in real-world road systems, only a small fraction of the road segments are deployed with sensors. For those road segments without sensors, previous methods may no longer work.

With the rapid growth of mobile technology, map applications (e.g., Baidu Map and Google Map) provide a rich source of information for traffic prediction. Figure 1 shows the average traffic speed and crowd query counts around the Capital Gym, Beijing on April 8, 2017. Query counts at time t are the numbers of queries issued for the Capital Gym, with the estimated time of arrival for these queries to be t. We can tell that the number of query counts (in red) are higher more than that of average historical query counts (in blue) around 18 PM, which predicts a sudden drop in the traffic speed. Note that several queries targeting the same destination in a short duration could lead to traffic jams at their designated destination after a while. Therefore, crowd map queries can provide early warnings (The number shown in [10] is 46 min) of traffic jams, which is good for many applications in ITS like route guidance, traffic control, and especially the congestion avoidance. Interestingly, the burst of map queries in a short duration normally indicates a "hotspot" being held at the targeting destination (the "hotspot" is "Fish Leong Concert" in Fig. 1). Therefore, the exploration of the hotspot from lots of crowd map queries brings a new way to explain the traffic jams.

We intend to make full use of crowd map queries in traffic speed prediction problems. However, there are two challenges to the integration of query and road traffic speed data: (1) **Spatiotemporal variation.** The raw crowd map queries targeting the same destination, like the Capital Gym, can be initiated

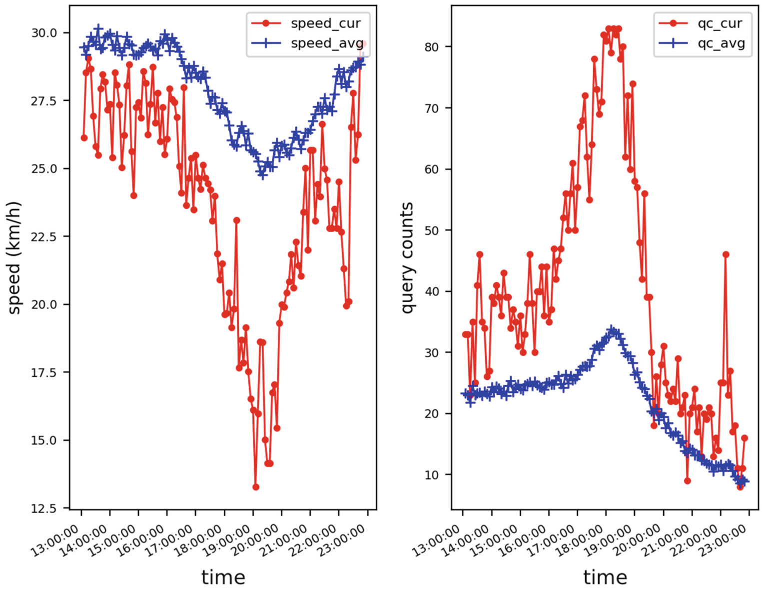

Fig. 1. The traffic speed (left) and crowd query counts (right) around the Capital Gym, Beijing on April 8, 2017. The red "dot" denotes the current traffic speed (query counts) while the blue "plus" represents the average historical traffic speed (query counts). At 19:00, there is the Fish Leong Concert in the Capital Gym. (Color figure online)

at different origins, at the different time and by individual users; (2) **Modality Gap.** The queries of traffic speed are from different modalities and have different distributions.

Enlighted by the idea that performance improvement can be achieved by properly integrating multi-modal information from various sources, this paper aims to predict future traffic speed with appropriate integration of current road traffic speed and crowd map queries. Technically, the contributions of this paper can be described in two aspects.

1. This paper proposes Multi-Seq2Seq-Att for hotspot traffic speed prediction. Multi-Seq2Seq-Att is a learning framework that deals with sequences of different modalities, (i.e., the query sequence and the traffic speed sequence). The main idea of Multi-Seq2Seq-Att attempts is to learn to fuse the multi-modal sequence with content attention. As a result, Multi-Seq2Seq-Att addresses the modality gap between the queries and the data of traffic speed.
2. Furthermore, the generality of Multi-Seq2Seq-Att makes it promising for many sequential multi-modal learning applications, such as the application of text and speech.

This paper is organized as follows: Sect. 2 presents the problem definition. Following that, in Sect. 3, we describe the proposed Multi-Seq2Seq-Att in detail.

Sect. 4 presents qualitative and quantitative results of different models. Finally, we make the conclusion in Sect. 5.

2 Problem Definition

Assume $\mathcal{L} = \{l^i | i = 1, 2, \ldots, K\}$ is a collection of K road segments, where a road segment is a section of a road. Let $\boldsymbol{v}^l = (v_1^l, v_2^l, \ldots, v_t^l)$ be the traffic speed of the road segment $l \in \mathcal{L}$, where v_t^l is a scalar that denotes the traffic speed of one specific road segment l at time t. Assume $\mathcal{Q} = \{q^i | i = 1, 2, \ldots, N\}$ is a corpus of N users' map query records. Each map query record q^i is defined by a triple $q^i = (t_s^i, s^i, d^i)$, which satisfies: (1) t_s^i is the starting time of query q^i; (2) s^i is the origin (source location) of q^i; (3) d^i is the destination. To simplify the problem, the superscript is removed without confusion in the remaining part of this paper.

Specifically, for the road segment l, given the previous traffic speed $V^p = (v_1^l, v_2^l, \ldots, v_t^l)$ and the query records $\mathcal{Q}$, our object is to maximize the conditional probability of observing the future traffic speed $V^f = (v_{t+1}^l, v_{t+2}^l, \ldots, v_{t+w}^l)$:

$$p_\theta(V^f | V^p, \mathcal{Q}) = \prod_{m=1}^{w} p_\theta(v_{t+m} | v_1, v_2, \ldots, v_{t+m-1}, \mathcal{Q}_{<=t}) \tag{1}$$

In the equation above, $\mathcal{Q}_{<=t} = \{q^i | t_s^i <= t\}$ and w is the prediction horizon and θ is a parameter. Given the map query records $\mathcal{Q}$, the previous traffic speed slot V^p and the future traffic speed V^f of K road segments, our training objective is to maximize the following log likelihood w.r.t. the model parameter θ:

$$\arg\min_\theta \quad -\frac{1}{K} \sum_{k=1}^{K} \log \; p_\theta(V^f | V^p, \mathcal{Q}) \tag{2}$$

3 Methods

Followed by our previous work [10], we address the spatiotemporal variation with hotspot discovery. In this section, we introduce the sequence to sequence model (Seq2Seq) and our multi-modal sequence to sequence learning with attention (Multi-Seq2Seq-Att) for hotspot traffic speed prediction.

3.1 Seq2Seq

For each hotspot, given the historical traffic speed slot $\boldsymbol{V}^c = (\boldsymbol{v}_{t-w+1}, \boldsymbol{v}_{t-w+2}, \ldots, \boldsymbol{v}_t)$ of selected k road segments, we aim to forecast their future traffic speed slot $\widehat{\boldsymbol{V}}^f = (\hat{\boldsymbol{v}}_{t+1}, \hat{\boldsymbol{v}}_{t+2}, \ldots, \hat{\boldsymbol{v}}_{t+w})$, where $\boldsymbol{v}_t = (v_t^1, v_t^2, \ldots, v_t^k)^T$ is the traffic speed of k road segments at time t, w is the prediction horizon.

As shown in Fig. 2, a sequence to sequence (Seq2Seq) network is applied to model the traffic speed. It consists of two LSTM [6] layers. The bottom LSTM layer (colored red) encodes information in the current traffic speed slot $\boldsymbol{V}^c$ while the second LSTM layer (colored green) decodes the encoding information of $\boldsymbol{V}^c$ to predict the future traffic speed slot $\widehat{\boldsymbol{V}}^f$.

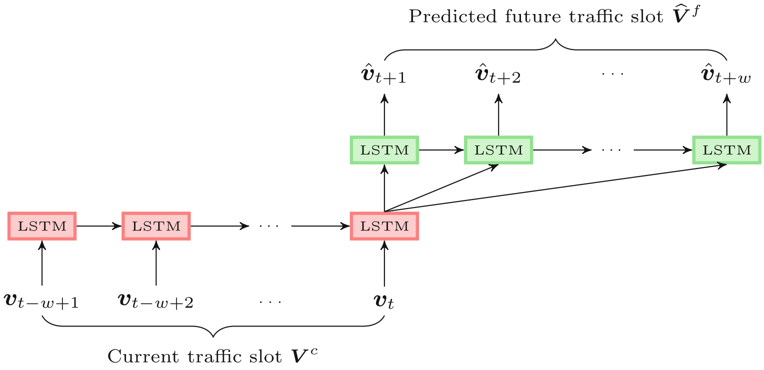

Fig. 2. The structure of the Seq2Seq network. (Color figure online)

3.2 Multi-Seq2Seq-Att

The map queries issued by users at a certain time can be utilized to foresee the traffic speed around the queried destination after a while. Considering a set of queries triggered earlier than 15:00 whose destinations are all around the Capital Gym, the arrival time towards the destinations can be estimated as [10]. The number of queries (ϕ_t) implies the information of how many individuals would arrive at the Capital Gym at the estimated arrival time, which would be beneficial to traffic speed prediction around the Capital Gym. Note that here only the map queries for the Capital Gym issued earlier than 15:00 is utilized to guarantee the foreseeable characteristic of map queries.

However, unleashing the power of the crowd map queries for traffic prediction is challenging due to the modality gap. To bridge the gap, we propose a multi-modal sequence to sequence framework with content attention as shown in Fig. 3. Motivated by [12], the idea of Multi-Seq2Seq-Att is to consider all the hidden states of the traffic encoder (self attention) and query encoder (multi-modal attention) when decoding the current target hidden state $\boldsymbol{h}_{t+1}$ through two content attention. Specifically, for self attention, given the current target hidden state $\boldsymbol{h}_{t+1}$, the traffic source hidden state $\bar{\boldsymbol{h}}_s^t$, the traffic context vector $\boldsymbol{c}_{t+1}^t$ is derived as follows:

$$\text{score}(\boldsymbol{h}_{t+1}, \bar{\boldsymbol{h}}_s^t) = \boldsymbol{h}_{t+1}^T \boldsymbol{W}_a^t \bar{\boldsymbol{h}}_s^t \tag{3}$$

$$\boldsymbol{a}_{t+1}^t = \text{self\_align}(\boldsymbol{h}_{t+1}, \bar{\boldsymbol{h}}_s^t) = \frac{\exp(\text{score}(\boldsymbol{h}_{t+1}, \bar{\boldsymbol{h}}_s^t))}{\sum_{s'} \exp(\text{score}(\boldsymbol{h}_{t+1}, \bar{\boldsymbol{h}}_{s'}^t))} \tag{4}$$

$$\boldsymbol{c}_{t+1}^t = (\boldsymbol{a}_{t+1}^t)^T \bar{\boldsymbol{h}}_s^t. \tag{5}$$

where W_a^t is the traffic attention weight matrix, a_{t+1}^t is the traffic alignment vector and c_{t+1}^t is the traffic context vector.

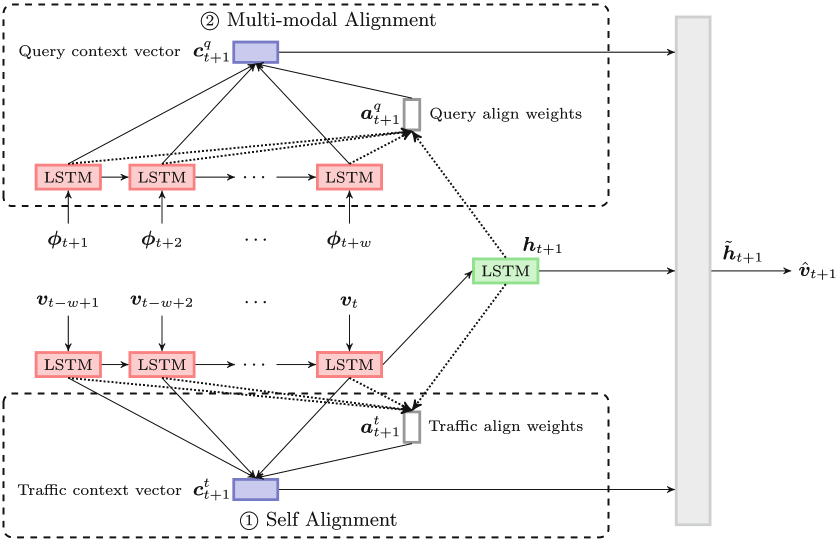

Fig. 3. The structure of the Multi-Seq2Seq-Att network. At time step t+1, the bottom (top) self (multi-modal) alignment infers an alignment weight vector $\boldsymbol{a}^t_{t+1}$ ($\boldsymbol{a}^q_{t+1}$) based on the current target state $\boldsymbol{h}_{t+1}$ and all source states $\bar{\boldsymbol{h}}^t_s$ ($\bar{\boldsymbol{h}}^q_s$). A global context vector $\boldsymbol{c}^t_{t+1}$ ($\boldsymbol{c}^q_{t+1}$) is then computed as the weighted average, according to $\boldsymbol{a}^t_{t+1}$ ($\boldsymbol{a}^q_{t+1}$), over all the source states.

For multi-modal attention, given the current target hidden state $\boldsymbol{h}_{t+1}$, the query source hidden state $\bar{\boldsymbol{h}}^q_s$, the query context vector $\boldsymbol{c}^q_{t+1}$ is derived as follows:

$$\text{score}(\boldsymbol{h}_{t+1}, \bar{\boldsymbol{h}}^q_s) = \boldsymbol{h}^T_{t+1}\boldsymbol{W}^q_a\bar{\boldsymbol{h}}^q_s \tag{6}$$

$$\boldsymbol{a}^q_{t+1} = \text{multi-modal\_align}(\boldsymbol{h}_{t+1}, \bar{\boldsymbol{h}}^q_s) = \frac{\exp(\text{score}(\boldsymbol{h}_{t+1}, \bar{\boldsymbol{h}}^q_s))}{\sum_{s'}\exp(\text{score}(\boldsymbol{h}_{t+1}, \bar{\boldsymbol{h}}^q_{s'}))} \tag{7}$$

$$\boldsymbol{c}^q_{t+1} = (\boldsymbol{a}^q_{t+1})^T\bar{\boldsymbol{h}}^q_s. \tag{8}$$

where W^q_a is the query attention weight matrix, a^q_{t+1} is the query alignment vector and c^q_{t+1} is the query context vector.

Given the target hidden state $\boldsymbol{h}_{t+1}$, the traffic context vector $\boldsymbol{c}^t_{t+1}$, and the query context vector $\boldsymbol{c}^q_{t+1}$, we utilize a simple concatenation layer to combine the information from all the vectors to produce an attentional hidden state as follows:

$$\tilde{\boldsymbol{h}}_{t+1} = \tanh(\boldsymbol{W}_c[\boldsymbol{h}_{t+1}; \boldsymbol{c}^t_{t+1}; \boldsymbol{c}^q_{t+1}]) \tag{9}$$

The attentional vector $\tilde{\boldsymbol{h}}_{t+1}$ is then fed into the linear-regression layer to produce the predicted traffic speed $\hat{\boldsymbol{v}}_{t+1}$ as:

$$\hat{\boldsymbol{v}}_{t+1} = \boldsymbol{W}_s\tilde{\boldsymbol{h}}_{t+1} \tag{10}$$

4 Experiments

4.1 Datasets

Data Pre-processing. Our experiments are implemented on two real-world datasets: (1) The traffic speed dataset used for traffic speed prediction. We collected the data from Baidu Map in the city of Beijing, China, starting on April 1, 2017 and ending on May 31, 2017. The dataset covers around 800,000 road segments. Traffic speed for each road segment is recorded per minute. Since it's a real word traffic speed dataset, traffic lights would have a huge impact on the traffic speed, which leads to dramatic traffic speed variance. In some cases, the traffic speed could differ as much as 20 km/h between two consecutive minutes due to the existence of traffic lights. To avoid the dramatic change in traffic speed, we average the traffic speed with a unit of 5 min and implement zero-phrase digital filtering [4] to make traffic speed smooth; and (2) map query dataset, which is normally used to discover hotspots and predict traffic speed. It is the query sub-dataset of the Q-Traffic dataset [10] and is collected from Baidu Map[1]. It includes around 114 million map queries calculated from April 1, 2017 to May 31, 2017, all in Beijing, China.

Correlation Analysis. The relation between traffic speed and query counts in the same spot is tested. For each spot, the average traffic speed of k ($k = 5$) adjacent road segments with a window of 5 min is used. And We also collect traffic speed together with map query counts data for all selected spots(hotspots). Rank correlation coefficient from Spearman is used in our test and the result $\rho = -0.57$ with a P-value$= 4.64 \times e^{-14}$ shows that a strong negative correlation exists between the query counts and the average traffic speed, which forecasts the potential to predict traffic speed statistics with our map query data.

4.2 Baselines

We compare our proposed model with the following methods.

- Random forests regression (RF) [9]: RF is a widely-used machine learning method for prediction and regression, which constructs a multitude of decision tree at training time and outputs the mean prediction of the individual trees;
- Support vector regression (SVR) [8]: SVR is a regression version of SVM, which is widely used in the area of traffic prediction;
- Gated recurrent unit (GRU) [2]: GRU is a variant version of RNN which has been widely used in regression and prediction tasks. We compare our model with the Seq2Seq model which consists of two GRUs;
- Seq2Seq: We compare our model with the basic Seq2Seq model which consists of two LSTMs;
- Seq2Seq_Att: Seq2Seq_Att is the Seq2Seq model with content attention;
- Init(Q): We use the encoding of the query to initialize the encoder of the Seq2Seq;

[1] http://map.baidu.com/.

- T⋆_Q⋆_OP: We compare our model with 11 $(2*2*3-1)$ variants T⋆_Q⋆_OP, where $\star = \{B, A\}$, $OP = \{Add, Stack, Att\}$ represents the fusion (add, stack or attention) of traffic (before or after encoding) and query (before or after encoding).

4.3 Evaluation Metrics

We implement two metrics to properly evaluate the performance of models proposed, which are mean absolute error and mean square error.

$$MSE = \frac{1}{T}\sum_{t=1}^{T}(\boldsymbol{v}_t - \hat{\boldsymbol{v}}_t)^2 \tag{11}$$

$$MAE = \frac{1}{T}\sum_{t=1}^{T}|\boldsymbol{v}_t - \hat{\boldsymbol{v}}_t| \tag{12}$$

where $\boldsymbol{v}_t$ and $\hat{\boldsymbol{v}}_t$ are the traffic speed ground truth and predicted speed at time t, respectively.

4.4 Results

Firstly, We show the effectiveness of our proposed Multi-Seq2Seq-Att for traffic prediction. Figure 4 demonstrates the performance, measured by MSE and MAE, of different methods. We can observe from the image that deep learning based methods (e.g. Seq2Seq and Multi-Seq2Seq-Att) outcompete many traditional methods like RF and SVR. It is observed that Seq2Seq model is slightly better than GRU, the major difference which separates these two is the sequence layer (one is LSTM, the other is GRU), based on that, we select LSTM as sequence

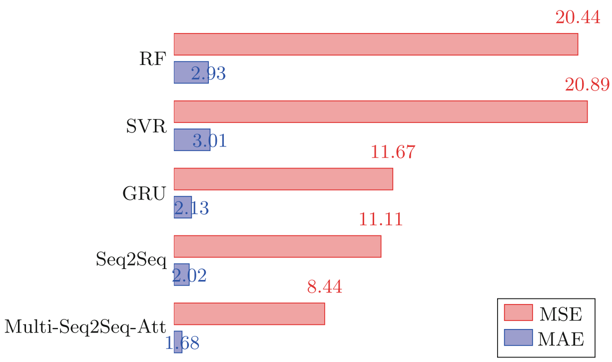

Fig. 4. Best performance of different methods for traffic speed prediction. Lower MSE (MAE) means better performance. (Best viewed in the electronic version)

layer in our final Seq2Seq network. Among all deep learning methods, our Multi-Seq2Seq-Att has the best performance. Compared with Seq2Seq, Multi-Seq2Seq-Att achieves 24% and 17% relative performance increase measured by MAE and MSE, respectively. This improvement may be explained in a way that our model better leverages the query features and the attention model.

Secondly, We show the effectiveness of the query and the attention mechanism. Table 1 compares the performance of all variants with the different number of hidden units. The performances of Seq2Seq_Att and T⋆_Q⋆_Att are better than that of Seq2Seq, which demonstrate the effectiveness of the attention mechanism and the query. In addition, T⋆_Q⋆_Att model outperforms all the T⋆_Q⋆_Add and T⋆_Q⋆_Stack, showing the power of attention. Moreover, the fusion after encoding is better than the fusion before encoding (raw input). As a result, our Multi-Seq2Seq-Att network which considers both the attention mechanism and the map query achieves the best performance. Furthermore, as we increase the number of hidden units, the performance will improve accordingly. However, as the hidden units number exceeds 32, the performance decreases.

Table 1. Comparison of all variants and our Multi-Seq2Seq-Att with different number of hidden units. The results with the best performance are marked in bold.

Methods	Num hidden units = 8		Num hidden units = 16		Num hidden units = 32	
	MSE	MAE	MSE	MAE	MSE	MAE
Seq2Seq	14.34	2.41	11.11	2.02	11.32	2.02
Seq2Seq_Att	15.10	2.51	10.14	2.02	9.83	1.87
Init(Q)	12.87	2.36	9.89	1.88	9.82	1.71
TB_QB_Stack	12.83	2.32	11.27	2.05	9.92	1.93
TB_QB_Add	14.20	2.51	13.69	2.43	12.92	2.28
TB_QB_Att	12.34	2.27	10.33	2.03	9.62	1.86
TB_QA_Stack	12.55	2.31	10.51	2.06	9.15	1.84
TB_QA_Add	12.61	2.29	11.29	2.07	9.41	1.85
TB_QA_Att	12.24	2.27	10.09	2.02	9.72	1.89
TA_QB_Stack	12.32	2.27	11.80	2.02	9.92	1.89
TA_QB_Add	13.50	2.38	12.28	2.17	9.54	1.85
TA_QB_Att	11.72	2.22	8.59	1.76	8.84	1.68
TA_QA_Stack	13.90	2.46	8.76	1.82	8.87	1.71
TA_QA_Add	13.73	2.42	8.70	1.85	9.41	1.71
Multi-Seq2Seq-Att	11.60	2.21	10.03	1.99	**8.44**	**1.68**

5 Conclusion

We study the problem of how to model map queries and to fully utilize them to assist traffic speed prediction. In this paper, we propose Multi-Seq2Seq-Att which integrates the sequence learning from different modalities. The attention part of Multi-Seq2Seq-Att fuses the information of map queries and traffic speed, addressing the modality gap between the map queries and traffic speed. As a result, Multi-Seq2Seq-Att achieves 24% relative improvement over other state-of-the-art methods on our datasets fetched from Baidu Map.

Acknowledgments. This work was supported in part by 973 program (No. 2015CB352302, 2015CB352300), the National Natural Science Foundation of China (Nos. 61625107, 61751209, U1611461), the Key Program of Zhejiang Province, China (No. 2015C01027) and Chinese Knowledge Center of Engineering Science and Technology (CKCEST).

References

1. Ahmed, M.S., Cook, A.R.: Analysis of freeway traffic time-series data by using Box-Jenkins techniques. No. 722 (1979)
2. Chung, J., Gulcehre, C., Cho, K., Bengio, Y.: Empirical evaluation of gated recurrent neural networks on sequence modeling. arXiv preprint arXiv:1412.3555 (2014)
3. Davis, G.A., Nihan, N.L.: Nonparametric regression and short-term freeway traffic forecasting. J. Transp. Eng. **117**(2), 178–188 (1991)
4. Gustafsson, F.: Determining the initial states in forward-backward filtering. IEEE Trans. Sig. Process. **44**(4), 988–992 (1996)
5. He, J., Shen, W., Divakaruni, P., Wynter, L., Lawrence, R.: Improving traffic prediction with tweet semantics. In: IJCAI, pp. 1387–1393 (2013)
6. Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural Comput. **9**(8), 1735–1780 (1997)
7. Huang, W., Song, G., Hong, H., Xie, K.: Deep architecture for traffic flow prediction: deep belief networks with multitask learning. IEEE Trans. Intell. Transp. Syst. **15**(5), 2191–2201 (2014)
8. Jin, X., Zhang, Y., Yao, D.: Simultaneously prediction of network traffic flow based on PCA-SVR. In: Liu, D., Fei, S., Hou, Z., Zhang, H., Sun, C. (eds.) ISNN 2007. LNCS, vol. 4492, pp. 1022–1031. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-72393-6_121
9. Leshem, G., Ritov, Y.: Traffic flow prediction using adaboost algorithm with random forests as a weak learner. Proc. World Acad. Sci. Eng. Technol. **19**, 193–198 (2007)
10. Liao, B., Zhang, J., Wu, C., McIlwraith, D., Chen, T., Yang, S., Guo, Y., Wu, F.: Deep sequence learning with auxiliary information for traffic prediction. In: Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM (2018)
11. Liu, X., Kong, X., Li, Y.: Collective traffic prediction with partially observed traffic history using location-based social media. In: Proceedings of the 25th ACM International on Conference on Information and Knowledge Management. pp. 2179–2184. ACM (2016)
12. Luong, T., Pham, H., Manning, C.D.: Effective approaches to attention-based neural machine translation. In: Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing, pp. 1412–1421 (2015)
13. Lv, Y., Duan, Y., Kang, W., Li, Z., Wang, F.Y.: Traffic flow prediction with big data: a deep learning approach. IEEE Trans. Intell. Transp. Syst. **16**(2), 865–873 (2015)

14. Ma, X., Tao, Z., Wang, Y., Yu, H., Wang, Y.: Long short-term memory neural network for traffic speed prediction using remote microwave sensor data. Transp. Res. Part C: Emerg. Technol. **54**, 187–197 (2015)
15. Tran, Q.T., Ma, Z., Li, H., Hao, L., Trinh, Q.K.: A multiplicative seasonal arima/garch model in evn traffic prediction. Int. J. Commun. Network Syst. Sci. **8**(04), 43 (2015)
16. Williams, B.M., Hoel, L.A.: Modeling and forecasting vehicular traffic flow as a seasonal arima process: theoretical basis and empirical results. J. Transp. Eng. **129**(6), 664–672 (2003)
17. Zhang, J., Wang, F.Y., Wang, K., Lin, W.H., Xu, X., Chen, C.: Data-driven intelligent transportation systems: a survey. IEEE Trans. Intell. Transp. Syst. **12**(4), 1624–1639 (2011)