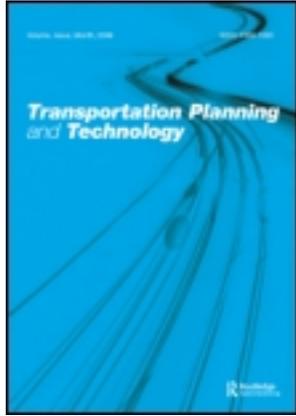


This article was downloaded by: [University of Waterloo]

On: 24 August 2011, At: 10:50

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Transportation Planning and Technology

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/gtpt20>

Estimation of time-dependent, stochastic route travel times using artificial neural networks

Liping Fu^a & L. R. Rilett^b

^a Department of Civil Engineering, University of Waterloo, Waterloo, Canada, N1L 3G1

^b Department of Civil Engineering and Texas Transportation Institute, Texas A&M University System, College Station, Texas, 77843-3135, USA

Available online: 21 Mar 2007

To cite this article: Liping Fu & L. R. Rilett (2000): Estimation of time-dependent, stochastic route travel times using artificial neural networks, *Transportation Planning and Technology*, 24:1, 25-48

To link to this article: <http://dx.doi.org/10.1080/03081060008717659>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.tandfonline.com/page/terms-and-conditions>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan, sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

ESTIMATION OF TIME-DEPENDENT, STOCHASTIC ROUTE TRAVEL TIMES USING ARTIFICIAL NEURAL NETWORKS

LIPING FU^{a,*} and L. R. RILETT^b

^aDepartment of Civil Engineering, University of Waterloo, Waterloo,
Canada, N2L 3G1; ^bDepartment of Civil Engineering and Texas
Transportation Institute, Texas A&M University System,
College Station, Texas, 77843-3135, USA

(Received 12 February 1999; In final form 31 May 2000)

This paper presents an artificial neural network (ANN) based method for estimating route travel times between individual locations in an urban traffic network. Fast and accurate estimation of route travel times is required by the vehicle routing and scheduling process involved in many fleet vehicle operation systems such as dial-a-ride paratransit, school bus, and private delivery services. The methodology developed in this paper assumes that route travel times are time-dependent and stochastic and their means and standard deviations need to be estimated. Three feed-forward neural networks are developed to model the travel time behaviour during different time periods of the day – the AM peak, the PM peak, and the off-peak. These models are subsequently trained and tested using data simulated on the road network for the City of Edmonton, Alberta. A comparison of the ANN model with a traditional distance-based model and a shortest path algorithm is then presented. The practical implication of the ANN method is subsequently demonstrated within a dial-a-ride paratransit vehicle routing and scheduling problem. The computational results show that the ANN-based route travel time estimation model is appropriate, with respect to accuracy and speed, for use in real applications.

Keywords: Artificial neural network; Travel time; Shortest path algorithm; Travel distance function; Vehicle routing scheduling; Dial-a-ride paratransit

*Corresponding author.

1. INTRODUCTION

The estimation of travel time from one location (origin) to another (destination), or O–D travel time, in a road traffic network is one of the key components in the vehicle routing and scheduling process of many fleet vehicle operations systems such as dial-a-ride paratransit, school bus, and private delivery systems (Savelsberch and Sol, 1995; Bodin *et al.*, 1983). Due to the inherent variation of urban traffic congestion, weather conditions and traffic incidents, O–D travel times are time-dependent (or dynamic) and stochastic. It has been increasingly recognised that such variations may have important impacts on the productivity and reliability of fleet vehicle operations and therefore should be explicitly considered in the routing and scheduling models (Dror and Powell 1993; Fu, 1996). Explicit consideration of these variations, however, requires a travel time estimation model that can provide a fast and accurate estimation of the descriptors that represent the dynamic and stochastic travel times. The primary objective of this paper is to demonstrate the feasibility of using artificial neural networks (ANN) for estimating the dynamic and stochastic O–D travel times within a vehicle routing and scheduling process. The dynamic and stochastic attributes of the O–D travel time are represented by the mean and standard deviation of the O–D travel time, both of which are a function of the time of day and the O–D locations.

In most of the existing vehicle routing and scheduling models, O–D travel time has traditionally been modelled as being static (time-independent) and deterministic in that a single value is used to represent the travel time between each O–D pair (Savelsberch and Sol, 1995; Bodin *et al.*, 1983). The methods used to estimate the O–D travel time commonly involve two steps. First, the distance between two locations is estimated based on location coordinates. In most practical applications, either rectangular distance or Euclidean distance is used to approximate the O–D travel distance (Love, 1988; Brimberg and Love, 1991). The estimate of the travel distance is then divided by an average speed to obtain the O–D travel time estimate. While it may be possible to extend this procedure to account for the dynamic and stochastic variation of travel times, the accuracy of the estimation will likely be in question because factors such as road network topology and traffic congestion usually have a highly non-linear impact on the O–D travel time.

A more accurate estimation method is to use a shortest path algorithm to calculate the travel time between two locations when the underlying road network and associated travel times or speeds on each link is available. If both the mean and standard deviation of the route travel time as functions of time of the day need to be estimated, the corresponding link travel time data in the network must be provided (Fu and Rilett, 1998). Apart from the need for such extensive network data, the major obstacle for using this method is that it could be too computationally intensive to be directly integrated in the routing and scheduling process for solving large sized problems (Shen *et al.*, 1995), as will also be demonstrated in this paper.

An intermediate method is to divide the service area into many small zones and subsequently calculate the travel time between individual zones using shortest path algorithms. Alternatively, the travel time between zones may be obtained based on past daily travel records, and pre-stored in computer memory. A trip from an origin to a destination would then be mapped as a trip from the zone where the origin is located to the zone where the destination is located. Although this method is potentially more efficient in terms of computational time, computer storage may be an issue if the number of zones utilised is large for more accurate representation of travel times between zones. For example, if the area of study is divided into 2000 zones and the time period of interest (*e.g.*, 6:00am~21:00) into 60 intervals (15 minutes per interval), the required computer memory to store both the mean and standard deviation between all zones for all time intervals would be in the order of 960 MB ($= 2000^2 * 60 * 2 * 2$ bytes).

Artificial Neural Networks (ANN) have become one of the most popular techniques in the Artificial Intelligence (AI) field during the last decade. The special architecture and computation mechanics inherent in ANN models make them useful for a wide variety of tasks such as image processing, pattern recognition, and solving combinatorial problems. ANN have been found to be very useful in modelling the relationship between quantitative and qualitative inputs and their related output. More detailed information may be found in other references (Rumelhart, 1986). The potential of using an artificial neural network to provide a quick and accurate estimation on the O-D travel time is the focus of this paper.

This paper first proposes three feed forward neural networks to model the O-D travel time behaviour (mean and standard deviation)

during the different time periods of a day: the AM peak, the PM peak and the off peak. The paper then discusses how the input attributes are identified, how the training data are represented, and how the “best” ANN models are developed and identified. These analyses were performed using data generated based on a traffic network for the City of Edmonton, Alberta. A comparison of the ANN model with the distance-based method and a shortest path algorithm is then presented. Lastly, the practical implications of the ANN-based method is illustrated through its application in a dial-a-ride paratransit vehicle routing and scheduling example.

2. NEURAL NETWORK BASED TRAVEL TIME ESTIMATION MODEL

This section examines the model building steps used to construct the artificial neural network. The topology of the ANN model and the representation of input and output data are first presented. The training and testing procedures are then discussed.

2.1. ANN Network Topology

The ANN used in this analysis is known as a back-propagation neural network and the general topology is illustrated in Figure 1. The ANN consists of three layers with the neighbouring layers fully connected. The output layer includes cells representing the variables to be estimated – in this situation the O–D travel time. The input layer represents factors which may have an impact on the O–D travel time. These factors may include such information as the geographic locations of the origin and destination and the departure time at the origin. The number of hidden nodes is a decision variable and determined during the training and testing stage.

In a typical urban environment the day is typically classified into three different time periods: the AM peak, the PM peak and the off peak. In each case the link travel time patterns may differ significantly. Instead of using a single ANN to map the travel time pattern for an entire day, three separate ANN models are developed for these three periods. The models developed are referred to as the AM Net, PM Net

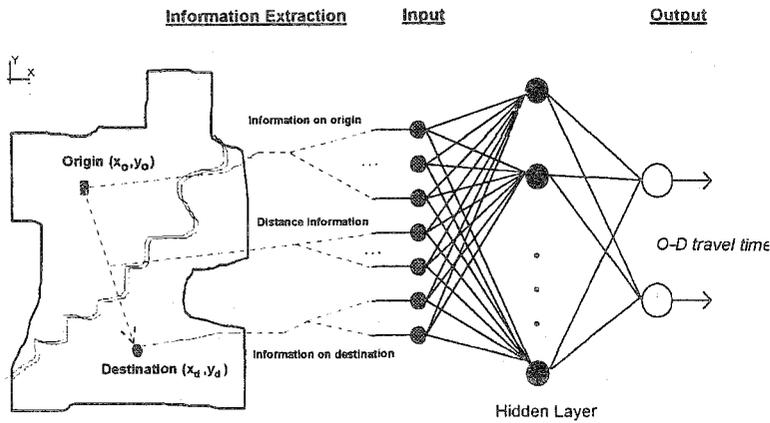


FIGURE 1 ANN topology for O-D travel time estimation.

and OFF Net in this paper. For the purposes of this research the AM peak was defined as lasting from 0600 to 0900 hrs, the PM peak was defined as lasting from 1500 to 1900 hrs and all other time periods belong to the off-peak period. In an actual implementation the selection of the time periods, their duration, and their number will obviously be a function of the data availability, the location of interest, and the type of vehicles being routed/scheduled.

2.2. Data Representation

The representation of the data is one of the most important steps in the development of a neural network model. There are two major steps in the data representation process. The first step is to identify the variables associated with the input and output as schematically illustrated in Figure 1. The output, O-D travel time (t_{od}), is represented by an estimate of the mean O-D travel time ($E[t_{od}]$) and an estimate of the standard deviation of the O-D travel time ($S[t_{od}]$). While the standard deviation may not be required in the routing and scheduling process it is useful to show the ability of ANN to estimate additional O-D travel information. In addition, it would be clearly advantageous to modify routing and scheduling algorithms to take advantage of this information.

Two different network design options were examined with respect to output estimation. The first option entailed using two separate neural

networks – one that was used for estimating the mean travel time and another one for estimating the standard deviation about this mean. In this option each network had one output cell. The second option entailed using a single neural network to estimate both the mean and standard deviation of the O–D travel time. In this situation two output cells were required.

Two input scenarios were tested for selection of appropriate input variables. The first, known as Scenario A, considered five input attributes: the coordinates of the origin and destination locations, *i.e.*, (x_o, y_o) and (x_d, y_d) in meters, and the departure time (T_0) in minutes after midnight. This scenario attempted to capture location effects, time of day effects, and to a limited degree, distance effects. In scenario B, two extra variables representing the estimated distance are added to those of Scenario A. These variables are directly related to the distance between the origin location and the destination location and have been historically used for distance estimation in the vehicle routing and scheduling problem. These two variables are the rectangular distance (l_1 , meters) and the Euclidean distance (l_2 , meters) from the origin location to the destination location and are defined in Eqs. (1) and (2) below.

$$l_1 = |x_o - x_d| + |y_o - y_d| \quad (1)$$

$$l_2 = \sqrt{(x_o - x_d)^2 + (y_o - y_d)^2} \quad (2)$$

Obviously, there are a number of possible combinations of inputs and in an actual implementation this step would involve examining more options.

The second step involved data normalisation. For the sake of learning effectiveness, all the input data are scaled into values that ranged between 0 and 1, while the outputs are scaled into values that ranged from 0.2 to 0.8. This transformation ensures that the output remains in the quasilinear part of the sigmoid function where learning is faster (Gallant, 1993).

2.3. Training and Testing Examples

In order to evaluate the ability of the proposed ANN models in capturing the O–D travel time patterns in a given urban area, the

municipal area of the City of Edmonton is used as a test bed in this paper. Ideally, the data used for training and testing the ANN should be direct field collected O–D travel times that represents the “true” travel time pattern in the area of interest. Because there were no such data available for the modelling purposes of this research, simulated travel time data based on the road network for the City of Edmonton were used. The “true” travel time between two locations were assumed to be the travel time on the shortest path between these locations in the given road network and were calculated using a shortest path algorithm. These travel time data were subsequently used in training, testing and evaluating the ANN models. It should be noted that, while the models developed in this study and the associated results are valid only for demonstration purposes, the methodology and principles are applicable for use when the appropriate field data become available.

The Edmonton network is composed of 3800 links and 1400 nodes and used primarily for planning applications. The link length and posted speed on the link are given as part of the network database. Because the given network data is not sufficient for generating time-dependent mean and standard deviation of route travel times, additional link travel time attributes were created. The dynamic and stochastic travel time patterns in the network were represented by hypothetical changes in the mean and standard deviation of travel time on each link. The following section describes how these link travel time data were created. It must be emphasised that the objective of this to generated data which are reasonably representative of actual condition and at the same time include as much variation as possible. The goal is to see whether ANN can model the non-linear travel time effects.

The urban area was divided into three sub-areas: Downtown, Midtown and Suburban with link travel times in each of these areas defined separately in order to consider the difference in the traffic patterns between them. It was assumed the dynamic and stochastic pattern on links in each area is the same and were defined in the following manner. The mean link travel time during the off-peak period is assumed to be equal to the link length divided by the posted speed while during the peak periods the mean travel time increases with percentages that are dependent on the area in which the link is located. Figure 2 shows the dynamic profile of travel times defined for the three sub-areas.

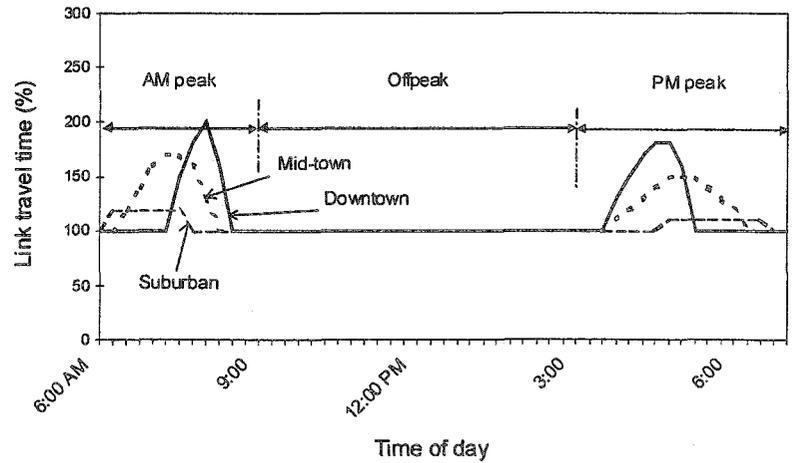


FIGURE 2 Dynamic link travel time pattern.

TABLE I Link travel time CV in different sub-areas

| <i>Subdivision</i> | <i>Downtown</i> | <i>Midtown</i> | <i>Suburban</i> |
|--------------------|-----------------|----------------|-----------------|
| Link Travel Time | 1.0 | 0.5 | 0.1 |
| CV | | | |

To generate the variation of the link travel time, it is assumed that the coefficient of variation (noted as CV and defined as the ratio of the standard deviation to the mean) of the link travel time is the same for all links in the same sub-areas. Table I gives the hypothesised link travel time CV for each sub-area. A larger value of CV is used for the links in the central area of the city based on the fact that the network has relatively more traffic controls and is denser. Given the mean and CV of the link travel time, the standard deviation (or variance) of the link travel time can then be determined. Note that, similar to the mean travel time, the resulting standard deviation is also time-dependent.

The data required for both training and testing an ANN model were developed as follows. The origin and destination locations of individual O-D pairs were randomly generated. For each O-D pair, six random departure times were generated during each of the AM, PM and off-peak periods, respectively. It was found that this number of departure times was sufficient to characterise the dynamic pattern of travel time during both the AM and PM peak periods for the given

network. For a network on which travel times change more rapidly, more departure times may be necessary. Once the departure times were identified, the means and standard deviations of the O–D travel times were calculated by a shortest path algorithm (Fu and Rilett, 1998). These data were then combined with the coordinate information for each O–D pair to form the training data. Table II shows an example of the training data for two O–D pairs.

2.4. Training, Testing and Results

The training of the neural network is done *via* the back-propagation learning algorithm (Rumelhart, 1986) with the overall objective being to find the “best” ANN to model the O–D travel time. The determination of the best ANN is commonly based on two criteria. The first one is related to the learning speed which is reflected by the number of iterations needed to completely train the neural network. In this situation the lower the number of iterations the better the model because it will require less computational effort. The second criteria is the root mean square error (RMS) defined in Eq. (3) where the closer the estimated data to the actual data, the lower the RMS and hence the better the model.

$$\text{RMS} = \sqrt{\frac{\sum_{k=1}^N \sum_{i=1}^M (Y_{ki} - D_{ki})^2}{N * M}} \quad (3)$$

where:

- N total number of O–D pairs to be trained;
- M number of output cells;
- Y_{ki} actual value at output cell i for example k ;
- D_{ki} estimated value at output cell i for example k .

For the purpose of this research, all the ANN models were trained using an off-line training program and therefore the time required to train a ANN is not considered as part of the evaluation criterion.

Travel times of 1000 O–D pairs were used as training examples and another 250 O–D pairs for testing. All these O–D pairs were generated using the procedure discussed previously. It should be noted that the number of O–D pairs required to train an ANN is a function

TABLE II Training data examples (original data)

| <i>Origin node coordinates (x_o, y_o) (meter)</i> | <i>Destination node coordinate (x_d, y_d) (meter)</i> | <i>Rectangular distance (l_1) (meter)</i> | <i>Euclidean distance (l_2) (meter)</i> | <i>Departure time (T_0) (minutes)</i> | <i>Travel time (tt) (seconds)</i> |
|--|--|--|--|--|--|
| (33814.1,34357.9) | (3011.3,43866.8) | 40312.7 | 30796.5 | 394 | 3847 |
| (33268.2,37791.5) | (2891.0,37372.2) | 32237.4 | 30380.1 | 511 | 1427 |

* Total minutes elapsed since midnight.

of the complexity of the travel time patterns that the ANN needs to learn. In most situations, the inherent patterns are often hidden and therefore a trial-and-error procedure is required. The procedure to identify the best ANN included three steps. The first step was to identify whether Scenario A or Scenario B gave the best results. It was found that the ANN with Scenario B (7 inputs) was clearly superior to that of Scenario A (5 inputs) in terms of both learning speed and prediction quality. While this should not be surprising because the former uses more input variables, it did indicate the importance of using “processed” information in improving the learning ability and training speed of an ANN model.

After selecting the input attributes, the second step identified the best representation of the location information. The locations of the origin and destination may be represented exactly by using their geographic coordinates or approximately by using zonal coordinates. The analysis showed that the two distance measures (l_1 , l_2) appeared to contain most of the distance information between the two locations and that an approximate representation of the locations was adequate. Therefore, the network area under analysis was uniformly divided into a 1000×1000 grid and each origin and destination were assigned to this grid system. It was found that using the exact coordinates significantly reduced the learning speed with only a very minor improvement in model quality.

The final step was to decide on the optimum number of hidden nodes for the two proposed network design options (*i.e.*, separate neural networks for estimating the mean and standard deviation and a single combined neural network) and to subsequently identify the preferred approach. As the first step, each model was trained with two to twenty hidden nodes. It was found that the combined model required much more hidden nodes in order to achieve the same level of estimation quality as the separate models. Figures 3 and 4 shows the estimated CV of the O–D travel time and the actual O–D travel time CV as a function of trip length from the separate ANN model with five hidden nodes and the combined ANN model with ten hidden nodes. The data were deliberately generated such that all the trips originated from a single location in the downtown. It can be seen that the separate neural network models achieved better results than the combined neural network model. It can also be seen in Figure 4 that

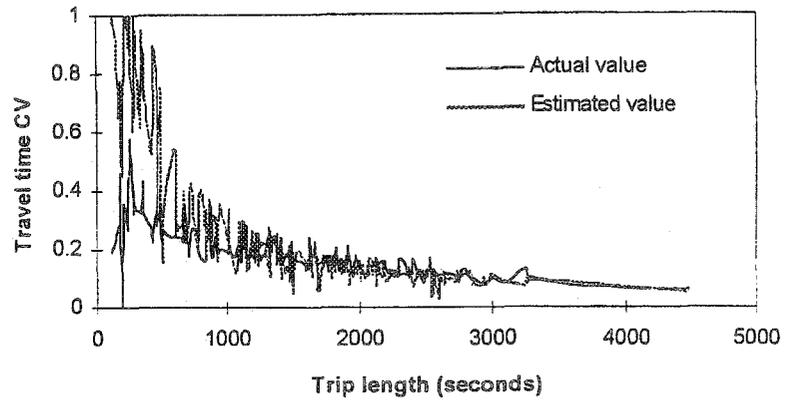


FIGURE 3 CV of O-D travel time by a joined network model: actual vs. estimated value as a function of trip length.

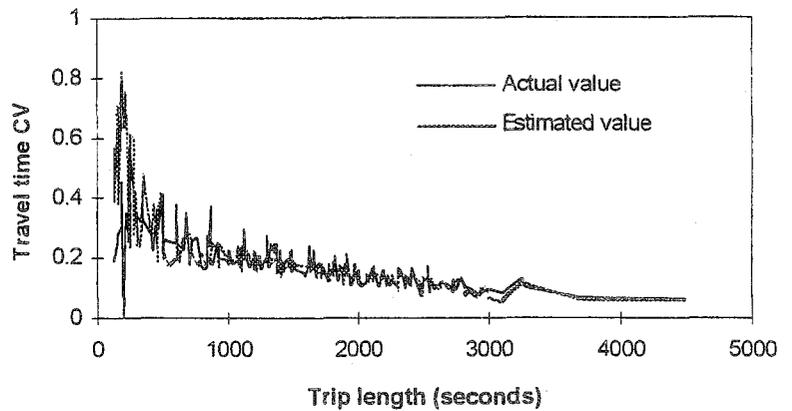


FIGURE 4 CV of O-D travel time by a separate network model: actual vs. estimated value as a function of trip length.

the combined neural network model over estimates the O-D travel time variance for short trips, but provides good estimates for trips longer than approximately 600 seconds. Based on these results it was decided to use the separate neural network option.

The next step involved training and analysing the neural networks, with two to twenty hidden nodes, for all three time periods. The neural network with five hidden nodes was found to be adequate to model both the AM and PM peak periods while a neural network with four

hidden nodes gave the best results for the off-peak period. As an example, the RMS as a function of the number of iterations for the mean travel time AM Net (the so-called learning curve) is shown in Figure 5.

The neural networks were subsequently tested on 250 randomly generated O-D pairs. Figure 6 illustrates the results from the O-D

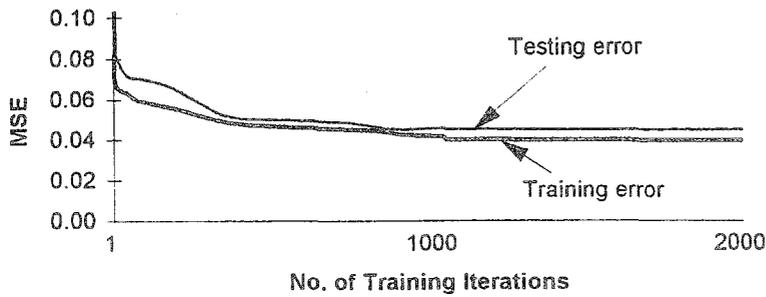


FIGURE 5 Learning progress curve for an AM net.

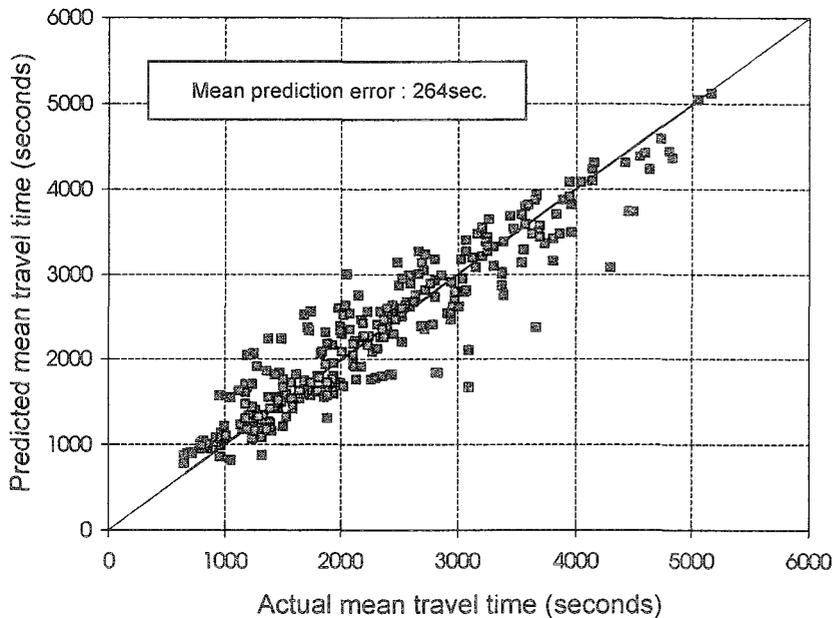


FIGURE 6 Actual mean travel time vs. travel time predicted by the AM net.

travel time estimation model as compared to the actual mean O-D travel time during the AM peak period. The average relative error (the difference between the estimated mean travel time with the actual mean travel time divided by the actual mean travel time) was found to be 12.1%. This relative error is equivalent to approximately 264 seconds for the average trip length of 2180 seconds.

In order to show the ability of the trained neural networks to model the dynamics of the travel time during a day, the predicted mean travel time and the actual mean travel time with different departure times of day are compared in Figure 7. This shows these travel times for two O-D trips, where the estimated value is from the three neural networks. One trip is from the northwest of the Suburban area to the

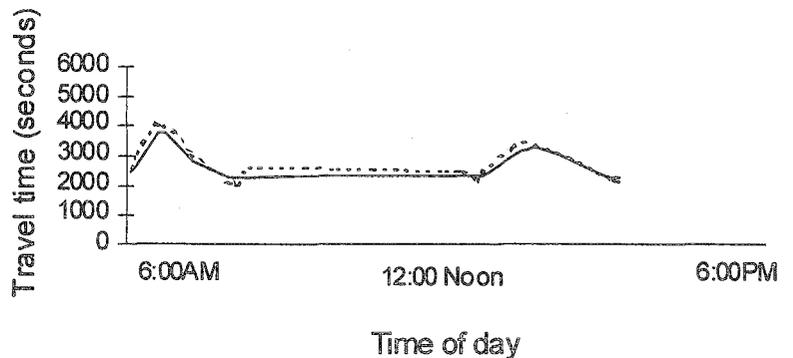
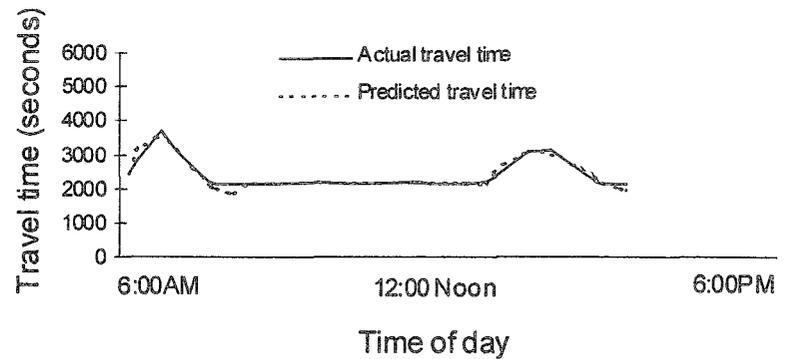


FIGURE 7 Actual travel time pattern vs. estimated travel time pattern: two O-D pairs.

south of the Mid-town area and another is from the south of the Mid-town area to the Downtown area. It can be seen that the non-linear relationship between the O–D travel time with time of day is tracked relatively well by the neural networks.

3. COMPARISON OF ANN AND TRADITIONAL METHODS

The comparison was based on the assumption that a rectangle and Euclidean distance function would be used to represent the traditional travel time estimation method, and that the shortest path algorithm would be used to represent the exact approach. We should note that the conclusions are by no means definitive and should be interpreted qualitatively because the inputs between these methods are different and better models may exist for the approximate method.

3.1. Distance-based Method

As discussed previously, the relationship between the O–D travel time and location information has historically been estimated using distance-based methods. In this approach, either the rectangular distance or the Euclidean distance is used to approximate the O–D travel distance, which is then divided by an average speed to obtain the travel time. It is therefore not “fair” to make a direct comparison between the ANN model and the distance-based method because the former uses more input variables than the later and therefore should naturally be superior in performance. The following two steps were therefore taken in order to make the comparison more reasonable.

First, the off-peak period is deliberately selected as the modelling period so that the non-linear impact of the departure time on the O–D travel time can be removed. In addition, the travel times are assumed to be deterministic and therefore an estimate of the standard deviation of the O–D travel time is not required. As a result, the O–D travel estimation problem is effectively the same as the O–D travel distance estimation problem in which distance-based models have been popularly applied.

Second, instead of using the rectangular distance or the Euclidean distance, a regression analysis was conducted to obtain the direct relationship between the O–D travel time and the influencing factors, including the coordinates of the origin and destination, a set of variables transformed from the coordinates, and the rectangular and Euclidean distances.

A total of 1000 O–D pairs were randomly generated in order to develop the distance-based model and the new ANN models. In addition, 800 randomly generated O–D pairs were used for testing. The topology of the ANN model used to model the off-peak period is the same as that shown in Figure 1 except that the departure time input cell and standard deviation output cell are removed. The ANN was trained following the procedure discussed previously using the 1000 O–D pairs.

A series of regression analyses were performed to establish the best equation between the O–D travel time and the independent variables listed above. It was found that there were no statistically significant or meaningful relationship between the O–D travel time and any subset of the coordinates variables and transformed variables. In addition, l_1 and l_2 were found to be statistically correlated and therefore should not be included in the same regression equation. As a result, the following calibrated models were found best to represent the travel time in the Edmonton network.

$$t_{od} = 0.0558 l_1, \quad R^2 = 0.81 \quad (4)$$

(147.0)

$$t_{od} = 0.0690 l_2, \quad R^2 = 0.80 \quad (5)$$

(152.2)

where:

t_{od} = travel time from origin (o) to destination (d), seconds;

l_1 = rectangular distance (Eq. (1));

l_2 = Euclidean distance (Eq. (2)).

Note that the values in brackets are the respective t - values. It may be seen that the variable is significant at the 95% level and the model explains a good deal of the variability. Note that the variable on the

right hand side of Eqs. (4) and (5) is in units of distance and the output variable on the left hand side is in units of time.

Both the trained ANN and the distance-based models were applied to the test data and the results are summarised in Table III. It can be seen that, with respect to both RMS and average relative estimation error, the estimation errors of the distance-based model is approximately doubled as compared to the neural network model.

It should be pointed out that the neural network had additional input variables and at first glance it appears that one cannot fairly compare the results. However, it was not feasible to add the additional information from both a statistical point of view and a practical perspective to the regression equations. It is this flexibility in specifying non-linear or correlated terms within the neural network that makes them appealing for applications discussed within this paper. In addition, the coordinates would not impose an additional data collection burden as they basically are already used as inputs in the traditional approach.

3.2. Shortest Path Approach

The purpose of this section is to demonstrate the computational efficiency of the ANN models as compared to the method of directly using the shortest path algorithm to calculate travel time. The comparison is based on the Edmonton road network. A label setting algorithm (LS) and A* algorithm elsewhere (Rilett *et al.*, 1994) are used to find the expected minimum O–D travel times. All the

TABLE III Comparison of prediction error of ANN model and distance-based model

| <i>Data</i> | <i>Measurement</i> | <i>ANN model</i> | <i>Distance-based model</i> | |
|---------------|-----------------------------|------------------|---------------------------------------|-------------------------------------|
| | | | <i>Rectangular distance (Eq. (4))</i> | <i>Euclidean distance (Eq. (5))</i> |
| Modeling data | RMS (seconds) | 162.5 | 350.4 | 362.2 |
| | Average relative error* (%) | 10.9 | 21.4 | 22.3 |
| Test data | RMS (seconds) | 182.5 | 364.7 | 367.0 |
| | Average relative error* (%) | 12.2 | 21.0 | 21.7 |

*relative error = |predicted value – actual value|/actual value

programs in this study are coded in C++ and executed on a Pentium-II computer with a 300 MHz clock speed and 128 MB RAM. A total of 100,000 O-D pairs were randomly generated and their travel time are estimated using the LS, A* and ANN approaches. The total CPU time for each algorithm to calculate the travel times for the given number of O-D pairs was recorded and the results are shown in Table IV. It can be seen that the ANN is approximately 5.4 times faster than the LS algorithm and 2.4 times faster than the A* algorithms. While the computation time is not really an issue for a simple route calculation when tens of thousands of calculations may be required (such as in a vehicle routing and scheduling problem) then this saving can be significant – and may offset the decrease in accuracy from using the ANN. Furthermore, it should be noted that the speed of the shortest

TABLE IV Computation time of shortest path algorithms and the ANN model

| <i>Method</i> | <i>Total CPU to calculate the travel times of 100,000 OD pairs (seconds)</i> |
|---------------|--|
| LS algorithm | 5.4 |
| A* algorithm | 2.4 |
| ANN model | 1.0 |

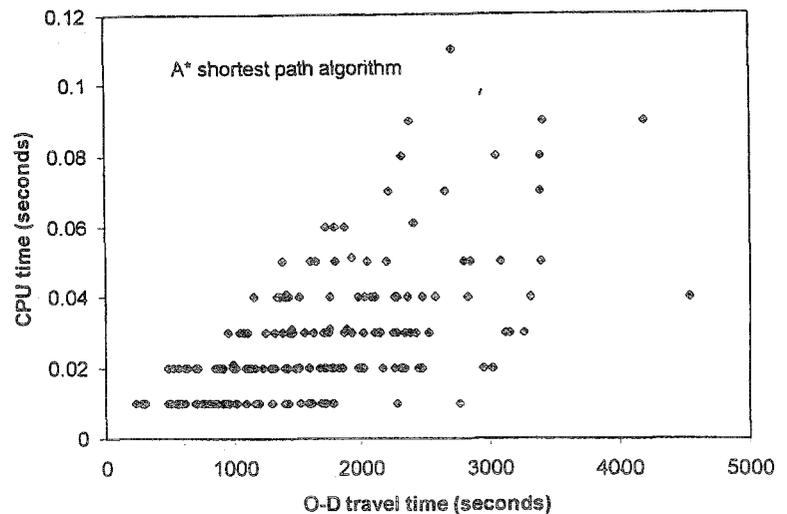


FIGURE 8 The relationship between computation time and travel time: A* shortest path algorithm.

path algorithm depends on the size of the road network. It can be expected that the relative efficiency of the ANN method as compared to the shortest path algorithm would be much larger when more detailed networks are used.

Another advantage of the ANN is that the time to compute the O-D travel time is the same for all O-D pairs regardless of their travel times. In contrast to the ANN model, the computation effort of a shortest path algorithm increases as the O-D travel time increases. Figure 8 shows the computation time of the A^* as a function of O-D travel time for the Edmonton network.

4. APPLICATION OF ANN O-D TRAVEL TIME ESTIMATION TO THE DIAL-A-RIDE PARATRANSIT VEHICLE ROUTING AND SCHEDULING PROBLEM

The objective of this section is to illustrate the application of the proposed ANN based O-D travel time estimation method in a vehicle routing and scheduling process. The vehicle routing and scheduling problem used in this analysis is from a dial-a-ride or shared-ride paratransit system although the technique is equally applicable for any related problem. The sub-problems in this paper consist of the subscriber dial-a-ride problem (*subscriber* DARP) and the real-time dial-a-ride problem (*real-time* DARP). In the former situation the task is to determine the assignment of all customers (or trips) to the available vehicles and their respective routes and schedules. This task is typically performed off-line and the users are informed some time interval after the initial request. In the latter case the objective of the process is to determine the assignment of a new customer into the existing schedule of a vehicle in real-time and inform the users whether their request can be handled at the same time it is made (*e.g.*, within, say, sixty seconds). A related situation is one in which the schedule has to be altered due to some external event. The detailed formulation and description of the dial-a-ride vehicle routing and scheduling problem with time-dependent and stochastic O-D travel time can be found elsewhere (Fu, 1996). The algorithm used to solve these two types of problem is a modified version of the traditional trip insertion algorithm that allows the use of time-dependent and

stochastic O-D travel time (Bodin *et al.*, 1983; Jaw *et al.*, 1986; Fu, 1996).

The test problems were derived from the Disabled Adult Transportation System (DATS) in the City of Edmonton, Alberta. Data on the weekday service (0600 ~ 2100 hrs) was obtained including 3000 trips of which approximately 75% trips are wheel-chair trips and the rest are ambulatories. Two sub-problems consisting of 500 and 1000 trips were generated by randomly selecting trips from the 3000 trips. A fleet of 106 vehicles is available to provide the service.

The origin and destination locations of the trips spread over the municipal area of the City of Edmonton. The same network previously described in Section 2.3 was used. The travel time on each link was assumed to be time-independent and deterministic, and was determined based on the length of the link and the posted speed on the link.

The routing and scheduling objective is assumed to minimise the expected total travel time. In Edmonton, the operating constraints consists of a maximum ride time of 90 minutes and a maximum service time deviation of 30 minutes to avoid excessive inconveniences to clients.

The above routing and scheduling problems were all solved using the same vehicle routing and scheduling algorithm but with the three different O-D travel time estimation methods discussed in the previous sections. For the ANN method, a single ANN with five hidden nodes was used to estimate the average travel time as needed in this specific case. In order to make the comparison impartial, the actual travel time used was based on the distance estimation method. That is, an additional step of using the distance function to estimate the travel time was included at the end of each of the three methods and consequently the same travel time was used in the routing and scheduling process. As a result, the schedules created by using those three methods were exactly the same. Figure 9 shows the relationship between the CPU time required to schedule all the trips as a function of the problem size for each of the three different O-D travel time estimation methods discussed in this paper. It can be seen that it required over 160 minutes of CPU time to solve the 3000 trip problem when the heuristic shortest path algorithm (A^*) was used in the vehicle routing and scheduling process. In contrast, the routing and scheduling algorithm with the ANN method solved the problem within

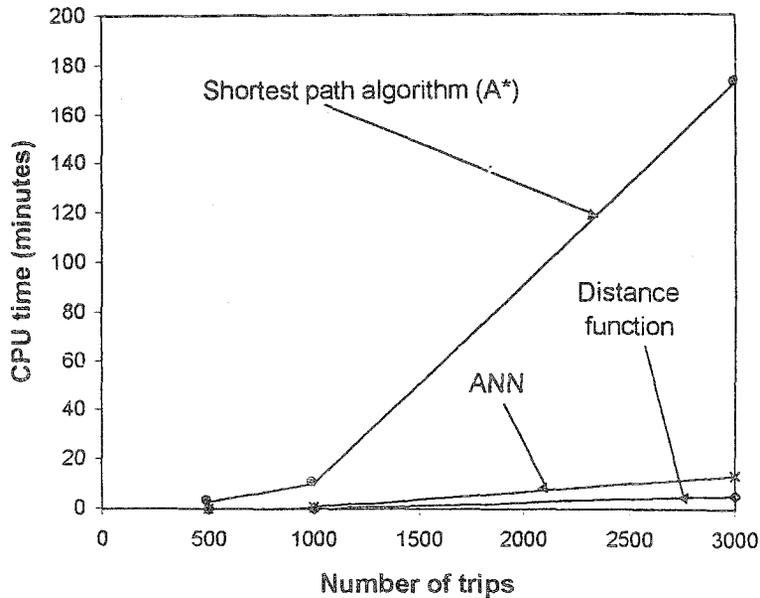


FIGURE 9 Relationship between computational time and problem size: *subscriber DARP*.

14 minutes of CPU time. This results in more than a ten times improvement over the approach using the SPP. As would be expected because of its simplicity, the distance-based regression formula provides the computationally fastest results. However, as discussed in Section 3, there is the possibility that the error in routing and scheduling would be higher because of its lower accuracy level.

In the real-time DARP the problem is relatively simple. Given that a request has been received by the operator, the request needs to be scheduled (or denied) in an acceptable time frame. The real-time scheduling is simulated by assuming that the new request is to be inserted into the existing schedules which were created by solving the subscriber DARP discussed previously and that the sequences of the original schedules are to be preserved. It was found that the CPU time required to insert a new trip into the existing schedules with up to 3000 trips on them were all less than one second when the distance-based and ANN-based O-D travel time estimation methods were used. This finding implies that the ANN method is efficient enough to be used in

a real-time scheduling situation. When the heuristic shortest path algorithm (A^*) was used, it required an average CPU time of two seconds to solve the 1000 trip problem and approximately eleven seconds to solve the 3000 trip problem. While it appears to be feasible in this case to use a shortest path algorithm in the real-time routing and scheduling process, it could impose a problem if the underlying road network is much larger and re-optimisation of the existing schedules is required (*e.g.*, exchange of trips between individual schedules). Furthermore, it should be noted that in order for the SPP approach to consider the dynamic and stochastic variations in travel time, extensive link travel times would be required.

5. CONCLUDING REMARKS

This paper introduced the concept of using ANN models for estimating the time-dependent and stochastic O-D travel time in an urban traffic network. Based on travel time samples from real network data and simulated link travel time patterns, a variety of ANN models were trained and evaluated. A detailed analysis of the performance of the ANN models as compared to the shortest path algorithm and distance-based methods was conducted. The main conclusions are summarised as follows.

It was found that an ANN model can be trained to map effectively the highly non-linear relationship between the O-D travel time and their location information in time-dependent and stochastic traffic networks. The success of an ANN technique for travel time estimation mainly depends on how the input information was abstracted and what type of network model was used. This study demonstrated that some enhanced data (for example, distance information) can be very helpful in improving the performance of an ANN, and that separate estimation models for different parameters to be estimated are much more effective than using a joined network model.

The solution quality of the ANN method was found to be significantly better than the traditional distance-based model for estimating O-D travel times. Therefore, better O-D estimates may be obtained for routing and scheduling applications by using the ANN O-D travel time estimation method instead of the traditional

distance-based method. While the ANN is not as accurate as the shortest path algorithms, it is much faster and independent of the size of the road network. For the example used in this paper the ANN was more than 500 times faster than the shortest path algorithms. Therefore it is useful in situations where travel time calculations are necessary but where the computation time is limited. The computational study showed that the ANN models proposed in this paper are feasible for use in the dial-a-ride vehicle routing and scheduling algorithm to solve realistic subscriber DARP and real-time DARP. On the other hand, the shortest path algorithm was found to be viable only in solving small problems, that is, problems with less than 50 trips.

Finally, it should be noted that there are other applications that potentially could make use of an estimation method of O-D travel times on time-dependent and stochastic networks. For example, it has been shown that a bi-directional shortest path algorithm is often faster than a uni-directional algorithm, but the former algorithm is only applicable in networks where the travel times are static (Kuznetsov, 1993; Rilett *et al.*, 1994). This is because in a time-dependent network, the bi-directional algorithm requires exact information about the departure time at an origin node and the arrival time at a destination node. An accurate estimate of the O-D travel time may make it possible to implement a bi-directional algorithm in a time-dependent network by using an estimate of the arrival time as an input to the process.

Acknowledgment

Funding for this research was supplied by the Natural Science and Engineering Research Council of Canada and is gratefully acknowledged.

References

- Savelsberch, W. P. M. and Sol, M. (1995) "The General Pickup and Delivery Problem", *Transportation Science*, 29(1), 17-29.
- Bodin, L., Golden, B., Assad, A. and Ball, M. (1983) "Routing and Scheduling of Vehicles and Crews. The State of Art", *Computers and Operations Research*, 10, 69-211.

- Dror, M. and Powell, W. (1993) "Stochastic and Dynamic Models in Transportation", *Operations Research*, 41(1), 11-14.
- Fu, L., "Improving Paratransit Scheduling by Accounting for Dynamic and Stochastic Variations in Travel Time", Accepted for presentation in *TRB annual meeting and publication in Transportation Research Record*, Transportation Research Board, National Research Council, Washington, D.C., 1998.
- Love, R. F., Morris, J. G. and Wesolowsky, G. O., *Location: Models & Methods Facilities*, North-Holland, 1988.
- Brimberg, J. and Love, R. F. (1991) "Estimating Travel Distances by the Weighted L_p Norm", *Naval Research Logistics*, 38, 241-259.
- Fu, L. and Rilett, L. R. (1998) "Shortest Path Problems in Traffic Networks with Dynamic and Stochastic Link Travel Times", *Transportation Research, Part B: Methodological*, 32B(7).
- Shen, Y., Potvis, J., Rousseau, J. M. and Roy, S. (1995) "A computer Assistant for Vehicle Dispatching with Learning Capabilities", *Annals of Operations Research*, 61, 189-211.
- Rumelhart, D. E. and McClelland, J. L. (Eds.), *Parallel Distributed Processing: Explorations in the Micro Structure of Cognition*, Vol. 1, MIT Press, Cambridge, 1986.
- Gallant, S. I., *Neural Network Learning and Expert Systems*, The MIT Press, Cambridge, Massachusetts, London, England, 1993.
- Fu, L., "Real-Time Vehicle Routing and Scheduling in Dynamic and Stochastic Traffic Networks", unpublished *Ph.D. Dissertation*, University of Alberta, Edmonton, Alberta, 1996.
- Jaw, J., Odoni, A. R., Psaraftis, H. N. and Wilson, N. H. M. (1986). "A Heuristic Algorithm for the Multi-vehicle Advance Request Dial-A-Ride Problem with Time Windows", *Transportation Research -B*, 20B(3), 243-257.
- Kuznetsov, T., "High Performance Routing for IVHS", *IVHS America 3rd Annual Meeting*, Washington, D. C., April, 1993.
- Rilett, L. R., Blumentritt, C. and Fu, L., "Minimum Path Algorithms for In-vehicle Route Guidance Systems", *Proceedings of the IVHS AMERICA*, 1994.