Transport Modal Choice Model with Adaptive Neuro-Fuzzy Utility Functions

Katia Andrade
Kenetsu Uchida
Seiichi Kagaya
Hokkaido University

Abstract: This paper suggests the hybrid Neuro-Fuzzy Multinomial Logit model for use in transportation studies. The model is applied for estimating travelers’ behavior in the context of the transport modal choice problem, where the modes of bus, subway and automobile are investigated. The model is evaluated by comparing its results with the results of a Multinomial Logit model. Moreover, the probabilities of selecting a mode obtained by applying the two models are compared with the actual transport modal choices, showing the efficiency of the suggested model in practical transportation problems.

1. Introduction

Travel behavior models are valuable tools in the field of traffic demand estimation since they provide information about travelers’ preferences. This information is required in the design and planning of transportation systems. For several decades, statistical models have been largely applied in solving traffic and transportation problems. Statistical models are most often developed on the basis of objective knowledge, i.e. formulae and equations that are used in solving daily life problems. However, this kind of modeling overlooks the fact that human reasoning is approximate rather than precise in nature [1]. Therefore, this uncertain knowledge, which can be named subjective knowledge or linguistic information and is used by human beings in day to day decisions, should also be considered while solving real-life problems [2]. In this context, researchers usually face difficulties in interpreting the uncertainties inherent to the real-life choice behavior problems while using exclusively mathematical models [3]. The uncertainties regarding travelers’ decisions are composed of two types: the randomness due to the non-deterministic nature of choice behavior and the vagueness due to the linguistic expression of the attributes of the transportation network. Thus, the travel behavior model should describe both kinds of uncertainties so as to be adequate.

Several models, which have been suggested since the 70s to deal with the different uncertainties, have formed a new category of travel behavior models. This category includes various techniques of soft computing, such as Fuzzy Logic and Neural Network and statistical methods where the familiar and highly-effective Logit models are highlighted. The new travel behavior models, sometimes called hybrid models, are suggested by some studies with promising results [4].

The aim of this study is to contribute to the development of the existing hybrid travel behavior models by suggesting a combined adaptive neuro-fuzzy multinomial logit model. The overall structure is based on the Multinomial Logit (MNL) model. However, the adaptive utility function is described by a neuro-fuzzy inference system. By using the neuro-fuzzy utility function into the MNL formulation, a better description of the combination of the vagueness uncertainty and the randomness uncertainty is expected. The model, called neuro-fuzzy multinomial logit (NFMNL) model, is developed for the transport modal choice problem. The preferences of shopping travelers regarding the modes of bus, subway and automobile are investigated. In order to evaluate the NFMNL model, a MNL model was also developed and the results were compared. The estimation results of both models are investigated and compared with the survey results, which showed better performance of the proposed model.

2. Basic Structure of Logit Models

Logit models, based on random utility theory, are well established as discrete choice models. By adopting this model, the basic assumption is that a traveler will select the transport mode which provides the maximum utility in the economic sense [5]. The utility of an alternative i for a person n is described in the following equation:

\[ U_{in} = V_{in} + \varepsilon_{in} \]  

(1)

where \( V_{in} \) usually described by a linear function, is the deterministic term of the utility for alternative i, while the second term (\( \varepsilon_{in} \)) is the random variable for the utility. By following the choice theory, the probability of selecting an alternative \( i \) in a multinomial process, is given by:

\[ P_{in} = \Pr(U_{in} \geq U_{jn}), \forall \ j \in C_n, \ j \neq i \]  

(2)
3. Database

The result of a questionnaire survey which was carried out by targeting shopping travelers who live in the district of Shin-Sapporo (12Km far from downtown Sapporo city) is used to investigate travelers’ behavior. Two commercial zones in downtown Sapporo city were assumed as destination zones, i.e. the Sapporo station area and the Odori station area.

A thousand questionnaires were randomly distributed along areas strategically selected according to the availability of public transportation facilities, i.e. bus stops and subway stations. 192 questionnaires were returned resulting in a recovery rate of 19.2 percent. Finally, 160 samples were used in the modeling process, since 32 questionnaires were incomplete or filled out with incorrect answers.

4. The Multinomial Logit Model

Initially, the data obtained by the survey was investigated so as to identify the most significant attribute variables in terms of modal choice. The well-known Multinomial Logit model was the basis for this analysis. The analysis showed that the attribute variables time-in-vehicle (TV), excess time (ET) and travel cost (C) were statistically the most significant attribute variables for modeling the transport modal choices. The Logit model developed in the present study was formulated as:

\[ U_{in} = \beta_i' + \beta_{TV} TV_i + \beta_{ET} ET_i + \beta_{C} C_i \]
\[ \forall i \in \{b : \text{bus}, s : \text{subway}, a : \text{auto}\} \]

\[ \Delta U^n_{ij} = (U_{in} - U_{jn}) \]  \hspace{1cm} (6)

\[ P_{in} = \frac{1}{1 + \sum_j e^{\Delta U^n_{ij}}, (\forall j \neq i)}, \]

where \( U_{in} \) is the utility of decision maker \( n \) for transport mode \( i \); \( \beta_i' \) is the parameter for alternative-specific constants for modes; \( \beta_{TV}, \beta_{ET}, \) and \( \beta_{C} \) are parameters for the measurable attribute variables of time-in-vehicle, excess time and travel cost, respectively; and \( P_{in} \) is the probability of selecting mode \( i \in A \) by decision maker \( n \), where \( A \) is the set of alternatives.

A thousand questionnaires were randomly distributed along areas strategically selected according to the availability of public transportation facilities, i.e. bus stops and subway stations. 192 questionnaires were returned resulting in a recovery rate of 19.2 percent. Finally, 160 samples were used in the modeling process, since 32 questionnaires were incomplete or filled out with incorrect answers.

5. Framework of the Neuro-Fuzzy Multinomial Logit Model

The neuro-fuzzy utility functions were estimated by using the optimization algorithm proposed by the Fuzzy Logic Toolbox (Version 2) of the software Matlab 7.0 [7] by the manner of trial and error. In the adaptive neuro-fuzzy optimization process, the named learning hybrid rule was applied for identifying the output parameters. The learning hybrid rule combines steepest descent (SD) and least squares estimator (LSE)
for identifying the parameters of the consequent part of the inferential rules \( i \). Specifically, three independent first-order Sugeno fuzzy inference systems \( i \) were developed and trained to describe the utilities of modes bus, subway and automobile. Each Fuzzy Inference System (FIS) is composed of nine inputs \((TV, ET, C)\) and one output \((Utility of mode i: Vi)\). The FISs were trained by using 80% of the eligible data (768 input-output pairs). The remaining 20% input-output pairs (192 pairs) were used as holdout data for verifying the model. The set of input variables were the same for the three FISs since the modal choice behavior of all travelers are assumed to be partially dependent on the characteristics of the three modes.

The general rule form of the fuzzy inference model for mode \( i \) in the \( k \) is:

\[
R_k^i : \text{If } TV_i \text{ is } A_{i,TV} \text{ and } ET_i \text{ is } A_{i,ET} \text{ and } C_i \text{ is } A_{i,C}, \ldots \text{then } s_{k}^i = \delta_{i,TV}^k TV_i + \delta_{i,ET}^k ET_i + \delta_{i,C}^k C_i + \varepsilon_k^i \quad (7)
\]

where \( TV, ET, \) and \( C \) are linguistic values corresponding to the input variables; \( s_{k}^i \) is a function in the consequent, which is assumed as the systematic component of the neuro-fuzzy utility functions; \( \delta_{i,Fk} (i \in (b, s, a), F \in (TV, ET, C)) \) and \( \varepsilon_k^i \) are the parameters of the membership functions, in which the latter reflects the unknown factors; \( K \) is the number of fuzzy rules; and \( A_{i,TV}, A_{i,ET}, \) and \( A_{i,C} \) are fuzzy sets in the antecedent of the fuzzy rules. \( A_{i,TV}, A_{i,ET}, \) and \( A_{i,C} \) characterize the possible attribute variables time-in-vehicle, excess time and travel cost, respectively, by bus, subway and automobile. The NFMNL model was based on the same attribute variables identified during the development of the MNL model as statistically the most significant for the transport modal choice so as to be compared with the classical MNL model. Figure 1 illustrates the Membership Functions (MF) for the attribute TV in the antecedent of the fuzzy rules. The other two attribute variables \((ET \text{ and } C)\) presented the same MF shape. In this study, \( s_{k}^i \) is a crisp function to be used for the Sugeno fuzzy models. In Sugeno fuzzy inference models, the crisp output function is frequently described by a polynomial in the fuzzy input variables, in which the most common is the first order polynomial \( i \).

The parameters of the antecedent part of the fuzzy inference rules were set up based on the evaluation of the characteristics of the input data set. As a process used by adaptive neuro-fuzzy inference systems, the initial values of the antecedent parameters can be defined in a way that the centers of the Membership Functions are equally spaced along the range of each input variable \( i \). Then, the parameters of the fuzzy rules, i.e. \( \delta_{i,Fk} (i \in (b, s, a), F \in (TV, ET, C)) \), are optimized to approach the final membership range.

In the NFMNL model, the subtractive clustering method was applied in order to deal with the problem of an exponentially-grown rulebase, which results from the large number of input variables. The reduced number of input variables included in the suggested model also contributed in reducing this problem. The subtractive clustering method as one of the fuzzy clustering methods is well-known as an effective tool for constructing linguistic interfaces based on experimental data \( i \). In short, the purpose of clustering is to partition a large data set into natural groups of data (clusters), which are able to represent the investigated system’s behavior. Based on the identified clusters, an initial Sugeno FIS can be developed with a reduced number of fuzzy rules to be used in the adaptive neuro-fuzzy training process \( i \). In this study, from the initial total number of 19,683 fuzzy rules \( i \) linguistic values), 31, 57 and 56 rules were extracted to represent the fuzzy systems of bus, subway and automobile respectively. This task was accomplished by using the subtractive clustering tool available in the Fuzzy Logic Toolbox of the software Matlab 7.0 \( i \).

The last stage in developing fuzzy systems is the defuzzification. The defuzzification is the process in which a crisp value is selected from the subset output of the fuzzy rules to be the representative value for the fuzzy system under interest \( i \). Since each rule for Sugeno inference system has a crisp output function, the defuzzification to be used for calculating the utility is simplified by applying the named weighted average method \( i \). Equation 8 shows the formulation used in this process. Then, the probability of individual \( i \) selecting mode \( i \) in the choice set \( C_0 \) can be written as in Equation 9.
\[ V_{in} = \left( \sum_{k=1}^{K} w_{ik}^s s_i^k \right) / \left( \sum_{k=1}^{K} w_{ik}^t \right), \forall i \in \{b, s, a\} \]

\[ = f_i(TV_i, ET_i, C_i, i \in \{b, s, a\} + e \quad (8) \]

\[ f_i: \text{linear function} \]

where:

\[ e = \sum_{k=1}^{K} w_{ik}^s \epsilon_i^k / \sum_{k=1}^{K} w_{ik}^t \text{ (error term)} \]

\[ P_{in} = 1 / \left[ 1 + \sum_j \exp^{V_i^j} \right], (\forall j \neq i), \quad (9) \]

where: \( \Delta V_i^{p} = (V_{a} - V_{i}) \): \( w_{ik}^s \) is the implication operator of the fuzzy rules, which is used for modeling the output fuzzy sets: \( s_i^k \) corresponds to the state variable obtained from the fuzzy rules, \( \epsilon_i^k \) reflects the unknown factors in the consequent of the fuzzy rules; and \( e \) is the error term. The calibration parameters can be obtained by minimizing Equation 4 with respect to \( \delta_{i,k} \) \( \in \{b, s, a\} \), \( F \in \{TV, ET, C\} \).

6. The Estimated Results

Firstly, the NFMNL model was evaluated by comparing its results with the results of the MNL model. For both models, the probabilities of selecting a transport mode were estimated and compared with the survey results. Figure 2(a) summarizes the utility functions and the transport modal choices obtained by applying the NFMNL model, while Figure 2(b) summarizes the results of the MNL model. These results show that both the NFMNL model and the MNL model were able to predict correctly higher preference by the mode subway, followed by the modes automobile and bus. Nonetheless, it is also noted that the NFMNL model presented better results for the estimation of modal choices if compared with the actual mode preferences. The index of root mean square (RMS) error was calculated in order to investigate the variation between the models estimation and the survey results, which is also shown in Figure 2. The obtained indexes confirmed better behavior of the proposed model. By comparing the NFMNL model estimation with the survey results, it was observed that the probabilities of selecting modes of bus and subway were underestimated while probability of selecting mode of automobile was overestimated. Despite this difference, the model demonstrated good performance.

Next, the performance of the NFMNL model was evaluated based on its predictive potential where the results are shown in Table 2. Overall, the NFMNL model succeeded in estimating 150 samples from the total of 192 data pairs, which provided 78.13% of fitness for the model. It was clear that the model presented good performance, especially by estimating the choices of modes subway and automobile. However, the number of correct predictions for bus was considered to be low. This may reflect a failure of the model in capturing the real influence of the value of walking time and waiting time on the choice of bus. In this study, walking time and waiting time were summed up to be the attribute variable of excess time as a restriction from the use of stated preference profiles in the data questionnaire. Moreover, there might be a possibility that the model overemphasized the poor reliability of mode bus in terms of time-in-vehicle, since it might vary widely as a result of delays particularly during the winter season if compared with the other modes.
Table 2 Estimation Results by the NFMNL Model

<table>
<thead>
<tr>
<th>Mode</th>
<th>Bus</th>
<th>Subway</th>
<th>Auto</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observation</td>
<td>4</td>
<td>11</td>
<td>6</td>
<td>21</td>
</tr>
<tr>
<td>Subway</td>
<td>1</td>
<td>111</td>
<td>19</td>
<td>131</td>
</tr>
<tr>
<td>Auto</td>
<td>0</td>
<td>5</td>
<td>35</td>
<td>40</td>
</tr>
<tr>
<td>Total</td>
<td>5</td>
<td>127</td>
<td>60</td>
<td>192</td>
</tr>
</tbody>
</table>

Samples correctly estimated 150
Fitness ratio (%) 78.13

7. Sensitivity Analysis
In order to demonstrate the influence of time variations of mode subway on the share probabilities, the sensitivity analysis was performed for both the NFMNL model and the MNL model (Figure 3). For this analysis, the values of the other attribute variables were assumed as shown in Figure 3(a). Figure 3(b) shows that the probabilities of selecting the three modes are almost constant until the time-in-vehicle of mode subway reaches 40 minutes in the NFMNL model. Then, the probability of selecting subway suddenly decreases while the probabilities of selecting bus and automobile increase. These results seem to be reliable when compared to the actual traveler’s preferences. Nevertheless, the modal share probabilities remain constant in the MNL model (Figure 3(c)). The results of the MNL model may not be accurate since changes of travelers’ behavior were not captured under significant changes of time of subway, although, time-in-vehicle was previously identified as a significant factor for explaining the modal choices. The difference of sensitivity between the models is caused by the different utility functions. On one hand, linear function was assumed for describing the utilities in the MNL model. On the other hand, more complex non-linear function was applied for describing the utilities in the NFMNL model, which seemed to be more realistic in modeling the share probabilities. Besides, the NFMNL model uses linguistic information for explaining the vagueness uncertainty regarding the modal choice. From the thresholds of the sensitive curves of NFMNL, primary insights on travelers’ sensitivity were captured. The results of the NFMNL model demonstrate that travelers shift to different modes whenever the travel time by mode subway becomes longer.

8. Conclusion
A multinomial logit model with adaptive neuro-fuzzy utility functions for the modal choice problem was suggested. The classical multinomial logit model is well able to explain the randomness uncertainty. However, it fails to explain the vagueness uncertainty inherent in the decision making process of shopping travelers. By incorporating linguistic variables in the NFMNL model, the vagueness uncertainty as well as the randomness uncertainty could be clarified. Consequently, transportation projects can be developed on the basis of more accurate transport modal shares’ estimation results.

A modest dataset is required for developing the NFMNL since it is developed based on disaggregate behavioral modeling. The NFMNL model has a more complex structure and needs more computing capabilities to be developed than the MNL model since it includes the neuro-fuzzy systems in its structure. Nevertheless, these extra efforts seem to be worth the gain in terms of the results accuracy. Additionally, the adaptive structure of neuro-fuzzy systems further enhances the development of travel behavior models. In adaptive systems, once the parameters
are identified they can be ported to fit any problem from a particular area.

The following points were noted regarding the results of the models. Firstly, the modal share probabilities demonstrated good performance for the proposed model compared to the results of the MNL model. Although the probability of selecting bus and subway were underestimated and the probability of selecting automobile was overestimated in the NFMNL model, the results seem to be reliable when compared with the survey results. The advantage of the NFMNL model was confirmed by the results of the root mean square error, which was considerably lower than that for the MNL model.

Secondly, the networks trained by using the NFMNL model showed a reasonable rate of correct answers in estimating transport modal choices in the validation data set. In particular, promising results were obtained for the estimation of subway and automobile. However, the model did not present a satisfactory performance in estimating bus choices. Possibly, the model was not able to capture properly the influence of time changes on the choice of bus. In order to deal with this problem, it is planned to improve the training data set.

Finally, the sensitivity analysis highlighted the differences in the behavior of the NFMNL model compared with the MNL model as a result of different utility functions. This analysis indicated reliable results for the NFMNL model as it could efficiently explain the relationship between explanatory attribute variables and utility of transport modes. The results of the NFMNL model suggest that travelers change their transport mode whenever the subway travel time lengthens. Moreover, these results showed that the longer the subway travel time becomes, the more sensitive the travelers are in choosing their transport mode. The variation on the travel time of the favorite transport mode might give primary insights on the traveler's sensitivity to changes of characteristics of the public transportation system. However, further analysis is necessary to evaluate the influence of other important explanatory factors for the modal choice.

The NFMNL model is a promising tool to be used in the transportation planning. Mainly, the success of the model is due to two reasons, in which the first is the possibility of dealing with the non-linear effects of the choice process, and the second is the ability of representing the vagueness uncertainty by including linguistic variables into the model. Other non-linear effects, such as squared functions and step functions, could also be included in the MNL modeling with the expectation of positive effects. However, the ambiguities relating to the vague process of choice making would not be effectively represented. Thus, satisfactory results could be reached by the hybrid NFMNL modeling.

For further studies, other important variables should be incorporated to the model, such as value of time, need for goods storage, and car ownership. The NFMNL model should also be compared to a more complex MNL model. Finally, the neuro-fuzzy multinomial logit model could be applied to other complex discrete choice problems in order to acquire further knowledge about different aspects of travel behavior.

References

Authors’ Contact
Katia Andrade kaandrade@gmail.com
Kenetsu Uchida uchida@eng.hokudai.ac.jp
Seiichi Kagaya kagayas@eng.hokudai.ac.jp