CA-RPT: Context-Aware Road Passage Time Estimation for Urban Traffic

Ying Liu[1,2], Zhenyu Cui[1,2][0000−0001−9857−2178], Tianlin Zhang, Jiaxu Leng, Weihong Xie[3], and Liang Zhang[4]

[1] School of Computer Science and Technology, University of Chinese Academy of Sciences, Beijing, 100190 China
[2] Key Lab of Big Data Mining and Knowledge Management, Chinese Academy of Sciences, Beijing, 100190 China
[3] School of Economics and Commerce, Guangdong University of Technology, Guangzhou, 510006 China
[4] School of Applied Mathematics, Guangdong University of Technology, Guangzhou, 510006 China

Abstract. Road passage time is an important measure of urban traffic. Accurate estimation of road passage time contributes to the route programming and the urban traffic planning. Currently, the estimation of road passage time for a particular road is usually based on its historical data which is simple to express the general law of road traffic. However, with the increase of the number of roads in the urban area, the connection between roads becomes more complex. The existing methods fail to make use of the connection between different roads and the road passage time, merely based on its own historical data. In this paper, we propose a road passage time estimating model, called "CA-RPT", which utilizes the contextual information between road connections as well as the date and time period. We evaluate our method based on a real geolocation information data set collected by mobile APP anonymously. The results demonstrate that our method is more accurate than the state-of-the-art methods.

Keywords: Road Passage Time · contextual information · road connection

1 Introduction

In recent years, the introduction of smart transportation has gradually changed the manner people travel. One of the widely used applications, in modern urban, is the estimation of road passage time. Accurately estimating road passage time is an important measure in guiding people's travel and avoiding the potential high blockage.

Road passage time estimation, which can reveals the congestion, is an important, but challenging task in smart traffic and has been widely used in travel time estimation [1] [2] which is of great importance for real-time traffic monitoring [3], driving directions choice [4] and transportation resource scheduling [5].

ICCS Camera Ready Version 2019
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-22750-0_64

Existing solutions, e.g., based on support vector regression [6] or spare flow capacity on the concerned road links [7], focus on the information of the road itself including the historical data rather than the contextual information from the connections between different roads.

Road passage time is affected by factors, such as the length, width, and the grade of the road, etc. In general, it takes longer time to travel on a longer, narrower, and lower-grade road under the same natural conditions and human activities.

However, all of the above are the objective characteristics of the road itself. Due to human activities, the judgment of the passage time of a road is inseparable from the time slots and date. Most of the previous studies used the historical data of a road to predict its future behavior, which uses a model to fit the recent or the historical data of a single node, and applies the given model to predict the upcoming road passage time.

In modern urban transportation, the connections between roads are more and more complicated and multiple factors may impact the road patency. The upstream and downstream, the number of adjacent roads, the congestion level of a road may affect the passage time estimation more or less. Therefore, the passage time estimation of a road is not a simple prediction with its own historical data or objective property. However, existing methods failed to utilize this part of the information, but by fitting historical data on a single road to fit the law of road passage time. The generalization ability of this approach is limited. Because it can only predict according to a general law, and does not have the ability to estimate the passage time based on temporary accidents, congestion, etc. which is occurring on some roads. Considering the influence of the following factors, we would like to use the road contextual information to finish the estimation.

Based on previous research, contextual information has been proven to be important in the field of psychology [8] [9] [10] and computer vision [11] [12] [13], which has also been well recognized for years that the contextual information is helpful in object recognition [14] [15] [16]. Inspired by the use of contextual information, this paper would like to propose an estimation model creatively, called CA-RPT which combines the road information with the contextual information into a novel structure thus to improve the prediction accuracy. The rest of the paper is organized as follow: Section.2 explain our approach; Section.3 is the experimental result and evaluation; Section.4 is the conclusion of this paper.

2 Our Approach

In this section, we will present our approach, including the overall architecture and the modeling process.

2.1 Overview

The architecture of the proposed CA-RPT is presented in Figure.1.

ICCS Camera Ready Version 2019
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-22750-0_64

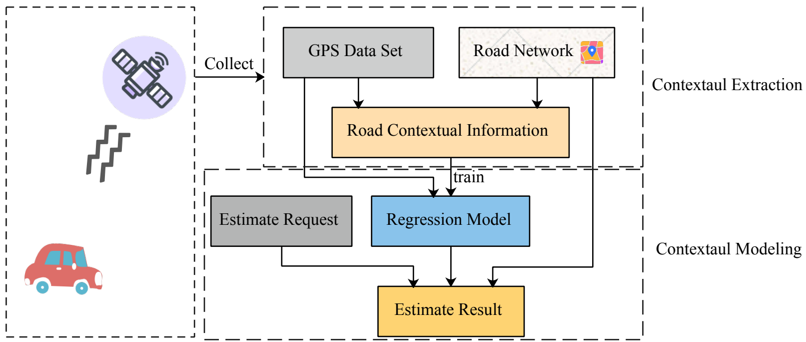

Fig. 1. The Architecture of CA-RPT.

CA-RPT is consist of two parts. The first part, called "Contextual Extraction", using GPS data, which collected in real time and processed into the passage time of each road, and road network to extract contextual information. The second part, which called "Contextual Modeling", use the former extracted contextual information and original GPS data set to fit a model, which utilizes the historical road passage time to estimate the one in the next few moments. The process of "Contextual Modeling" could be considered as a regression model, which use the road passage time for a period of time before a certain time point to predict the possible road passage time in the future. Since the process of contextual extraction does not add new dimensions, we could take the existing appropriate regression method for prediction.

In the following, we will explain the two parts of CA-RPT in turn.

2.2 Contextual Information

Inspired by contextual information for image recognition, we believe that the same approach can be used in passage time estimation. Before contextual extraction, we should firstly define some context concepts in traffic to be mentioned as follows. Similarly, we first define the context of the road as shown in Equation.1.

r_i: a road wait to be estimated
r_i': a road adjacent to r_i.
$F_(r_i)$: the sum of the contextual information of the road r_i.
$\psi\ (r_i')$: the road passage time of r_i'.
$\lambda\ _i$: the impact scale factor ($\lambda\ _i \in (0,1)$).

$$F(r_i) = \sum_{r_i'(road\ adjacent\ to\ r_i)} \lambda_i * (\psi(r_i') + F(r_i')) \tag{1}$$

ICCS Camera Ready Version 2019
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-22750-0_64

The equation indicates that all the information available to a road consists of two parts, the road passage time and the contextual information obtained from its neighboring roads. Compared with existing methods, which only use the historical data of one road, our approach adds the road's own historical time to the adjacent road multiply by the coefficient λ. This formula is obviously a recursive formula. In order to use the context information, we must first find a simple way of obtaining contextual information. Hence we get partial derivative of Equation.1 with respect to variable r_i' and obtain the Formula.2, which shows that the contextual information of each road can be obtained by multiplying the passage time of each adjacent road by the influence factor separately.

$$\frac{\partial F(r_i)}{\partial r_i'} = \lambda_i * \psi'(r_i') + \lambda_i * F'(r_i') \tag{2}$$

Based on the above facts, the contextual information of different roads can be added independently. Thus, for each r_i' in the Equation.1, we can put them into the same equation recursively. Then we can recursively get the Equation.3, where $\omega_{r_i'}$ represents the adjacent trajectory between r_i and r_i'.

$$F(r_i) = \sum_{r_i' \in (roads)} \psi(r_i') * \prod_{r_j' \in \omega_{r_i'}} \lambda_j \tag{3}$$

According to Equation.3, the contextual information of each road is the accumulation of the road passage time obtained by a series of roads adjacent to it. The contextual information acquired by each street is multiplied by the impact factor on its corresponding adjacent trajectory, which are obviously exponentially related, to form the contribution of the road to be estimated. In this way, we get complete context information for a road, which includes contextual information of all the other roads.

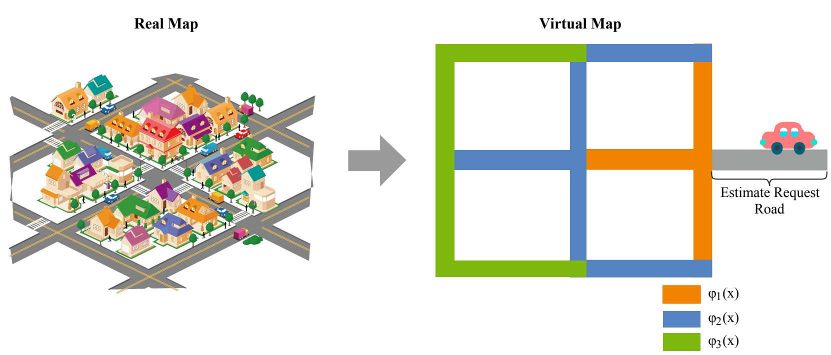

Fig. 2. Contextual information based on distance.

ICCS Camera Ready Version 2019
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-22750-0_64

For an intuitive view, as shown in Figure.2, $\phi_i(x)$ represents the sum of contextual information of road x, and i represent there are i roads belong to the trajectory from each road to road x.

2.3 Contextual Extraction

Based on the conclusions of the previous section, we perform contextual extraction to obtain contextual information for a road to assist us in making road passage time estimation. Obviously, for each road, the cumulative impact of other roads on it is unique and independent. Therefore, the core of contextual extraction is to calculate the sum of contextual information for each road. For a road to be estimated, we design an algorithm to mine the contextual information on its adjacent roads, called "Contextual Information Extraction in Traffic Network"(CIETN), as shown in Algorithm.1. Here, an additional parameter must be used, called Dep, to limit the number of ϕ_i we calculate.

Algorithm 1: CIETN

Input: The index of road to be estimated(i);The depth of context-aware information extraction(Dep).
Output: The contextual information extracted by CIETN(Σ).

Function *CIETN(Index i,Deep Dep)*
  $List_A \leftarrow$ Push(i,$\lambda = 1$);
  $Visit_{r_i}$ = True;
  while $Dep \geq 0$ **do**
    Dep = Dep-1;
    for *each* (i,λ) *in* $List_A$ **do**
      $\phi_i \leftarrow \lambda * RoadPassageTime\ in\ History(r_i)$;
      $\Sigma \leftarrow \Sigma + \phi_i$;
      for $r_j adjacenttor_i$ **do**
        if *(r_j) and (r_i to r_j) is Not Visited* **then**
          $List_B \leftarrow Push(j, \lambda * \lambda_j)$;
    $Visit_{each(r_i)}$ = True;
    $Visit_{each(r_i\ to\ r_j)}$ = True;
    $List_A \leftarrow List_B$;
    clear($List_B$);
  return Σ;

Algorithm1 is the core algorithm of contextual extraction, which uses hierarchical traversal to calculate the contextual information on the road to be extracted. There are three points that need special explanation. For the first one, the additional parameter, Dep, is necessary based on the fact that if the distance between two roads is far enough, the mutual impact can be ignored. In other words, we must only extract contextual information within a certain distance from the road to be extracted. Secondly, the parameter λ is calculated based on the basic properties of the road, including the length, width and level

ICCS Camera Ready Version 2019
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-22750-0_64

of the road. That is, the λ of a road is fixed and does not change over time. Last but not least, when we get the passage time of a road, we get historical data for a period of time. These fixed-length historical data are processed by the same sampling process and become a fixed-length one-dimensional tensor, which makes the contextual information Σ returned by the algorithm also a one-dimensional tensor. And this tensor is consistent with the length of the original data.

The complexity of this algorithm is O(N^{Dep}), where N represent the connected roads and Dep represent the depth of extraction, but it is far from this level in practical applications.

Considering the crossing of modern cities, there are often only four roads at the junction. As a result, the N in the above paragraph is small. Moreover, we often set the Dep to a number within 5, which ensures that the time cost of the algorithm is small.

After extraction algorithm CIETN, contextual information for each road, which will be used in contextual modeling along with the original road historical passage time.

2.4 Contextual Modeling

The task of contextual modeling is to train a predictor to predicts future passage time of a road. According to experience, road passage time is likely to show a trend in different time periods and different dates. In order to verify our conjecture, we first randomly selected several roads with different timing characteristics for analysis.

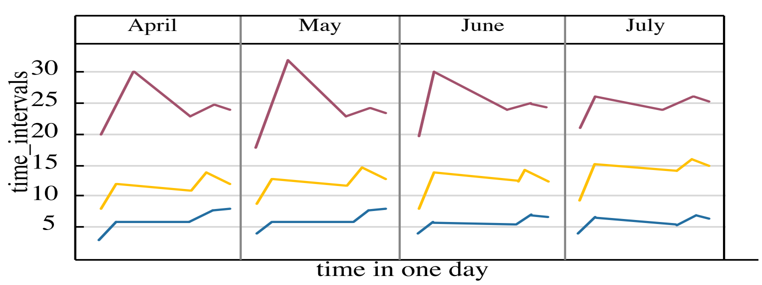

Fig. 3. Contextual passage time statistics on partial roads in different months in average.

DOI: 10.1007/978-3-030-22750-0_64

As shown in Figure.3, the ordinate represents the contextual passage time of a road in history, and its unit is 10 seconds while the abscissa represents the time in one day from day to night.

After cluster analysis, we found that the contextual passage time of most roads show a similar trend within single hour, however, what is not obvious is the information to show trend of the roads between the months. So we consider further subdividing the data to mine more potential features that can be distinguished.

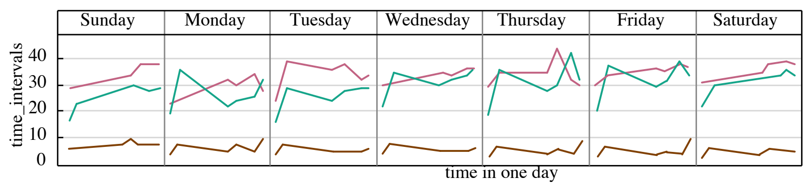

Fig. 4. Contextual passage time statistics on partial roads in different months.

Then, in each week, we clustered each of the randomly selected streets, as shown in Figure.4. The results show that the contextual passage time of the same road is relatively obvious due to the factors of the week. Therefore, different days of a week should be considered separately when establishing a regression model during the contextual modeling phase.

What's more, during the cluster analysis experiment, we also found that the contextual passage time of the road in one week day are often grouped into one class, which shows a strong correlation between road transit time and time. Therefore, we consider using the LSTM(Long Short Term Memory Network) model [17] to solve this regression problem, which is applied in multiple time-related regression problems successfully [18] [19] [20]. LSTM changes the weight of the self-loop by increasing the input threshold, forgetting the threshold and outputting the threshold. In this way, when the model parameters are fixed, the integral scale at different times can be dynamically changed, thereby avoiding the problem of gradient disappearance or gradient expansion.

In our approach, we use the original road passage time data and the road contextual information extracted by contextual extraction to train on the LSTM network. These two parts of the data are of the same dimension and the same length, so our modeling process can be easily performed on the LSTM network.

ICCS Camera Ready Version 2019
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-22750-0_64

3 Experimental Result and Evaluation

To verify the effectiveness of our approach, we applied our algorithm in a public data set. This data set is provided by the Guizhou Provincial Big Data Development Authority which is composed of the urban road network information and the geographical location information of users, which collected anonymously by the mobile APP in real time. These raw data are processed and merged to form traffic information for the city at full time without blind spots. It is composed of three data sets, which described by three tables: table.1, table.2, table.3.

Table 1. The description of data set.1.

Attribute Name	Type	Tip
link_ID	string	Unique identifier for each road segment(link)
length	double	Length of the link(m)
width	double	Width of the link(m)
link_class	int	The level of link, e.g., number 1 represents main road

Table 2. The description of data set.2.

Attribute Name	Type	Tip
link_ID	string	Unique identifier for each road segment(link)
in_links	string	Direct upstream link of the link, link_IDs were split by #
out_links	string	Direct downstream link of the link, link_IDs were split by #

Table 3. The description of data set.3.

Attribute Name	Type	Tip
link_ID	string	Unique identifier for each road segment(link)
date_time	date	Date, e.g., '2015-10-01'
time_interval	string	Time_interval, e.g., [2015-09-01 00:00:00, 2015-09-01 00:00:10]
travel_time	double	Average passage time of the vehicles on the road(s)

Table.1 is the attribute data set of road (link), which is a data set describing the connection of major roads in the city. Wherein, each traffic direction of each road is composed of a plurality of road segments(links), and the data set provides a unique identifier, length, width, and road type of each link. Table.2 is a link upstream and downstream connection data set, which is a data set describing the connection between links. The links includes the upstream and downstream connection according to the direction in which the vehicle is allowed to pass. The data set includes the direct upstream link and the direct downstream link of each

ICCS Camera Ready Version 2019
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-22750-0_64

link. Table.3 is a link historical passage time data set. The data set records the average passage time on each link in different time periods of the history (2 min is a time period). The average passage time of each time period is determined by the time period when the vehicle entering the link.

The data set we use provides a total of 132 links of static information, as well as upstream and downstream connections between these links which also includes the average passage time of each road on each time period from March 2016 to May 2016, March 2017 to June 2017 and July 2016.

The evaluation index we use is MAPE(mean absolute percent error). Generally, the lower the MAPE value, the higher the accuracy of the model, such as the Formula.4, where ttp means the estimate value, ttr means the real value, N means the number of estimate value, T_i means the number of estimate time intervals in link i.

$$MAPE = \frac{sum_{i=1}^{N} sum_{j=1}^{T_i} |ttp_{i,j} - ttr_{i,j}|}{sum_{i=1}^{N} T_i} \tag{4}$$

We randomly selected the time period to make predictions and ensure that there is sufficient data for training. In the experiment, we used some common methods to make estimation and use the model(LSTM) which has the best performance in estimation accuracy as our regression model for contextual modeling. The experimental results are shown in Table.4.

Table 4. The MAPEs of different methods.

Method	MAPE
Linear Regression	0.5451
Logistic Regression	0.3324
Xgboost	0.3159
LSTM	0.2881
Context-Aware LSTM	0.2762

The experiment result shows that the MAPE value of our approach decreased by 0.0119 compared with the most accurate existing method, which shows that our context-aware modeling method can effectively improve the accuracy of estimation which is better than any existing method.

4 Conclusion

In this article, inspired by the successful application of contextual mining in multiple domains, we proposed a road passage time estimation framework(CA-RPT) whose process is divided into two parts, contextual extraction and contextual modeling. The former part use historical data of roads and road network information to extract contextual information in traffic. Based on these information, the later one make estimation by context-aware modeling. Experiments

ICCS Camera Ready Version 2019
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-22750-0_64

show that the road passage time estimated by CA-RPT is more accurate than existing methods, by obtaining and effectively utilizing road contextual information. What is more, in the experiment, the λ and the depth of contextual extraction(Dep) we used is an empirically based constant, which does not necessarily produce the best results in our modeling process. It can be expected that if these constant can be trainable, the estimation accuracy might be further improved.

5 Acknowledgement

This project was partially supported by Grants from Natural Science Foundation of China 71671178, 9154620, and also supported by Guangdong Provincial Science and Technology Project 2016B010127004.

References

1. Wang Y , Zheng Y , Xue Y . Travel time estimation of a path using sparse trajectories[J]. 2014.
2. Zheng Wang , Kun Fu , Jieping Ye . Learning to Estimate the Travel Time. Proceedings of the 24rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, 2018.
3. Chawla S , Zheng Y , Hu J . Inferring the Root Cause in Road Traffic Anomalies[C]// IEEE International Conference on Data Mining. IEEE, 2013.
4. Jing Y, Yu Z, Zhang C, et al. T-drive:driving directions based on taxi trajectories[C]// Sigspatial International Conference on Advances in Geographic Information Systems. 2010.
5. Yuan N J , Zheng Y , Zhang L , et al. T-Finder: A Recommender System for Finding Passengers and Vacant Taxis[J]. IEEE Transactions on Knowledge and Data Engineering, 2013, 25(10):2390-2403.
6. Wu C H , Ho J M , Lee D T . Travel-time prediction with support vector regression[J]. IEEE Transactions on Intelligent Transportation Systems, 2004, 5(4):276-281.
7. Liang Z , Wakahara Y . Real-time urban traffic amount prediction models for dynamic route guidance systems[J]. EURASIP Journal on Wireless Communications and Networking, 2014, 2014(1):85.
8. Bar M. Visual objects in context.[J]. Nature Reviews Neuroscience, 2004, 5(8):617-629.
9. Oliva A , Torralba A . The role of context in object recognition[J]. Trends in Cognitive Sciences, 2007, 11(12):0-527.
10. Palmer E. The effects of contextual scenes on the identication of objects[J]. Memory Cognition, 1975, 3: 519-526.
11. Carbonetto P, De Freitas N, Barnard K. A statistical model for general contextual object recognition.[J]. Lecture Notes in Computer Science, 2004, 3021:350-362.
12. Chen L C , Papandreou G , Kokkinos I , et al. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2016, 40(4):834-848.

ICCS Camera Ready Version 2019
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-22750-0_64

13. Galleguillos C , Belongie S . Context based object categorization: A critical survey[J]. Computer Vision and Image Understanding, 2010, 114(6):712-722.
14. Divvala S K , Hoiem D , Hays J H , et al. An empirical study of context in object detection[J]. 2009.
15. Galleguillos C , Rabinovich A , Belongie S . Object categorization using co-occurrence, location and appearance[C]// IEEE Conference on Computer Vision Pattern Recognition. IEEE, 2008.
16. Leng J , Liu Y . An enhanced SSD with feature fusion and visual reasoning for object detection[J]. Neural Computing and Applications, 2018.
17. GERS, Felix A, SCHMIDHUBER, et al. Learning to forget : Continual prediction with LSTM[J]. Neural Computation, 2000, 12(10):2451-2471.
18. Duan Y, Lv Y, Wang F Y. Travel time prediction with LSTM neural network[C]// IEEE International Conference on Intelligent Transportation Systems. 2016.
19. Selvin S, Vinayakumar R, Gopalakrishnan E A, et al. Stock price prediction using LSTM, RNN and CNN-sliding window model[C]// International Conference on Advances in Computing. 2017.
20. Marchi E, Ferroni G, Eyben F, et al. Multi-resolution linear prediction based features for audio onset detection with bidirectional LSTM neural networks[C]// IEEE International Conference on Acoustics. 2014.

ICCS Camera Ready Version 2019
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-22750-0_64