Modeling Individualized Travel Time with a Back Propagation Neural Network

Qing Tang
Department of Civil, Architectural and Environmental Engineering
Missouri University of Science and Technology, Rolla, MO 65409, USA
Phone: (573) 202-0472
Email: qt7b6@mst.edu

Xianbiao Hu, Corresponding Author
Department of Civil, Architectural and Environmental Engineering
Missouri University of Science and Technology, Rolla, MO 65409, USA
Phone:(573) 341-6178
Email: xbhu@mst.edu

Keywords: individualized traffic information, travel time prediction, advanced traveler information system, BP neural network

1. PROBLEM STATEMENT

Current ATIS is designed to provide the "average" traffic conditions although different travelers exhibit heterogeneous driving behaviors, and thus, each individual driver may end up with their own different "individualized" travel time. For example, Google Maps provides the same travel time for everyone who travels from the same origin to the same destination, but the reality is that some travelers may arrive sooner than predicted while others may arrive later, depending on their driving behaviors. The discrepancy between individual drivers' experienced travel time and the system-predicted average travel time leads to errors in travel time prediction and subsequently degrade the user experience.

2. RESEARCH OBJECTIVES

To improve travel time prediction accuracy and enhance user experience, this paper explores the modeling and provisioning of individualized travel time for ATIS that considers the heterogeneous, individual driving behaviors. Specifically, it aims to bridge research gaps in two aspects. First, at the link level, given multiple individual driver' travel time observations (e.g., three probe vehicles driving on the same road segment at the same time), how to derive a more accurate aggregated estimation of traffic dynamics accounting for individualized driving behaviors. We would like to use an example to illustrate the idea: on a road segment at a certain time slot, if driver A is found to be driving at a speed of 65 mph and that is the only data that can be collected, current ATIS practices would conclude an average traffic speed of 65 mph for this road segment. However, if through analyzing accumulated historical driving behavior data and driver A is found to be driving on average 3 mph slower than average traffic, would it not be more reasonable to derive an average traffic speed of 68 mph (65 mph plus the 3 mph individual difference) instead of 65mph?

The second research question to be investigated in this paper is that given the derived aggregated traffic dynamics estimation, how should systems compute individualized route travel time for each driver while considering his or her driving behavior? For example, if driver B is considering taking the same route that driver A is traveling on, and we know driver B is driving on average 5 mph faster than other drivers, would it not be more reasonable to use a traffic speed of 73 mph (68 mph plus the 5 mph individual difference) to derive the predicted route travel time for driver B, instead of using 65 mph or 68 mph?

We expect that this system design would not only enable personalized information modeling and provisioning for ATIS (which would enhance user experience), but also improve the overall system accuracy. This research paper aims to analyze and model the individualized behavioral differences and test these hypotheses with the ultimate goal of improving system prediction accuracy and enhancing user experience.

3. METHODOLOGICAL APPROACH

To achieve the goal of improving travel time prediction accuracy by considering heterogeneous driving behaviors, three modules are designed as illustrated in Figure 1. The basic ideas of the proposed approach are to 1) characterize individualized driving behavior based on historical data, 2) derive real-time link traffic information considering these behavioral differences, and 3) estimate individualized real-time route travel information accounting for behavioral differences.

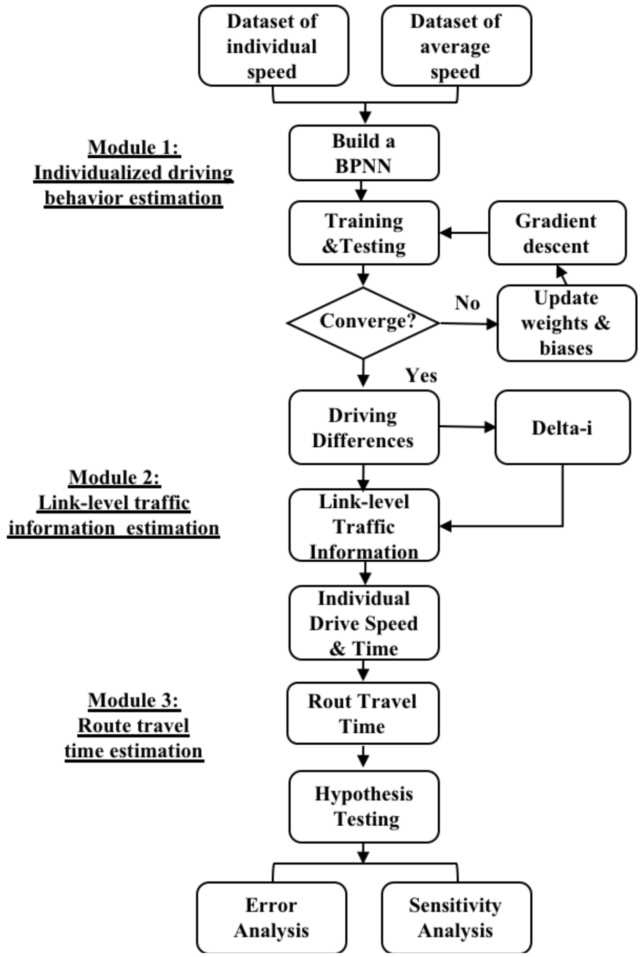


Figure 1 General Framework.

2.1 Individualized driving behavior characterization model

In this module, a back propagation neural network (BPNN) is built to characterize the relationship between individualized driving behaviors and average traffic conditions. Neural network is guaranteed to output $y(x)$ or some close approximation for every possible input x [1] and can be very powerful even with simple network architectures. As illustrated in Figure 2, this particular neural network has three layers: an input layer, a hidden layer, and an output layer. In this module, i, j, k represent the node of the input, hidden and output layer, respectively; w_{ji}^2 represents the weights between the j^{th} node in the hidden layer and the i^{th} node in the input layer; z_j^2 and a_j^2 denote the input and output of the j^{th} node in the hidden layer respectively; w_{kj}^3 stands for weights between the k^{th} node in the output layer and the j^{th} node in the hidden layer; z^3 and a^3 denote input and output in the output layer.

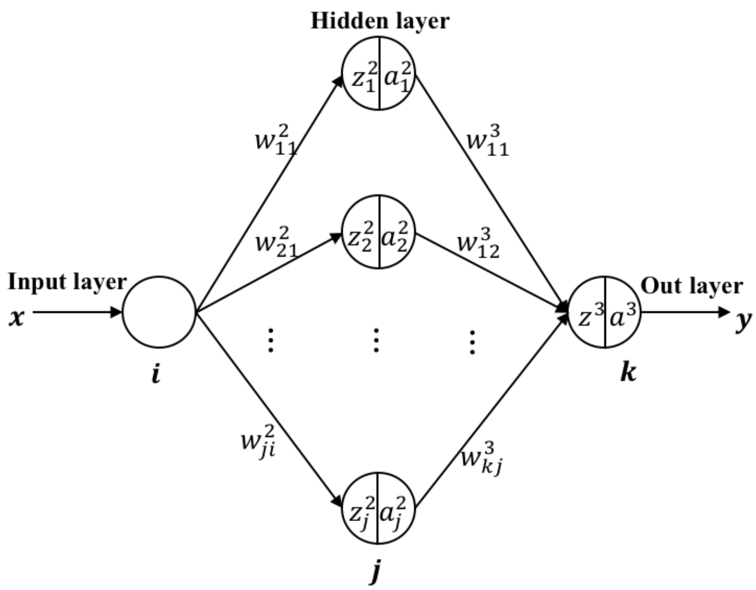


Figure 2 BP Neural Network.

2.2 Link-level traffic estimation model

In this module, we proposed a link travel time estimation algorithm that uses the delta to derive link-level traffic information based on the multiple individual driver's travel time observations. We suppose that driver e wants to travel from origin O to destination D at a given time (say as t) and h_g vehicles are observed on link g at the same time. The speed of these vehicles is $v_1, v_2 \dots v_{h_g}$, respectively.

2.3 Individualized route travel time prediction model

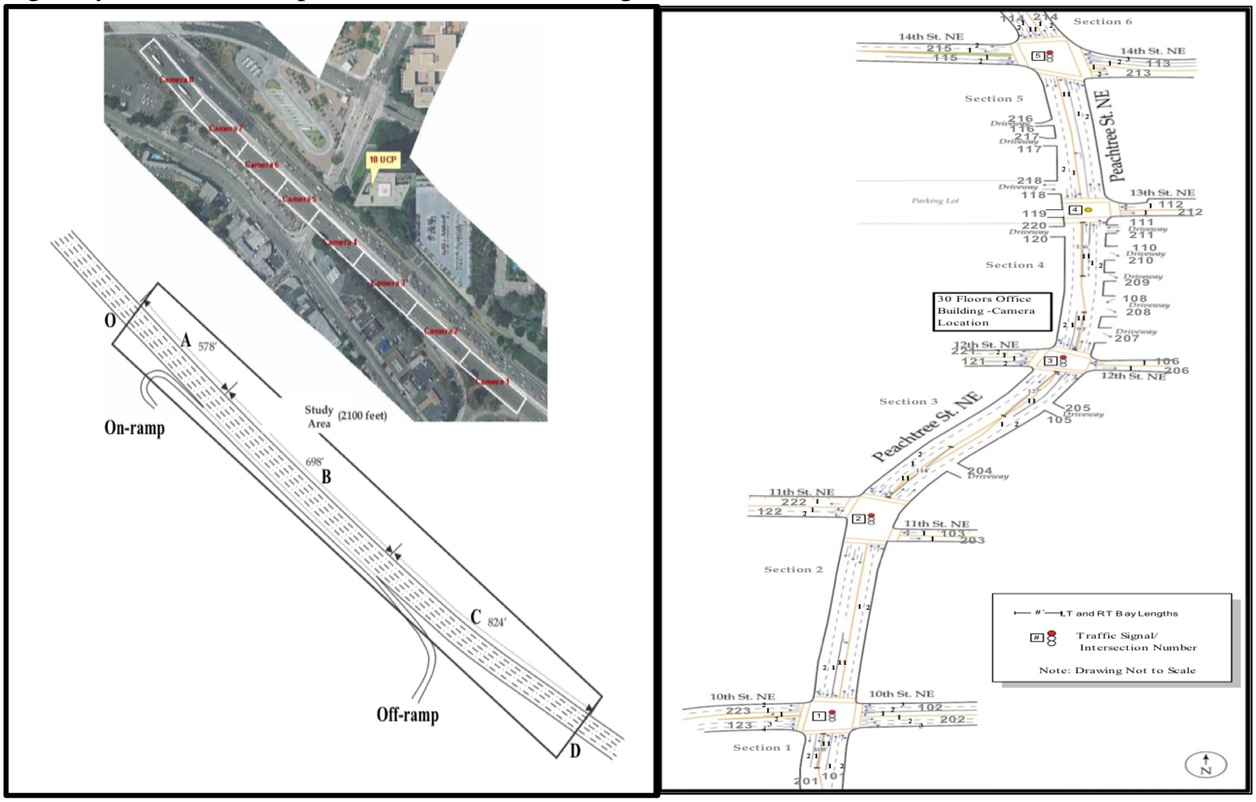
Suppose driver e is about to start a trip and the system needs to estimate route travel time. After module 1 and 2, the following information has been derived: 1) the delta for this particular driver e based on his or her historical driving data, and 2) link-level aggregated traffic information based on the real-time observations from those who are already traveling in the roadway network.

We first estimate the link-level individualized drive speed $\tilde{v}_e^g$ for driver e on the g^{th} link by considering the driver's own delta and traffic information $\bar{v}_g'$ as mentioned in module 2. Next, we aggregate the number of links between that route in a time-dependent fashion. Route travel time is predicted by aggregating every link's travel time instead of using the mathematical average speed to calculate travel time.

4. Preliminary Results

4.1. Dataset Description

We used the real data from the Federal Highway Administration's Next Generation Simulation (NGSIM) dataset to train the BP neural network model. Our study areas consist of an uninterrupted highway and an interrupted arterial shown in Figure 3.



(a) (b)

Figure 3 (a) Study area schematic on Highway 101 in Los Angeles, CA (b) Study area schematic Peachtree Street in Atlanta, GA.

4.2.Hypothesis testing

The basic assumption we have made in this research is that driver's behavior is heterogeneous, and thus using the average value to describe individual performance becomes problematic. In this section, we aim to validate this assumption with the NGSIM field data. A simple analysis found that 323 (40.4%) of drivers' speed is higher than the average speed, 254 (31.8%) of drivers' speed is lower than average, and 223 (27.8%) of drivers' speed is about the same as average ($\pm$1ft/sec) on Highway 101. In the same way, 156 (38.5%) of drivers' speed is higher than the average speed, 162 (40.0%) of drivers' speed is lower than average, and 87 (21.5%) of drivers is about the same as average ($\pm$1ft/sec) on Peachtree Street.

We reject the null hypothesis (H_0) and accept the alternative hypothesis (H_1) that individual speed is not identically distributed with average speed, which validates our previous assumption that individual driving behavior is heterogeneous and using the average value to describe individual performance is indeed problematic.

4.3. Modeling result analysis

Through training, testing and predicting, the results show that 662 (82.8%) drivers' APE dropped for highway case, which means that the proposed BPNN model is able to improve travel time prediction accuracy for most drivers in an uninterrupted traffic flow scenario. On average, their APE dropped by 7.7%, and there are 18% drivers whose APE dropped by over 12%. Considering traffic on highway is uninterrupted flow and traffic prediction research has been relatively mature, a 7.7% average improvement can be considered significant.

On the other hand, for the arterial case the travel time prediction for 69.2% drivers have been improved by 21.3% on average, and 17.5% travelers experience a travel time prediction accuracy improvement of over 30%. While it is well understood that due to the nature of interrupted traffic flow, the traffic time prediction is more complicated and difficult than that of uninterrupted flow, such 21.3% average improvement demonstrates the significant benefits brought by the proposed model. This result proves the previous suspicion that for a heterogeneous system, if we are able to use different values to characterize or predict each driver's performance instead of using an average value, the prediction error due to the heterogeneity and the averaging calculation will be reduced or eliminated. By comparing the interrupted and uninterrupted flow cases, it can be observed that the proposed model can improve traffic prediction accuracy for both cases, and it brings more significant benefits to the arterial case which has higher heterogeneity.