Route choice modelling in urban transportation networks using fuzzy logic and logistic regression methods

Yetis Sazi Murat* and Nurcan Uludag

Department of Civil Engineering, Faculty of Engineering, Pamukkale University, 20070, Denizli, Turkey

Received 20 October 2006; revised 19 October 2007; accepted 22 October 2007

This study presents route choice model of transportation network in Denizli using fuzzy logic model (FLM), logistic regression model (LRM) and survey-based data. Four important parameters (traffic safety, travel time, congestion and environmental effects) were used in the models. Fuzzy logic concepts are especially used to take note of imprecision, vagueness and uncertainty characteristics of route choice behaviour. FLM is compared with LRM and encouraging results are obtained with FLM.

Keywords: Fuzzy logic method, Logistic regression method, Route choice, Traffic safety, Travel time

Introduction

Route choice behaviour depends on travel costs, travel time, traffic safety, comfort, habits, and socio-economic and demographic characteristics, of which travel time is generally the key factor of all\(^1\). Overall approach for the route choice is to evaluate costs of each route and to choose the route with lowest travel cost. Primary parameters while making a choice between an origin and destination are travel time, traffic safety, intervals, cost (fuel), traffic signs, jam and queuing, road type, landscape, road constructions and habitual effects. Cost is generally proportional with travel distance. Random structure of the problem can be modelled by some methods based on utility maximisation. However, in these methods, crisp values are used so that they do not answer uncertainty of the problem.

Route Choice Modelling

Initially, about route choice topic, User Equilibrium (UE) models were developed. Variation of driver goals and expectations\(^2\) leads to various route choice decisions, and this variation appears as a stochastic component in modelling. In order to relax restrictive assumptions in UE models, logit and probit models were developed\(^3,4\). In these models, drivers are accepted as rational decision makers who maximise their utility\(^5\). Ben-Akiva & Lerman\(^6\) dealt with route choice phenomenon in detail and proposed some models including discrete choice analysis. General expression of logit model\(^7\) is as follows:

\[
P_i = \frac{e^{V_i}}{\sum_{j=1}^{n} e^{V_j}}
\]

where, \(P_i\) = choice probability of \(i_{th}\) route, \(V_i\) = deterministic part of \(i_{th}\) route’s utility.

Lo & Lam\(^8\) studied on using transportation information systems on Multinomial Logit (MNL) route choice models. Abdel-Aty et al\(^9\) studied statistical analysis of route choice. Mahmassani & Srinivasan\(^10\) modelled perception of travel time by road users with an online information system. Do et al\(^11\) revealed that drivers are very conservative about route choice decision even though they have long-term learning.

Fuzzy Modelling

Murat & Gedizlioglu\(^12\) applied fuzzy logic (FL) on traffic signal control and obtained promising results. Yin et al\(^13\) used fuzzy neural network (FNN) in urban traffic flow prediction. Murat\(^14\) also used FL and FNN in modelling vehicle delays at signalized intersections. Lan & Huang\(^15\) considered traffic incident detection using...
FNN approach. Chiou & Lan\textsuperscript{16} regarded build-operate-transfer transportation projects and developed royalty models for uncertainties. Teodorovic & Kikuchi\textsuperscript{17} developed a model in which travel time is perceived as a fuzzy parameter for drivers. Akiyama & Tsuibo\textsuperscript{18} developed a multi-step model in which perceived travel time, congestion and accident risks are taken as fuzzy parameters. Lin & Lee\textsuperscript{19} used FNN. The model results were more flexible than logit model especially in decision process. Henn\textsuperscript{20} developed a traffic assignment model by taking the uncertainty and vagueness in dynamic choice process.

Lee \textit{et al}\textsuperscript{21} developed a model where Latent Class Multinomial Logit (LCML) and FL results were combined in order to evaluate the drivers’ randomness and uncertainties together. Binetti & De Mitri\textsuperscript{22} proposed a model in which the costs are represented with fuzzy values and, road users make their choices by comparing these fuzzy costs. Vythoulkas & Koutsopoulos\textsuperscript{23} studied on introducing the rule weights to FL models. Palma and Picard\textsuperscript{24} investigated route choice behaviour of road users considering uncertainty in travel time. Henn & Ottomanelli\textsuperscript{25} denoted the importance of uncertainty in traffic assignment models and the lack of classical utility models. Arslan & Khisty\textsuperscript{26} developed a hybrid model where route choice decision-making was described in a hierarchy that uses concepts from fuzzy logic and analytical hierarchy process (AHP). Individual travel behaviors are analyzed by Asakura & Iryo\textsuperscript{2} using mobile phones without information from base stations. Morikawa & Miwa\textsuperscript{28} interested in dynamic route choice behavior of road users and showed that probe vehicle can be used for modeling decision process during the trip. Fuzzy set theory\textsuperscript{29,30} provides strong and meaningful tools especially for representation of uncertain conceptions.

**Fuzzy Logic Model (FLM) Parameters**

Route choice problem, because of the fuzzy character of the parameters and perceptual differences of the users, is not suitable to be modelled by the classical mathematical models. Road users have subjective information about the parameters on route choice and perceptions are different. These subjective information and perceptions cause some uncertainties on route choice decision. Because of crisp approach, it is very difficult to model these intuitive uncertainties using conventional methods. For instance, travel time of a route can be defined in a different way by each road user and the interval can be ranged in a wide scope. This interval can easily be defined by FL approach using membership functions. But, in crisp approach, this can be defined only one value and perceptions of most of the road users’ are ignored. This is a deficiency in modeling approach. Therefore, FLM is preferred to take the note of imprecision, vagueness and uncertainty characteristics of route choice behaviours.

In this study, fuzzy logic and logistic regression route choice models are proposed with four essential parameters — travel time, traffic safety, congestion and environmental effects.

**Proposed Fuzzy Logic Model**

**Survey Study for the Model**

The data used for the model was obtained from a route choice survey\textsuperscript{31} arranged for Denizli, Turkey. Survey was organized to 500 road users (university students, university staff and other employee from different business sectors). In the survey, decisions of routes for two directions from Kampus to Çınar and Çınar to Kampus (Fig. 1) were asked from road users. Importance rates of the reasons of this decision and route choice decision for peak hour traffic conditions were also inquired. There were also some questions about maximum and minimum travel times of different transportation modes serving on these routes.

With the help of questions and grades, public transportation and private car users’ evaluations for four routes were displayed. Road users’ perceptions about travel time, traffic safety, congestion and environmental effects for different routes were displayed by grade. The boundaries of grades are selected as 0-20 for all of the parameters. The effect of these parameters on the

![Fig. 1 — Network modelled in the study](image-url)
decisions of road users and the route choices at normal and peak times were determined. Private car ownership, mode choices of attendee, parameters effecting mode choice, and changes of route choice decisions at peak times were obtained with survey study. The parameters in the survey were related with each other; by a statistical examination, it was seen that correlation between the parameters is significant (Table 1). The data set is partitioned as training and test sets (400 data for training; 100 data for test). Test data are selected randomly from whole data set.

Most effective parameter is travel time, followed by traffic safety, congestion and environmental effects (Fig. 2). Environmental factors include general scene and views along the routes. Indifference in determining the best route (Fig. 3) was related to subjective perception of road users, because of which same route is perceived as the best for two directions. Evaluation of the routes with traffic safety parameter (Fig. 4) and their route choices are parallel. Shortage of travel time does not directly affect the route choice behaviour (Fig. 5). Therefore, this study concerns some additional parameters (traffic congestion, traffic safety and environmental effects) in modelling route choice behaviour of road users.

Structure of the Fuzzy Logic Model
In FLM, membership functions of input and output parameters are determined based on survey results and by the help of an expert’s knowledge. For input parameters (travel time, traffic safety, congestion and environmental effects), triangular and trapezoidal membership functions are used, while for output parameter (route utility), triangular is used (Fig. 6). Most convenient membership functions are defined by examination of statistical dispersion of the survey results and after some trials.
Fig. 4 — Evaluation of routes with traffic safety parameter

Fig. 5 — Evaluation of routes with travel time parameter

Fig. 6 — Fuzzy logic route choice model parameters and membership functions
Input parameters (Fig. 7) that are obtained from survey study were used in the first stage, given values of these parameters are converted to fuzzy using membership functions in the second stage, the rule base was used for inference in the third stage, using rule base a fuzzy output that shows the utility of route was obtained and the fuzzy output was converted to crisp numbers using a defuzzification method in the fourth stage. At the end, route utility was obtained as a crisp number. After these processes, choice probabilities of routes were determined by multinomial logit formula.

Rule base used in the fuzzy logic model is formed after statistical analysis of the input and output parameters. Rule base consists of 3*3*3*3=81 rules; 3 for traffic safety; 3 for congestion, 3 for travel time and 3 for environmental effects. These IF-THEN rules are collated with AND operator (Table 2). The model output is gained where Mamdani inference mechanism is used. In order to compare fuzzy results of FLM with crisp results of mathematical models, fuzzy output is defuzzified using Centroid method as

\[ z^* = \frac{\int \mu_c(z) z dz}{\int \mu_c(z) dz} \]  

where, \( z^* \) = crisp value (value for x axis), \( \mu_c \) = membership value for given crisp value.

By defuzzification of output, values of each route for each attendee are determined that helped to evaluate the 4 routes globally for the whole survey.

**Logistic Regression Model (LRM)**

The results of survey were analysed with LRM that is used for binomial situations. The survey data was examined in the meaning of traffic safety, congestion, travel time and environmental effects. For each route, an independent LRM was formed in which the selected route was denoted with 1, and non selected with 0 using 500 survey attendees’ evaluations for 4 parameters and logistic regression choice functions were calculated. Logistic regression choice probability (P) is given as

\[ P = \frac{1}{1 + e^{-(b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + b_4 x_4)}} \]  

where, P: probability of chosen, x1: traffic safety, x2: congestion, x3: travel time, x4: environmental effects b0, b1, b2, b3, b4 are regression factors.

Survey data was analysed with LRM, which is a statistical method used for binomial situations, and
regression factors were found (Table 3). By using these factors (Table 3), utility of each individual route is determined. LRMs are developed corresponding to 95% confidence interval. All of the models developed satisfy these levels. Maximum likelihood method is used in estimation of logistic regression parameters. Level of significance of the models is determined using chi-square test that is used for determination of goodness of fit and significances of parameters. Looking into chi-square values of the coefficients (Table 4), all of the models are found significant and can be used for modelling route choice behaviour.

### Results and Discussion

#### Comparison of Fuzzy Logic and Logistic Regression Models

Results of FLM and LRM are compared considering the training and test set. Compatibility of results obtained with LRM and FLM with the whole survey data is computed. If the route, which logistic regression (LR) results showed as the highest choice value and the real selected one, are the same, this meant the compatibility between LR and real data. Likewise, compatibility between FL and real data is computed. The function used for this accuracy computation is

\[
\text{Accuracy}\% = \frac{100}{N} \sum y_n \quad ...(4)
\]

where, \( N \) = number of examples; \( y_n = 1, \text{ if } n_{th} \text{ individual’s choice in survey and model result is the same, otherwise 0.} \)

Validations of the models are investigated before comparisons. The training data set (400 data) is used for validation search. Comparisons are made in two levels as aggregate and disaggregate. The choices of the routes are considered in aggregate level whereas the choices probabilities are taken into account in disaggregate level. Accuracy rates of LRMs and FLM for training data are reasonable (Table 5), therefore, both of these models are used in modelling route choice behaviour.

FLM results are also compared with real data in order to search validation. All of the data are used (500 data) in this validation search. After defuzzification of fuzzy output of FLM, utility of each route is acquired as crisp value. Comparison (Table 6) of the accuracies of FLM and LRMs results shows that FLM provides higher accuracy rates than LRM. Some sample comparisons for disaggregate level are also presented (Fig. 8) for Camlik & K Sehitler Routes using test data.

#### Conclusions and Recommendations

Comparing with current mathematical models, results of fuzzy approach are more realistic because of its
Table 5 — Route choice probabilities

<table>
<thead>
<tr>
<th>Route choice/</th>
<th>Route choice/</th>
</tr>
</thead>
<tbody>
<tr>
<td>by survey, %</td>
<td>by Fuzzy model, %</td>
</tr>
<tr>
<td>Çamlik</td>
<td>31.00</td>
</tr>
<tr>
<td>K.Sehitler</td>
<td>41.60</td>
</tr>
<tr>
<td>İstiklal</td>
<td>23.60</td>
</tr>
<tr>
<td>Ç.Yolu</td>
<td>3.80</td>
</tr>
</tbody>
</table>

Table 6 — Accuracy rates of the models for training and test data in aggregate level

<table>
<thead>
<tr>
<th>Destinations</th>
<th>Logistic regression</th>
<th>Fuzzy logic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Training data</td>
<td>Test data</td>
</tr>
<tr>
<td>Çamlik</td>
<td>83.25</td>
<td>85</td>
</tr>
<tr>
<td>K.Sehitler</td>
<td>83.25</td>
<td>86</td>
</tr>
<tr>
<td>İstiklal</td>
<td>90.00</td>
<td>94</td>
</tr>
<tr>
<td>Ç.Yolu</td>
<td>99.75</td>
<td>100</td>
</tr>
</tbody>
</table>

Fig. 8 — Route choice probabilities for Çamlik (a) and K Sehitler (b) routes in disaggregate level
characteristics on modelling imprecision and uncertainty in route choice. Accuracy of FLM results is higher than LRM results. Average accuracy of FLM for all routes from Kampus to Çınar is 91.75%, and LRM gives 89.06% for training data in aggregate level. On the other hand, these values are 93.5% and 91.25% for test data respectively. Results for disaggregate level also supports this finding.

Traffic safety and environmental effects have an importance effect on route choice decision of road users. Considering all of these parameters by fuzzy logic approach provides better results comparing with logistic regression. According to the survey results, the most preferred route is K1br1s ^ehitler, followed by Çaml1k, Östiklal and Çevre Yolu. Consequently, FLM results are realistic, consistent with real data, and valid for both directions. Survey results showed that the routes with shortest and longest travel times are perceived identical for both directions. This result is related to subjective perceptions of road users, because of which same route is perceived as the best for two directions. According to the answers of attendees, most effective parameters for route choice decision is travel time, followed by traffic safety, congestion and environmental effects. Even at peak periods, while there is a congestion problem, road users do not change their routes.

For future studies, weather condition can be taken into account as an important parameter on route choice decision. It can be integrated with environmental factors or it can be used as a new parameter. Because comfort of the trip is related to road type and road users can prefer more comfortable route. Therefore, road type can be effective on route choice decision of road users. Membership function boundaries can be re-determined by new methods, such as Neural Networks, Genetic Algorithms in order to get better results. The weights of fuzzy rules can be changed for better results. In another study, a new fuzzy model combining route and mode choice can be developed. For the route choice models, definition of drivers’ decision behaviour is very important. The models developed for route choice can be used for transportation planning and traffic management especially. Furthermore, route choice probability determined with this model can be used for developing traffic assignment models. With a combined model of Fuzzy Logic and traffic assignment, some scenarios can be constituted.

References


24 Palma A & Picard N, Route choice decision under uncertainty, in *10th Int Conf on Travel Behaviour Res* (Lucerne, Switzerland) 2003, 60-72.


