IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 24, Issue 3, Ser. II (May. −June. 2022), PP 25-30
www.iosrjournals.org

Urban Road Traffic Link Travel Time Estimation Based on Sparse Data

Vu Huy Luong[1], Nguyen Thi Mai Phuong[2], Nguyen Kim Son[3]
Thai Nguyen University of Information and Communication Technology, Vietnam ([1]vhluong@ictu.edu.vn, [2]ntmphuong@ictu.edu.vn, [3]nkson@ictu.edu.vn)

Abstract - *Travel time data is collected at irregular intervals and is less frequent for links on non-major roads. For many periods of time, in a particular link, there may not be any current travel time data available. This is a problem for any model that uses recent travel time as an input variable. A neighbouring link inference method is proposed which uses the sparse data from nearby links to infer the designated link. The motivation of the method is to provide accurate travel time in near real-time for links, which do not have recent travel time data, to support the development of a large-scale traffic model. The method utilises Feed Forward Back-propagation Neural Networks (FFB-ANNs) to learn the relationship between neighbouring links from sparse historic travel time data. The relationships between neighbouring links in urban areas are complex because traffic light signal cycles usually heavily affect travel times. Additionally, in particular, on minor links, there is little travel time data available. Results show that FFB-ANNs are capable of estimating the travel time on major roads. However, even on non-major roads with more sparse data, the performance of the proposed FFB-ANNs estimators shows promising results.*
Keywords *–*

Date of Submission: 01-06-2022 Date of Acceptance: 13-06-2022

I. Introduction

Travel delays due to traffic congestion cause a huge waste of money increased human stress and unsafe traffic situations. They also increase negative environmental and societal side effects [1]. Congestion can be defined as the traffic demand exceeding the roadway capacity. In urban areas, transportation infrastructure development is constrained by land resource availability and financial limit [2]. In order to deal with growth, advanced dynamic traffic management systems are needed to manage existing transportation systems efficiently. Such systems require highly efficient and dynamic models. Real-time traffic information which comes from traffic models can be used to optimise signal control settings [3] and to help commuters avoid traffic congestion. A valuable and objective type of traffic information is the travel time [1].

There are two main sources to collect travel time data. They are moving observers and stationary observers. The stationary observers, for example, video surveillance, can produce travel time data at regular and frequent intervals for a particular highway or major road. In contrast, the moving observers include floating cars and probe cars, which use the Global Navigation Satellite System (GNSS) to trace positions from actual cars across the entire traffic network, which can produce travel time at irregular and less frequent intervals. For many periods of time, in a particular link, there may not be any current travel time data available [4, 5]. Especially, travel time data are more sparse on links of non-major roads. This is a problem for any model that uses recent travel time as an input variable.

We propose a method for dealing with sparse travel time data called a neighbouring link inference method. It utilizes Feed Forward Back-propagation Neural Networks (FFB-ANNs) to learn the traffic urban links' relationship from sparse historic travel time data. After the learning process, it is able to use travel time data from nearby links to estimate accurately near real-time travel times of designated links. We found that data sparse rate is an important factor that mainly affects the performance of the model.

II. Related Works

The stationary observers can produce real travel time data at regular and frequent intervals [4]. However, perhaps because stationary traffic observers are more expensive [5, 6] and are therefore only available on some particular highway or arterial road. In contrast, the moving observers can produce travel time at irregular and less frequent intervals. Because they use GNSS to trace positions from actual cars across the entire traffic network, they can cover almost any link as needed [7].

Travel time data collected from moving observers is naturally less frequent for links on non-major roads [4]. It means that for many periods of time, in a particular link, there may not be any current travel time data available. This is a problem for any model that uses recent travel time as an input variable.

The travel times on freeways regularly show relatively low variability, especially in congested conditions. They mainly depend on the geometrical characteristics of freeways, such as the number of ramps weaving sections per unit of road length [8]. In contrast, urban travel times can reveal very high variability caused by traffic light signal cycles as queuing delays usually heavily affect travel time. Signalised and unsignalised intersections, different speed conditions, public transport blockage, pedestrian and cyclist disturbance, transit priority, and parking also often affect travel time [1, 5]. High-variability urban travel time is a barrier to build an accurate travel time model.

There have been numerous methods for modeling travel time. They are Linear Regression [9, 10], Queue Theory [11-13], Bayesian Inference [14], nearest neighbourhood [15, 16], Neural Network [17-19], Support Vector Machine [4], Fuzzy Logic [20, 21], combination of mathematical and computation intelligence [3, 1] and Bayesian Network [1, 22, 23]. Most of these studies were less attention to urban travel than free-way travel time. Urban travel times are complex because traffic light signal cycles usually heavily affect travel times. These studies also did not focus to model large-scale traffic networks.

Few researchers have focussed on the challenges presented by sparse and missing travel time data where many periods of time, in a particular link, there may not be any current travel time data available. This is a problem for any model that uses recent travel time as an input variable. We propose a method for dealing with this problem called a neighbouring link inference method, which utilizes Feed Forward Back-propagation Neural Networks to learn urban traffic links' relationship from sparse historic travel time data. Accurate near real-time travel times of the designated link are estimated by nearby links' travel times.

The kriging method, which is used to predict travel time based on sparse floating car data, has been proposed in [24]. The prediction relies on the feature that nearby points will have almost the same travel time. Hence, travel time of a route can be estimated from nearby routes. The kriging method shows that it can deal with sparse floating car data. However, it is only able to predict travel time for long routes. The work presented in [4] also used information from nearby links to predict travel time called geospatial inference. Their model is a Support Vector Machine model that relies on floating car data, a time series data, from specific nearby links to predict future travel time data in a selected link. However, they were not deal with the problem of sparse data. The results show that the geospatial predictor performs better in greater congestion [4].

The main characteristics of our work that differentiate it from other and past work are:

- The use of sparse historic travel time data.
- The ability to estimate the travel time from nearby link travel time data.
- The ability to apply in a large-scale traffic network.

III. Neighbouring Link Inference Method

Assume that the traffic parameters of a link potentially have a correlation with traffic parameters of directly connected links. This correlation can be used to estimate travel times of a link using travel times of other links.

In order to investigate the assumption, a FFB-ANNs model has been selected. We have chosen Feed forward ANN with Backpropagation learning algorithm because it is widely used in solving various classification, estimation, and forecasting problems [17-19].

Forty links are randomly chosen in the first implementation. They consist of twenty major links including Motorway, A or B links, and twenty minor links which are divided as Classified unnumbered and Unclassified links. Link types are defined by Department for Transport in [25].

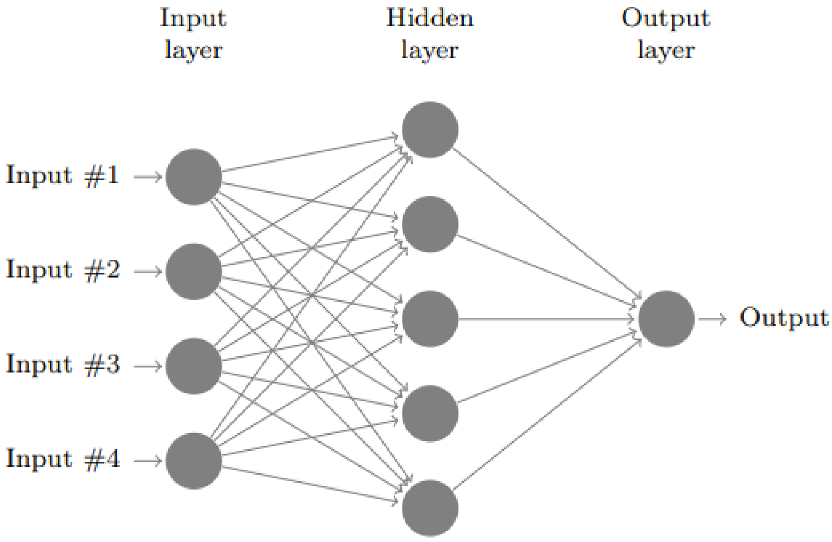


Fig 1 Feedforward Back Propagation Neural Network (FF-BP ANN)

The FFB-ANN model with a single hidden layer has been chosen and illustrated in Figure 1. The first layer has N+3 input neurons, where N is a number of neighbouring links. There are M nodes in the Hidden layer. M will be selected later using experiments. The third layer is designed with an output neuron. FFB-ANNs are used to learn an approximate estimate of the travel time value for each sample, which seems to be a regression problem. So Linear activation functions are applied for neurons in the hidden and output layer in the first experiment setup. Activation steepness is 0.1f.

The model is trained on the most six years of sparse historic data of Leicestershire from 2006 to 2012. They have been acquired from Trafficmaster and Norwich Union data sources. The data are contained in CSV format files. The data contains daily average link travel times, where data exists, for each 15-minute interval.

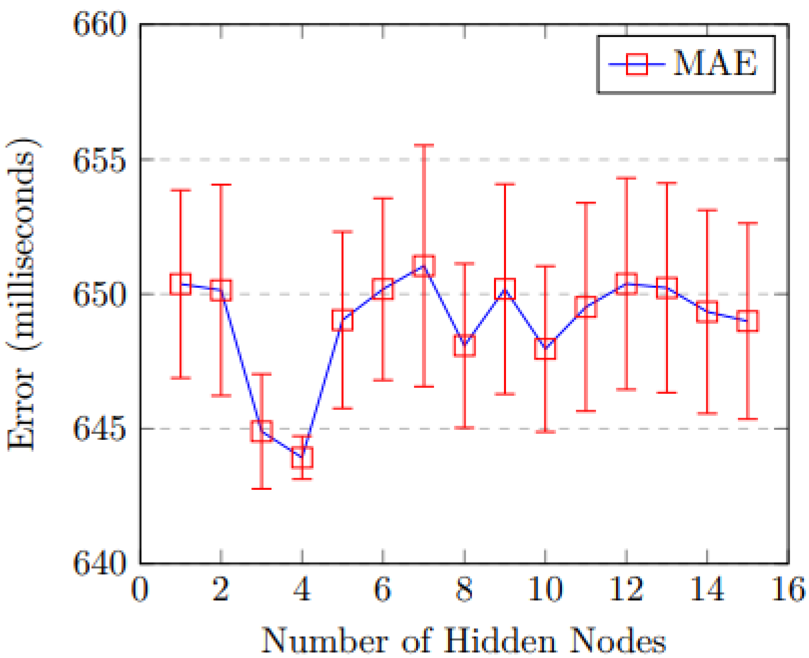


Fig 2 Relationship between the number of hidden nodes and mean absolute error (MAE) as tested in unseen data.

Data samples for training and testing FFB-ANNs are parsed from the sparse historic data where travel time in designated links and directly connected links are available at the same time slot on the same day over six years. Typically, there was only one sample available at different time slots each day. Therefore, the data cannot be considered a time series. Hence, the number of samples is quite small, especially in minor links. The data sample in every single designated link is divided into two parts. Eight ninths of them is used to train the FFB-ANN model and the rest is used to evaluate the model.

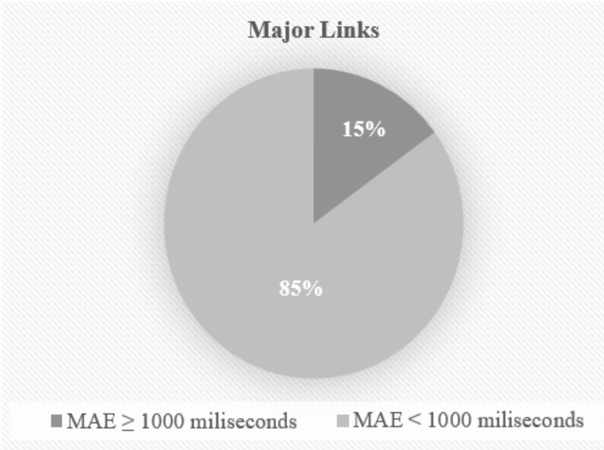
IV. Results

The number of hidden neurons is an important parameter that affects FFB-ANN training and estimating performances. Figure 2 shows the experimental results in terms of Mean Absolute Error(MAE) when FFB-ANNs are trained on a subset of sparse travel time data set with varying numbers of hidden neurons. Results showed that MAE and standard deviation of MEAs are the smallest when the number of hidden neurons is four. Based on the results, we always set the number of hidden neurons to four for all FFB-ANNs in other experiments.

Forty links have been selected from the Leicestershire traffic network. A link will be selected if its number of data samples for training and testing FFB-ANN is greater than one hundred. A large number of experiments have been conducted to test the travel estimation capability of the FFB-ANNs model. The results consist of Mean Absolute Error (MAE), Standard Deviation of absolute errors (StdDev), Minimum of absolute errors (Min), and Maximum of absolute errors (Max) in travel time in milliseconds.

Table 1. Performances of FFB-ANN model on unseen data.

ID	Major Links				Minor Links			
	MAE	**StdDev**	**Min**	**Max**	**MAE**	**StdDev**	**Min**	**Max**
1	361.233	661.227	0	24918	2696.32	3110.33	37	35960
2	331.014	1093.4	0	54901	2824.53	8202.68	37	168456
3	140.345	911.138	0	64350	2113.65	2524.97	8	30977
4	142.969	354.045	0	22302	2514.23	2052.96	9	17968
5	478.595	2240.67	0	83277	3929.5	2694.38	42	15033
6	1088.91	4165.17	0	118372	3810.55	3077.09	42	18036
7	234.682	579.104	0	49057	1989.49	1876.99	0	13320
8	733.568	3987.24	0	315120	2025.15	2317.99	0	19078
9	446.091	1272.28	0	206040	968.689	952.102	1	6002
10	429.182	773.087	0	16648	937.814	1261.44	31	9391
11	536.646	2051.08	0	185706	3731.14	2319.42	19	21232
12	649.363	2339.17	0	121710	3979.41	2326.54	24	18046
13	860.602	2395.17	1	26053	1142.83	1164.99	0	13785
14	1300.71	1043.67	47	5662	1026.36	2290.31	2	40895
15	1675.42	2707.79	0	164942	5515.32	2925.98	1742	32917
16	1879.06	3105.22	8	60112	5541.08	3271.06	1946	36467
17	892.946	997.286	0	36152	880.523	1141.48	0	16365
18	812.135	911.339	0	17231	1376.68	1978.87	2	46206
19	565.841	1958.07	1	33318	278.943	575.982	0	7499
20	310.578	624.514	0	25888	291.072	642.248	0	7675
Mean:	695.506	1734.350	2.714	83331.380	2382.172	2336.339	197.1	29627.55



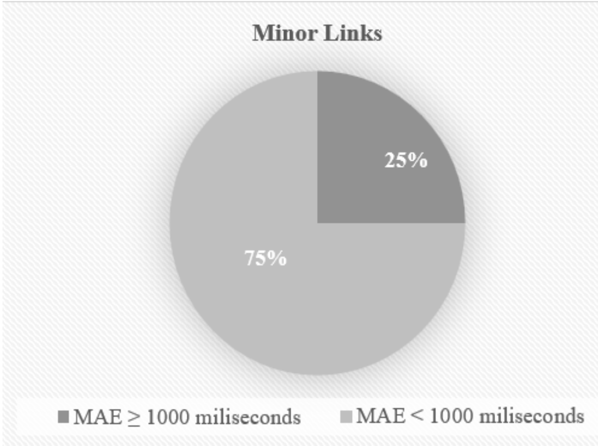


Fig 3 Mean Absolute Error (MAE) of the FFB-ANN models on Major links and Minor links

Table 1 illustrates the results for twenty major links and the twenty minor links in Leicestershire, UK. The results show that the introduced FFB-ANN models are capable of estimating travel times in designated links using travel of neighbouring links. They generally perform better on major than minor links. In major links, FFB-ANN estimators produce maximum MAE value at 1879.06 milliseconds and minimum MAE value at 140.345 milliseconds. More than 80% of estimators in major links have MAE value that is less than 1000 milliseconds. Figure 3 shows that FBB-ANN estimators perform less accurately in minor links. However, most of the MEA values are acceptable. The estimators still show promising results.

The StdDevs values in most of the investigated links are extremely high. They might be impacted by abnormal traffic data and traffic light cycles involved in the training and testing data sets.

V. Conclusions

In this paper, we showed that FFB-ANNs are capable of estimating the travel time on major roads. However, even on non-major roads with more sparse data rate, the performance of the proposed FFB-ANNs estimators shows promising results. It is difficult to accurately estimate urban travel time because traffic light signal cycles usually heavily affect travel times, especially in a short link.

In future work, we will focus on improving the performance of the Neighbouring Link Inference Method, especially in minor links, as the vast majority (70%) of links in the UK fall within the minor link category [25].

References

[1]. C. a. H. A. a. L. J. a. v. Z. H. Hinsbergen, "Bayesian neural networks for the prediction of stochastic travel times in urban networks," *IET Intelligent Transport Systems,* vol. 5, pp. 259-265, December 2011.

[2]. N. P. a. A. Stevanovic, "Traffic Congestion Analysis Visualisation Tool," in *2015 IEEE 18th International Conference on Intelligent Transportation Systems*, 2015.

[3]. G.-L. a. A. K.Singh, "Modeling Arterial Travel Time with Limited Traffic Variables using Conditional Independence Graphs & State-Space Neural Networks," in *Procedia - Social and Behavioral Sciences, 6th International Symposium on Highway Capacity and Quality of Service*, Stockholm, Sweden, 2011.

[4]. Y. G. D. N. a. T. H. M. Jones, "Predicting link travel times from floating car data," in *16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013)*, The Hague, Netherlands, 2013.

[5]. X. M. a. H. N. Koutsopoulos, "A New Online Travel Time Estimation Approach using Distorted Automatic Vehicle Identification Data," in *11th International IEEE Conference on Intelligent Transportation Systems, 2008*, Beijing, China, 2008.

[6]. J. W. a. P. Přibyl, "Real-time travel time estimation on highways using loop detector data and license plate recognition," in *2012 ELEKTRO*, Rajecke Teplice, Slovakia, 2012.

[7]. Travel Time Data Collection Handbook, Texas Transportation Institute, 1998.

[8]. H. v. L. a. H. V. Z. Huizhao Tu, "Travel Time Variability versus Freeway Characteristics," in *2006 IEEE Intelligent Transportation Systems Conference*, Toronto, ON, Canada, 2006.

[9]. L. H. a. M. Barth, "A Novel Loglinear Model for Freeway Travel Time Prediction," in *11th International IEEE Conference on Intelligent Transportation Systems*, Beijing, China, 2008.

[10]. J. R. a. E. v. Zwet, "A simple and effective method for predicting travel times on freeways," *IEEE Transactions on Intelligent Transportation Systems,* vol. 5, no. 3, pp. 200-207, 2004.

[11]. J. H. a. B. S. J. Lee, "Variable Speed Limits Model to Minimize Confliction without Change in Total Travel Time Using Queuing Theory," in *IEEE 18th International Conference on Intelligent Transportation Systems*, Gran Canaria, Spain, 2015.

[12]. X. C. Z. L. a. L. Z. L. Li, "Freeway Travel-Time Estimation Based on Temporal–Spatial Queueing Model," *IEEE Transactions on Intelligent Transportation Systems,* vol. 14, no. 3, pp. 1536-1541, 2013.

[13]. J. Y. a. B. Ko, "Comparison of Travel Time Estimation Using Analysis and Queuing Theory to Field Data Along Freeways," in *International Conference on Multimedia and Ubiquitous Engineering (MUE'07)*, Seoul, Korea (South), 2007.

[14]. X. F. a. C.-C. L. a. K. Liu, "A bayesian dynamic linear model approach for real-time short-term freeway travel time prediction," *Transportation Research Part C: Emerging Technologies,* vol. 19, pp. 1306-1318, 2011.

[15]. H. a. L. Y. a. Y. B. a. B. S. Chang, "Dynamic near-term traffic flow prediction: System-oriented approach based on past experiences," *Intelligent Transport Systems, IET,* vol. 6, pp. 292-305, 2012.
[16]. K. I. W. a. Y. Y. C. J. Y. Wang, "Short-term travel time estimation and prediction for long freeway corridor using NN and regression," in *15th International IEEE Conference on Intelligent Transportation Systems*, Anchorage, AK, USA, 2012.
[17]. G. R. Ruimin Li, "Incorporating uncertainty into short-term travel time predictions," *Transportation Research Part C: Emerging Technologies,* vol. 19, no. 6, pp. 1006-1018, 2011.
[18]. B. a. E. D. a. C. F. a. W. S. a. G. E. Passow, "Adapting traffic simulation for traffic management: A neural network approach," in *Intelligent Transportation Systems - (ITSC), 2013 16th International IEEE Conference on*, 2013.
[19]. Y. H. a. J. X. Haibin Su, "Travel time estimating algorithms based on Fuzzy Radial Basis Function Neural Networks," in *Chinese Control and Decision Conference, 2010*, Xuzhou, China, 2010.
[20]. P. a. M. M. Zheng, "Estimation of travel time using fuzzy clustering method," *Intelligent Transport Systems, IET,* vol. 3, pp. 77-86, 2009.
[21]. Z. H. a. R. Z. L. Deng, "The Bus Travel Time Prediction Based on Bayesian Networks," in *2013 International Conference on Information Technology and Applications*, Chengdu, China, 2013.
[22]. A. a. B. Y. Derbel, "Bayesian network for traffic management application: Estimated the travel time," 2015.
[23]. A. a. B. Y. Derbel, "Bayesian network for traffic management application: Estimated the travel time," in *2015 2nd World Symposium on Web Applications and Networking (WSWAN)*, 2015.
[24]. H. Miura, " A study of travel time prediction using universal kriging," in *TOP 18, 257–270 (2010). https://doi.org/10.1007/s11750-009-0103-6,*, 2010.
[25]. U. Department of Transport, " Guidance on road classification and the primary route network," January 2012.

Vu Huy Luong. "Urban Road Traffic Link Travel Time Estimation Based On Sparse Data." *IOSR Journal of Computer Engineering (IOSR-JCE),* 24(3), 2022, pp. 25-30.