May 25, 2010
Professor José VIEGAS
Chair of 12th WCTR conference on Transport Research
Institute Superior Técnico (IST), Lisbon, Portugal

Dear Professor José VIEGAS,

I would like to submit the attached revised manuscript, " A Joint Model of Destination and Mode Choice for Urban Trips: A Disaggregate Approach," for presentation in 12th WCTR conference and for possible publication in a reputed scientific journal.
The manuscript was revised for reviewers' comments after full review. Responses to reviewers' comments are attached at the end of this manuscript. The manuscript was also edited for proper English language, grammar, punctuation, spelling, and overall style by American Journal Experts Co. and it should be English-ready for publication. The editorial certification is attached to this e-mail.

Sincerely,

Yousef Shafahi
Associate Professor and Chair
Department of Civil Engineering
Sharif University of Technology
P.O. Box 11365-9313
Azadi Avenue, Tehran, Iran
Tel: (98) (21) 66005818
Fax: (98) (21) 66004828
E-mail: shafahi@sharif.edu

# A JOINT MODEL OF DESTINATION AND MODE CHOICE FOR URBAN TRIPS: A DISAGGREGATE APPROACH

*SEYEDABRISHAMI, Seyedehsan, PhD Candidate in Transportation, Sharif University of technology, e_abrishami@civil.sharif.edu*

*SHAFAHI, Yousef, Associate Professor of Civil Engineering Department, Sharif University of Technology, shafahi@sharif.edu*

## ABSTRACT

Trip destination and mode choice are highly influenced by travelers' perceptions and behaviors; selecting a destination and a vehicle for a trip are two interdependent problems. This article presents and applies a disaggregate joint model for traveler destination and mode choice. The choice model uses fuzzy set and probability theory to deal with the uncertainty embedded in travelers' perceptions and behaviors. The model is structured as a decision tree in which fuzzy and non-fuzzy classification of influential variables regarding destination selection and mode choice expand the tree. Probability theory is utilized to extract choice probabilities from the decision tree. The most influential explanatory variables among all of the variables categorized for travelers' household, trip, and living zone specifications are selected based on minimizing the fuzzy entropy value. An aggregation method is designed to provide aggregate estimates for transportation planning based on a disaggregate choice model. A data set containing travelers' information from more than 9000 households in Shiraz, a large city in Iran, is used for model construction and evaluation. When compared with actual travel demand, the model's aggregate estimates of trip generation, distribution, and modal split indicate acceptable accuracy in terms of learning and the model is accurate enough to provide meaningful information and to enable generalization of the model's findings.

*Keywords: destination, disaggregate, fuzzy decision tree, joint model, mode choice*

## 1- INTRODUCTION

Current travel demand models in practice are estimated at the disaggregate level but used for aggregate level forecasting. Disaggregate models conceptualize travel demand apart from its modeling. Two kinds of uncertainty, travelers' individual perceptions and the randomness of their behavior, complicate disaggregate estimation of travel demand. The uncertainty of individual perception arises from the different perceptions held by travelers with regard to one reality. For example, actual travel time between an origin-destination pair may be perceived differently by different travelers. Uncertainty surrounding the randomness of travelers' behavior arises from differences in the behavior of different travelers with the same perceptions. For example, two similar travelers with the same perceptions of destination attractiveness may choose different destinations for their purposes.

Previous research has utilized probability theory to consider decision uncertainty. Choice models based on probability theory postulate that the probability of an individual choosing a given option is a function of socioeconomic characteristics and the relative attractiveness of the option [Ortuzar 2001]. The most common choice model based on probability theory is the multinomial logistic model (MNL), in which utility functions using linear equations are calibrated for each alternative and then the probability of choosing each alternative is calculated. However, some other choice models such as Probit, conditional Logit (CL), and nested Logit (NL) utilize for choice models. The main assumption of MNL is that unforeseen factors conform to a probability distribution. However, probability theory in conjunction with fuzzy theory can appropriately take all disaggregate models' uncertainties into account. According to the work of Zadeh [Zadeh 1995, 1996], probability theory is very useful when dealing with the uncertainty inherent in measurements or objects that can be measured; however, it is not very useful for dealing with the uncertainty embedded in human perceptions. The former issue involves crisp sets, while the latter involves fuzzy sets [Ross et al., 2002].

In recent years, many modelers have tried to take into account the uncertainties of travel demand modeling using fuzzy concepts, but many modelers ignore the possibility of using probability theory to complement fuzzy theory. This paper aims to use both probability and fuzzy theories for traveler destination and mode choice modeling; a fuzzy-probability disaggregated joint model that utilizes a decision tree analysis to extract an if-then rule base structure is presented and applied to a real-world problem. Destination and mode choices in traditional view of travel demand modeling (4-step) determines in two separate sequential steps. Since in this paper we forecast these two choices in a unique structure, therefore the structure called joint model. The decision tree (DT) is one of the most popular algorithms used in data mining and machine learning. The suggested algorithm used to construct a DT works by choosing variables from among all of the influential variables for travelers' destination and mode choice, categorized in three groups: household, trip, and traffic zone specifications. Some of the variables used in the DT structure are fuzzy; thus, the DT is called a fuzzy decision tree (FDT). Probability theory is used to determine travelers' destination and mode choice probabilities. Then a new aggregation method is used to provide aggregate estimates that are mostly appropriate for transportation planning. A data set containing trip data from more than 9000 households belonging to individuals living in Shiraz, a large city in Iran, is used to construct and evaluate model learning and to assess the generalizability of the aggregate estimations of this model.

Fuzzy theory and decision trees have been applied in previous work on travel demand modeling. Fuzzy theory has been used extensively to estimate travel demand at the aggregate and disaggregate levels [Teodrovic 1998, 1999]. Dell 'Orco and Kikuchi proposed a framework that utilized possibility and probability theories in transportation choice models. The former theory used for uncertainties embedded in travelers' decision making and the latter for analyst's uncertainty [Dell 'Orco and Kikuchi 2003]. Avineri believed that uncertainty involved in travel choices can be supported by a prospect theory. This theory models responses to uncertain situation as "gains" and "losses" over reference point. He utilized fuzzy theory to represent meaning of reference-based perceptions in the mind of traveler [Avineri 2008, 2009].

DT algorithms construct a tree structure or rule-base as a model using both discrete and continuous outputs. DTs are widely used for decision modeling because of their simplicity, interpretability, usefulness even with small data sets, and their ability to be combined with other decision techniques. DTs can be divided into three types: classification trees (when the outcome is discrete), regression trees (when the output is continuous) and classification and regression trees (CART) that include both kinds of outcomes [Breiman et al. 1984]. Discrete DT models are usually built based on information gain computed by the concept of entropy. These decision trees contain some nodes that stand for variables with a significant influence on decision-making; branches show the divisions defined for each variable. Algorithms like C4.5 [Quinlan 1993], C4 [Quinlan 1987], and ID3 [Mitchell 1997] are typically used to generate DT models.

A limited number of studies have applied C4 and C4.5 algorithms to construct predictive models of travel demand. Wets et al. applied the C4 algorithm to derive a decision tree for transport mode choice in the context of activity scheduling using a large activity diary data set [Wets et al. 2000]. Yamamoto et al. analyzed drivers' route choice behavior using the C4.5 algorithm. The results of a comparative analysis between the C4.5 algorithm and discrete choice models indicate the superior ability of former to represent driver route choice [Yamamoto 2002]. Thill and Wheeler analyzed the applicability of Ross Quinlan's C4.5 decision tree inference algorithm to the class of problems involving the choice among travel destinations within an urban area [Thill & Wheeler 2000]. Xie et al. investigated the capability and performance of work travel mode choice modeling with two emerging pattern-recognition data mining methods: decision trees and neural networks. The modeling results showed that the DT model demonstrates the highest estimation efficiency and the best interpretability compared to neural networks and the traditional multinomial logit model [Xie et al. 2003].

The remainder of this paper is organized in four sections. Section 2 discusses FDT construction and the inference algorithm and section 3 applies the constructed FDT to a real-world problem and suggests a new aggregation method. The FDT model is evaluated in section 4, and section 5 summarizes the main conclusions of the paper.

# 2- FDT CONSTRUCTION AND INFERENCE ALGORITHMS

The algorithms used to construct decision trees work by selecting the best variable at each step for splitting the decision trees' nodes into branches. The "best" variable is defined by how well the variable splits the set into subsets that have the same value of the target variable. Different algorithms use different formulae to assess variable quality. In this article, fuzzy entropy is used as a measurement to identify the best variables.

Before describing the suggested algorithm for FDT construction and inference, some notations and assumptions should be introduced:

- The set of variables is denoted by $V = \{V_1, V_2, \ldots, V_n\}$, and the value of each variable is shown by $v_i$.
- Fuzzy variables are indicated by a tilde (~) placed over a variable $(\tilde{V}_i)$.
- The set of linguistic terms assigned to a fuzzy variable ($\tilde{V}_i$) or the set of categories assigned to a non-fuzzy variable ($V_i$) are indicated by $T = \{t_1, \ldots, t_l, \ldots, t_p\}$. For example, if travel time is assumed to be a fuzzy variable, the set $T$ can be $\{Short, Long\}$; household car ownership is a non-fuzzy variable, so the set $T$ can be $\{0, 1, 2+\}$.
- $v_{t_l}^i$ shows the $l$ th linguistic term ($t_l \in T$) assigned to the $i$ th variable $\tilde{V}_i$ (e.g., $v_{Short}^{TravelTime}$) or the $l$ th category ($t_l \in T$) assigned to the $i$ th variable $V_i$ (e.g., $v_{2+}^{Household\ Car\,ownership}$).
- $D = \{d_1, \ldots, d_k, \ldots, d_m\}$ denotes the set of decision alternatives (destination and mode choice).
- The training data set is indicated by $E$, which contains $e_j(\in E) = (v_1, \ldots, v_i, \ldots, v_n, D)$.
- $N$ stands for the decision tree's node.
  - $NP$ stands for the parent node of $N$.
  - $NC$ stands for the child node of $N$.
  - $N^{Root}$ stands for the root node.
- $\chi^N$ is the set of membership degrees of training elements in node $N$.
  - $\chi^N(e_j)$ is the membership degree of $e_j(\in E)$ in node $N$.
  - $\chi^N(e_j)\big|v_{t_l}^i$ is the membership degree of $e_j(\in E)$ in node $N$ with regard to $v_{t_l}^i$.
- $\mu_{t_l}^{ij}$ for fuzzy variables is the membership degree of the $j$ th element of the training set ($e_j \in E$) belonging to the linguistic term $t_l(\in T)$ of the $i$ th variable ($\tilde{V}_i \in V$), however for non-fuzzy variables expressed as:
  - $\mu_{t_l}^{ij}$=1 if the $j$ th element of the training set ($e_j \in E$) belongs to category $t_l(\in T)$ of the $i$ th variable ($V_i \in V$); otherwise, $\mu_{t_l}^{ij}$=0.
- $\mu_{d_k}^{j}$ is the membership degree describing the choice possibility of the $k$ th decision alternative ($d_k \in D$) by the $j$ th element of the training set ($e_j \in E$).
- $V^N$ is the set of variables that appear on the path leading to node $N$.
- $C^N\big|v_{t_l}^i$ is the cumulative membership degree of training elements in node $N$ with regard to $v_{t_l}^i$.
  - $C_k^N\big|v_{t_l}^i$ is the cumulative membership degree of training elements in node $N$ for the $k$ th decision alternative ($d_k \in D$) with regard to $v_{t_l}^i$,

- $I^N$ is the entropy of node $N$.

The algorithm used for FDT construction is similar to Janikow's algorithm [Janikow 1998] with some changes to make it compatible with destination and mode choice problems. The algorithm's steps are as follows:

Step 1- Start with the tree root node ($N^{Root}$) and all elements of the training data set. Let $N = N^{Root}$ and $V^N = \phi$.

Step 2- At any node $N$ that must still be expanded, compute $\chi^N$ for the remaining variables ($V - V^N$):

$$\chi^N(e_j)\Big|v_{t_l}^i = Min\left\{\mu_{t_l}^{ij}, \chi^{NP}\left(e_j\right)\right\}.$$

(Note 1: when the parent node is the tree root, $\chi^{NP}\left(e_j\right)=1$ for all $j$).

Step 3- Compute $C_k^N\Big|v_{t_l}^i$ and $C^N$ for each linguistic term or category assigned to the fuzzy and non-fuzzy variables, respectively:

$$C_k^N\Big|v_{t_l}^i = \sum_{j=1}^{|E|}\chi^N\Big|v_{t_l}^i\left(e_j\right)\times\mu_{d_k}^j \quad \forall\ d_k \in D \text{ And } (\tilde{V}_i \in V \text{ or} V_i \in V)$$

$$C^N\Big|v_{t_l}^i = \sum_{k=1}^{m}C_k^N\Big|v_{t_l}^i .$$

Step 4- Compute $I^N$:

$$I^N = \frac{\sum_{l=1}^{p} I^N\Big|v_{t_l}^i \times C^N\Big|v_{t_l}^i}{\sum_{i=1}^{l} C^N\Big|v_{t_l}^i} ,$$

where:

$$I^N\Big|v_{t_l}^i = -\sum_{k=1}^{m}\left(\frac{C_k^N\Big|v_{t_l}^i}{C^N\Big|v_{t_l}^i}\times\log\frac{C_k^N\Big|v_{t_l}^i}{C^N\Big|v_{t_l}^i}\right) .$$

Step 5- If no variables remain or if any other stopping criterion is met, tree expansion stops and all nodes at the end of branches change to leaves; otherwise, one of the remaining variables that minimizes the node's value of $I^N$ is selected to split node $N$, and the process returns to step 2.

This algorithm starts from the root node of the tree and expands the FDT according to information contained in the training data set. The differences between FDTs and DTs include the following: the DT elements of the training data set belong to one branch of each node, but in an FDT, these belong to some of the node's branches with probability degrees; additionally, while the typical DT node is divided into two branches, FDT nodes may be divided into two or more branches based on the linguistic terms assigned to fuzzy or categories assigned to non-fuzzy variable.

Leaves located at the ends of branches specify decision alternatives selected by decision makers. The major difference between FDT and DT leaves is that the latter makes a distinct prediction of decision for the decision maker, while the former predicts a set of decisions with different probabilities.

One similarity between the two decision trees is the concept of entropy that is used to select influential variables to expand the trees. The FDT uses a minimum operator to compute elements' degree of membership for each node and a multiplication operator to calculate the cumulative membership degrees of elements at each node. The two operators are t-norm operators.

The set of choice probabilities at each leaf is extracted from the training data set elements that have non-zero membership degrees in the leaf. This set may be computed by different methods. This paper uses three different methods based on probability theory to take travelers' randomness of behaviors into account.

The first method considers all training elements reaching to the leaf and computes the probabilities according to the equation below:

$$P_l^k = \frac{n_l^k}{n_l},$$

where:

$$n_l^k = \left|E_l^k\right| \quad where \;\; E_l^k(\subset E) = \left\{e_j \middle| \chi^l(e_j) \neq 0 \; and \; \underset{z}{Max}\left(\mu_{d_z}^j\right) = \mu_{d_k}^j\right\}.$$

$E_l^k$ is a subset of $E$ in which training elements have a non-zero membership degree at leaf $l$ and select decision alternative $k$ with the maximum possible degree. $n_l^k$ is the size of the $E_l^k$ set.

In the second method, choice probabilities are influenced by an element's degree of belonging to leaves. Therefore, choice probabilities are computed using the following equation:

$$P_l^k(G_i) = \frac{n_l^k(G_i)}{n_l(G_i)}.$$

The degree of belonging to leaves can be divided into three groups, $G_1$= (0, 0.33), $G_2$= [0.33 0.66), and $G_3$= [0.66 1.0]. $E_l(G_i)$ is the set of elements with a degree of belonging to leaf $l$ in the range of group $G_i$.

$$n_l^k(G_i) = \left|E_l^k(G_i)\right| \qquad E_l^k(G_i) = \left\{e_j \middle| \chi^l(e_j) \in G_i \; and \; \underset{z}{Max}\left(\mu_{d_z}^j\right) = \mu_{d_k}^j\right\}$$

$P_l^k(G_i)$ is the choice probability of decision alternative $k$ for elements with a degree of belonging to leaf $l$ in the range of group $G_i$.

In the third method, membership degrees of elements to decision alternatives ($\mu_{d_k}^j$) are adjusted by the elements' degree of membership for each leaf. The choice probability of decision alternative $k$ at leaf $l$ is computed by the following equation:

$$P_l^k = \frac{\dfrac{1}{\left|E_l^k\right|} \times \sum\limits_{e_j \in E_l^k} \left(\chi^l(e_j) \times \mu_{d_k}^j\right)}{\sum\limits_{k=1}^{m} \sum\limits_{e_j \in E_l^k} \left(\chi^l(e_j) \times \mu_{d_k}^j\right)}.$$

The different $P_l^k$ definitions are used for the inference algorithm from FDT. However, some other definitions may be possible. The following inference algorithm has been applied to draw conclusions from the FDT. Assume that the data set $R$ contains elements $r_j = (v_1, \ldots, v_n)$. The following steps show how FDT estimates choice probabilities for each element.

Step 1- Calculate for each $r_j(\in R)$ the membership degree of belonging to each leaf $l$. If the total number of leaves is assumed to be equal to $L$, a vector with $L$ elements will indicate membership degrees for each $r_j$. The membership degree for an element $j$ at leaf $l$ is:

$$\chi^l(r_j) = \underset{i}{Min}\,\{\mu_{t_l}^{ij}\} \quad \forall V_i \in V^l .$$

The notation definitions for data set $R$ are the same as those for data set $E$ described in previous paragraphs.

Step 2- Compute the cumulative choice probability of decision alternatives over all leaves. Assume that $P(r_j, k)$ is the choice probability of decision alternative $k$ by decision maker $r_j$:

$$P(r_j, k) = \frac{\sum_{l=1}^{L}\left[\chi^l(r_j) \times P_l^k\right]}{\sum_{k=1}^{m}\sum_{l=1}^{L}\left[\chi^l(r_j) \times P_l^k\right]} .$$

Finally, the choice probabilities of available decision alternatives are estimated in step 2. However, an aggregation method is suggested to aggregate travel demand estimates of FDT to provide appropriate estimations for transportation planning.

# 3- FDT APPLICATION IN SHIRAZ

This study uses trip data obtained from the Shiraz Comprehensive Transportation Study (SCTS) conducted in 2000. The Shiraz study area is 20550 km$^2$ and includes 1.98 million people. The city of Shiraz, which has a population of 1.15 million people, is in the center of the study area. Shiraz is divided into 156 traffic zones in 15 regions. An Origin/Destination survey with household interviewing by trained people was conducted in the SCTS area. The data were gathered from 4 percent of households in Shiraz and validated by observations from 4 screen lines in the study area.

The data gathered from the randomly selected households are divided into two main categories: general household information (e.g., household members' age/sex/occupation, number of vehicles owned by household members, address) and information on all motorized travel by household members on the date of interviews (origin and destination of travel, ending and starting time, travel mode and purpose, etc.). Travel information was gathered from more than 65000 members of over 9000 households. Private mode and transit travel times between pair of origin and destination were obtained from EMME/2 software.

Since more than 85 percent of trips are home-based[1] (HB) and over 20 percent of HB trips occur in the morning peak-hour, daily travel information is converted to daily travel HB morning peak-hour trips (HBP). The HBP data set contains information from 9177 travelers. Approximately 75 percent (6883 elements) of total data is used to construct the FDT; the rest of the data (2294 elements) are used to test the FDT results.

## 3-1- FDT Construction

The candidate variables assumed to influence destination and mode choice in this study are:

- Household characteristics: household car ownership (HC), household size (HS)
- Trip specification: purpose
- Zone characteristics: zonal car ownership (Zcar), zone distance to CBD (Dist)

The household car ownership variable is categorized into three groups: without private car, one private car, and two or more private cars. Household size is also categorized into three groups: small (1 to 3 members), medium (4 to 6 members), and large (7 or more members). The purpose variable is a categorical variable that includes work, school, shopping, recreation, and personal. All of these variables are assumed to be non-fuzzy variables because these variables are usually exact and not influenced by travelers' perceptions. The zone-related variables, which are usually inexact, are considered as fuzzy variables defined by linguistic concepts because zone-related variables may have a variety of effects on traveler destination and mode choice. Zonal car ownership that is average car ownership at each zone can be an indicator of household income level at each zone. The household income level is not easily obtainable but experts can estimate this variable by using linguistic terms for zonal car ownership. Fuzzy is an appropriate tool that can define linguistic terms thus we use fuzzy to define this variable. Common sense suggests that two linguistic terms: low and high, assign to zonal car ownership. Two linguistic terms: close and far assign to variable distance to CBD according to common sense.
The fuzzy C-means (FCM) [Bezdek 1981] algorithm is used to compute membership function parameters. Figures 1 and 2 illustrate the fuzzy sets defined for each variable.

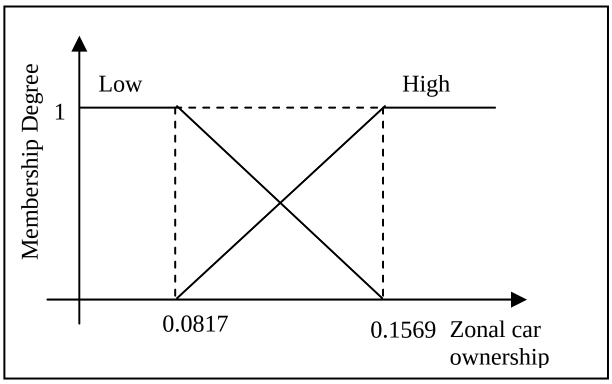


Figure 1- Zonal car ownership membership functions

[1] Originate or end at home zones.

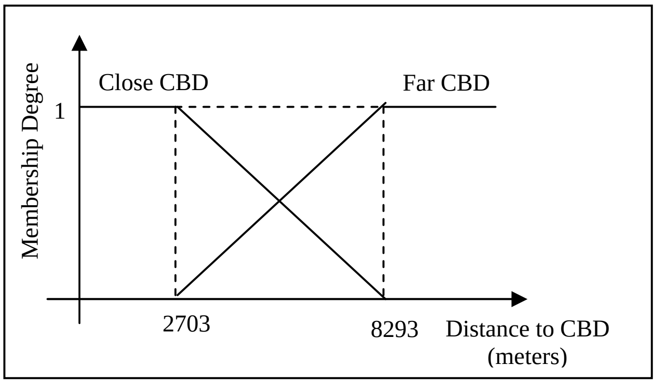



Figure 2- Distance to CBD membership functions

The definitions of decision alternatives will effectively influence FDT performance. In this study, destination zone distance to CBD in addition to travel time between origin and destination are selected as criteria for defining decision alternatives. Distance to CBD usually indicates zone accessibility by travel modes and zone's socio-economic activities. Private and transit ravel time between origin-destination pair are assumed. Membership functions are used to define these two variables, distance to CBD and travel time. Membership functions for zone distance to CBD is the same with Figure 2. Travel time is classified using two linguistic terms: short and long. Figure 3 provides the membership functions. The process of labeling decision alternatives is illustrated in Figure 4; eight decision alternatives are named in the boxes in the right column. These names stand for linguistic terms assigned to both variables: destination distance to CBD and travel time between the origin-destination pair.

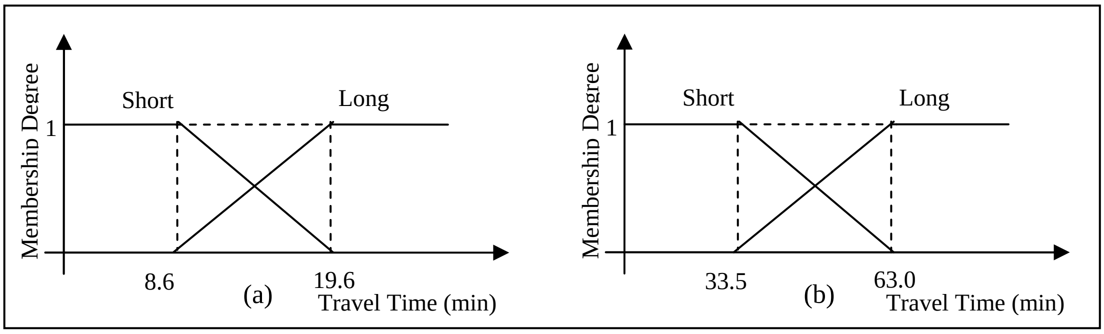



Figure 3- Travel time membership functions (a) private (b) transit

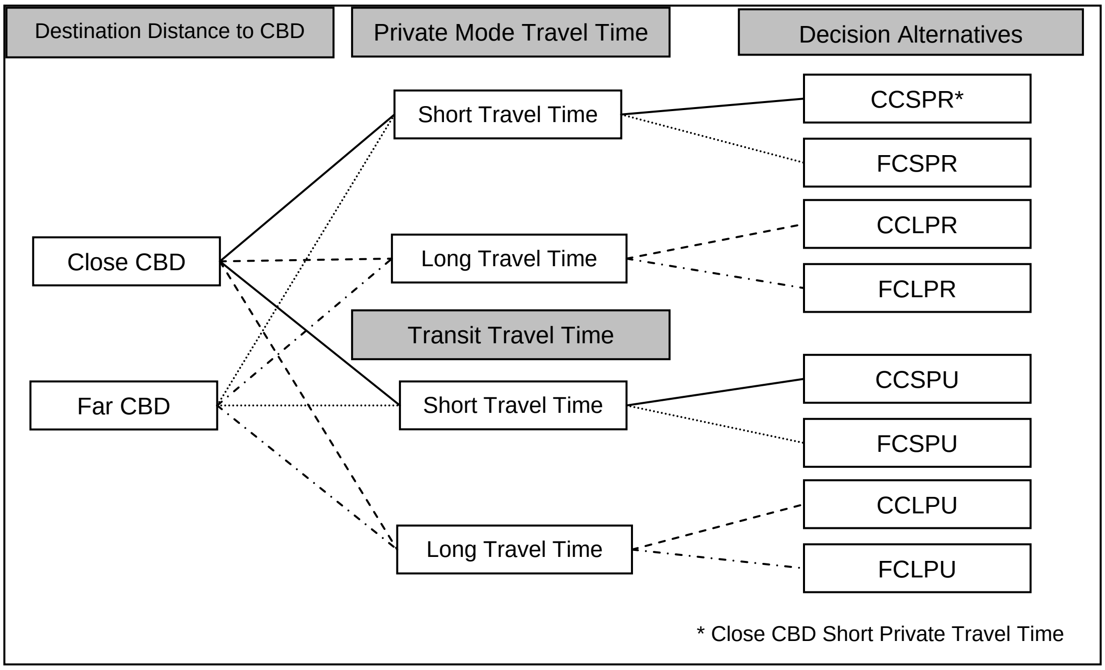



Figure 4-Labeling of decision alternatives for the destination and mode choice problem

Each training element's degree of belonging to decision alternatives is the product of membership degrees of travel time between the origin and destination with the destination distance to CBD. An example is presented (Tables 1 to 3) to clarify how membership degrees for both input and decision alternatives are computed; five elements from the training data set are used as examples. Tables 1 and 2 show how the data are pre-processed and prepared for FDT construction.

Table 1-Data before pre-processing for some examples

| No. | Input Variables | | | | | | Destination & Mode | | Travel Time (min) | |
|---|---|---|---|---|---|---|---|---|---|---|
| | HC | HS | Purpose | HZ@ | Dist (meters) | Zcar | Travel Mode$ | Destination Zone@ | Private | Public |
| 1 | 2 | 5 | Shopping | 35 | 2010.6 | 0.178 | 2 | 7 | 8.10 | 31.30 |
| 2 | 2 | 3 | Work | 47 | 437.8 | 0.123 | 1 | 43 | 4.80 | 18.30 |
| 3 | 0 | 6 | Recreation | 99 | 540.7 | 0.091 | 1 | 142 | 17.00 | 52.80 |
| 4 | 1 | 8 | Recreation | 110 | 2310.1 | 0.047 | 2 | 9 | 8.40 | 32.00 |
| 5 | 1 | 2 | School | 117 | 899.7 | 0.109 | 2 | 28 | 6.70 | 26.80 |

$1=private, 2=public   @ zone number (1-156)

Table 2- Input data after pre-processing for some examples

| No. | Input Variables | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | HC | HS | Dist | | Zcar | | Purpose | | | | |
| | | | Close | Far | Low | High | Work | School | Shopping | Personal | Recreation |
| 1 | 2+ | Medium | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 |
| 2 | 2+ | Small | 0.94 | 0.06 | 0.44 | 0.56 | 1 | 0 | 0 | 0 | 0 |
| 3 | 0 | Medium | 0.29 | 0.71 | 0.86 | 0.14 | 0 | 0 | 0 | 0 | 1 |
| 4 | 1 | Large | 0.33 | 0.67 | 1 | 0 | 0 | 0 | 0 | 0 | 1 |
| 5 | 1 | Small | 0 | 1 | 0.63 | 0.37 | 0 | 1 | 0 | 0 | 0 |

Table 3- Definitions of decision alternatives for available data for some examples

| Destination Zone | | | Travel Time | | | | Decision Alternatives | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Distance to CBD | | | Private | | Public | | Private | | | | Public | | | |
| Distance | Close (E) | Far (F) | Short (A) | Long (B) | Short (C) | Long (D) | CCSPR (A×E) | FCSPR (A×F) | CCLPR (B×E) | FCLPR (B×F) | CCSPU (C×E) | FCSPU (C×F) | CCLPU (D×E) | FCLPU (D×F) |
| 536.5 | 1.00 | 0 | 1.00 | 0 | 1.00 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 899.7 | 1.00 | 0 | 1.00 | 0 | 1.00 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8351.2 | 0 | 1.0 | 0.24 | 0.76 | 0.35 | 0.65 | 0 | 0.24 | 0 | 0.76 | 0 | 0 | 0 | 0 |
| 1260 | 1.00 | 0 | 1.00 | 0 | 1.00 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 1602.4 | 1.00 | 0 | 1.00 | 0 | 1.00 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |

Table 2 presents the input data preparation and pre-processing for FDT construction and inference, and Table 3 illustrates the process of computing degrees of belonging to eight alternatives.

In the first step, pre-processed training data are used to select the root node's variable for expanding the node. Entropy values for each variable at the root node are displayed in Table 4. The purpose variable has the minimum entropy; thus, this variable is selected for FDT root node expansion. The root node divides into five branches, with each branch specifying a trip purpose as shown in Figure 5.

Table 4- Entropy for all variables at the root node

| Variable | HC | HS | Purpose | Dist | Zcar |
|---|---|---|---|---|---|
| Entropy | 0.81 | 0.77 | 0.74 | 0.76 | 0.78 |

In the next step, the variable related to each of the five new nodes appearing under the root node according to the training data should be selected. Training data are divided into five subsets based on trip purpose. Entropy calculation at each node specifies the selected variable for each node. Home zone distance to CBD (Dist) is selected for all five nodes as indicated in Figure 5. The stopping criterion for FDT expansion is assumed to be the number of training data set members at each node; thus, when the number of members falls below 1000, node expansion stops. The FDT branching results in 33 nodes with 52 leaves. In other words, the number of rules included in the rule base structure of the model is 52, equal to the number of leaves. Figure 5 shows all of the branches in detail. Each leaf specifies choice probabilities for eight decision alternatives, for example Table 5 shows these probabilities for work trips. Zones in CBD or close to CBD usually attract many of work trips, thus decision alternatives indicate zones close to CBD have higher probabilities. Additionally, travelers usually select destination within a short travel time from home zone for work trips. Therefore, choice probabilities of short travel time with private or public transportation are higher than the rest. Choice probabilities for "CCSPR" show that leaves 4 to 7 and 10 to 12 that related to households with one or more than one private car usually utilize their own car for work trips.

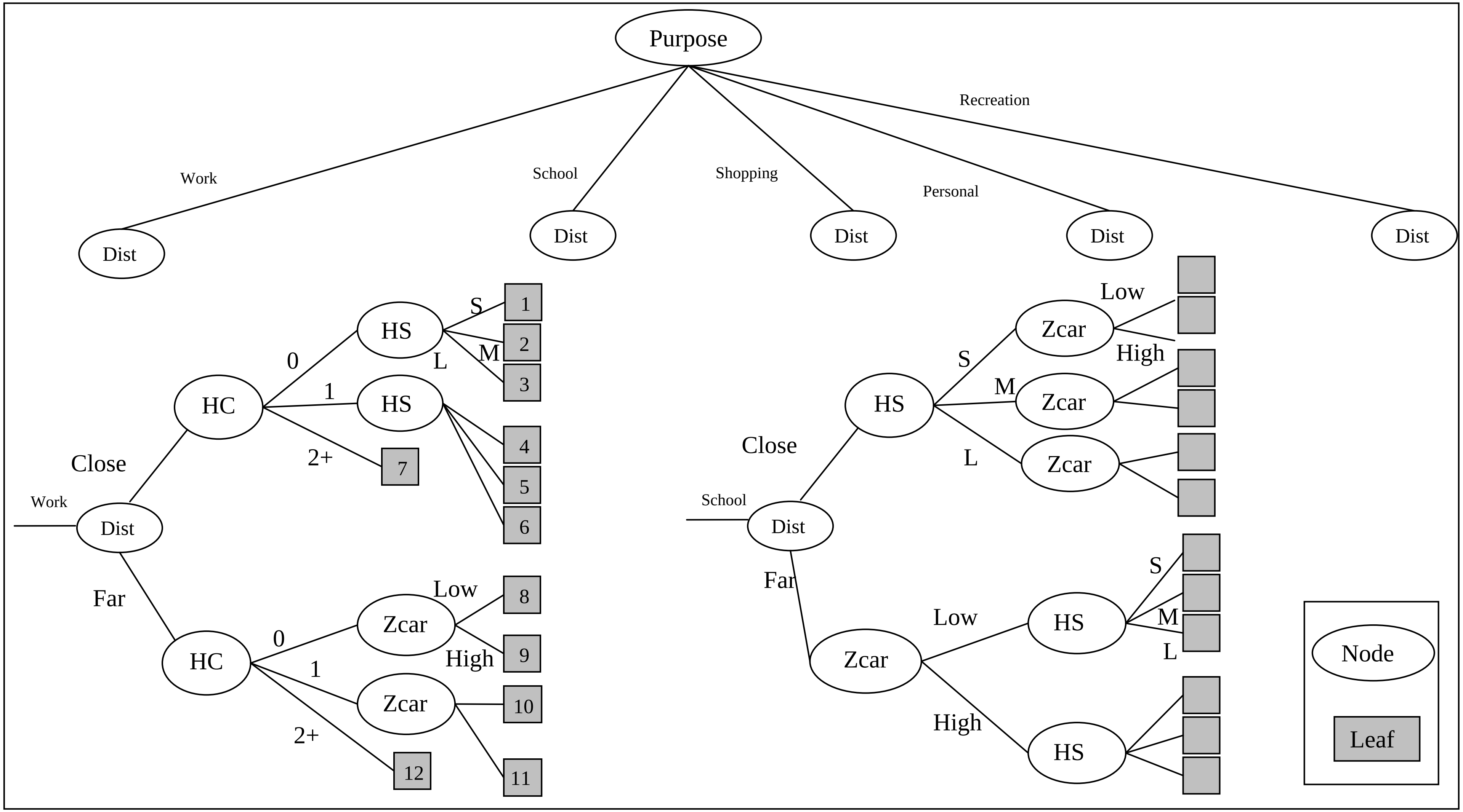



Figure 5- FDT structure details.

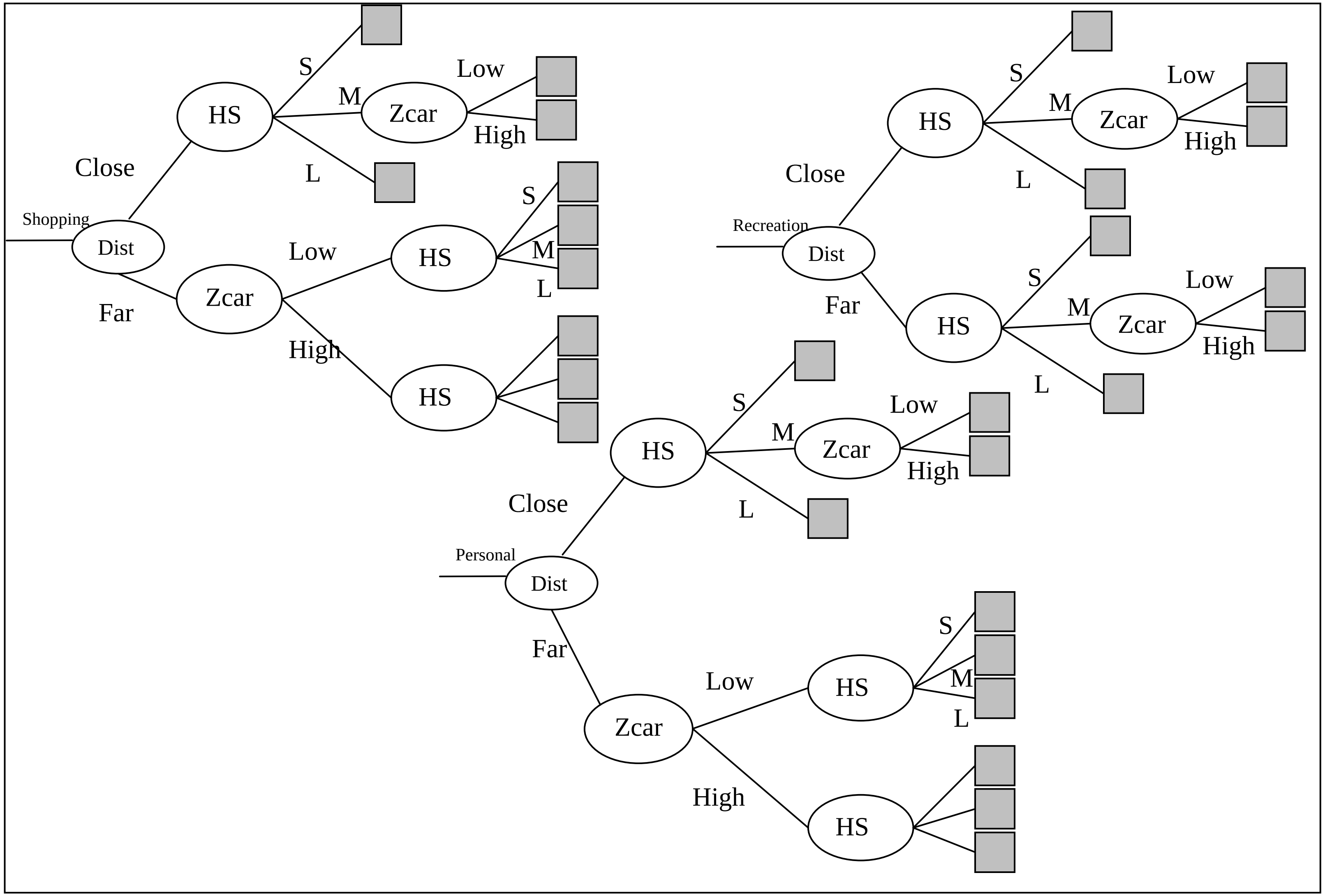



Figure 5 (continued) - FDT structure details

Table 5- Choice Probabilities for Work Trips in Different Leaves

| Leaf No. | CCSPR | FCSPR | CCLPR | FCLPR | CCSPU | FCSPU | CCLPU | FCLPU |
|---|---|---|---|---|---|---|---|---|
| 1 | 0.02 | 0.00 | 0.01 | 0.00 | 0.58 | 0.09 | 0.20 | 0.11 |
| 2 | 0.02 | 0.00 | 0.01 | 0.00 | 0.53 | 0.10 | 0.22 | 0.11 |
| 3 | 0.01 | 0.00 | 0.00 | 0.00 | 0.53 | 0.11 | 0.23 | 0.11 |
| 4 | 0.45 | 0.08 | 0.19 | 0.05 | 0.14 | 0.02 | 0.05 | 0.01 |
| 5 | 0.37 | 0.08 | 0.18 | 0.06 | 0.19 | 0.03 | 0.07 | 0.02 |
| 6 | 0.27 | 0.09 | 0.13 | 0.03 | 0.25 | 0.05 | 0.13 | 0.05 |
| 7 | 0.46 | 0.08 | 0.19 | 0.07 | 0.12 | 0.01 | 0.06 | 0.02 |
| 8 | 0.01 | 0.00 | 0.01 | 0.00 | 0.44 | 0.12 | 0.30 | 0.12 |
| 9 | 0.01 | 0.01 | 0.01 | 0.00 | 0.41 | 0.15 | 0.31 | 0.10 |
| 10 | 0.29 | 0.10 | 0.24 | 0.05 | 0.15 | 0.04 | 0.11 | 0.02 |
| 11 | 0.29 | 0.11 | 0.26 | 0.05 | 0.14 | 0.04 | 0.10 | 0.02 |
| 12 | 0.32 | 0.10 | 0.31 | 0.07 | 0.07 | 0.01 | 0.09 | 0.02 |

Distance to CBD is an important factor in selecting trip destination and mode because distance indicates the access of home zone inhabitants to transit systems (transit systems are usually more accessible near CBD) as well as their income level. The FDT shows that travelers usually use private cars for work trips because household car ownership is more significant than other variables in work trip branch. Household size plays a key role in all branches.

## 3-2-FDT Aggregation Method

The estimates of disaggregate models should be aggregated for application to transportation planning. Five procedures are typically used in the literature for aggregation of disaggregate models: average individual, classification, statistical differentials, explicit integration, and sample enumeration. Each of the methods has strengths and weaknesses [Ben-Akiva & Lerman 1985]. These aggregation methods cannot be used for the FDT model because of the special structure of the fuzzy decision tree. Thus, this article presents and applies a new aggregation method as follows:

Step 1 (input definition) - Define the inputs of the FDT model for the aggregation method:

- Number of households in each zone and household trip rates: The rows and columns in this table indicate household size and household car ownership, respectively. The table elements are number of households in each zone categorized by household size and car ownership. Figure 6 schematically illustrates the required table. A table similar to the one shown in Figure 6 is also used to specify household trip rates differentiated according to trip purpose; the rates in this table are computed using data from the field survey.

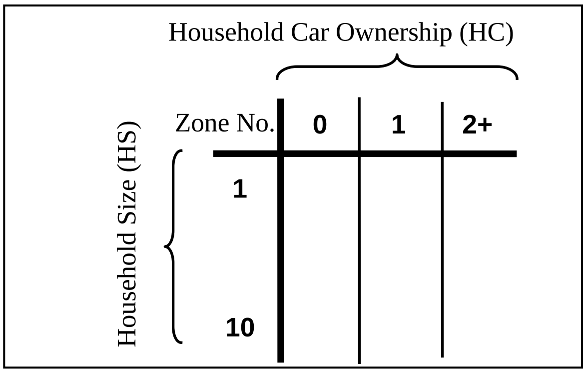



Figure 6- Schematic table for the number of households in each zone categorized by household size and car ownership

- Zone specification table: Figure 7 shows a zone specification table that indicates the membership degrees for zone distance to CBD as well as zonal car ownership using the relative linguistic terms defined for both variables. $\mu_i^{Close}$, $\mu_i^{Far}$, $\mu_i^{Low}$, and $\mu_i^{High}$ indicate zone $i$ membership degrees using the linguistic terms "close", "far", "low", and "high", respectively. These membership degrees are computed using the membership functions shown in Figures 1 and 2. The first two linguistic terms define zone distance to CBD; the second two define zonal car ownership.

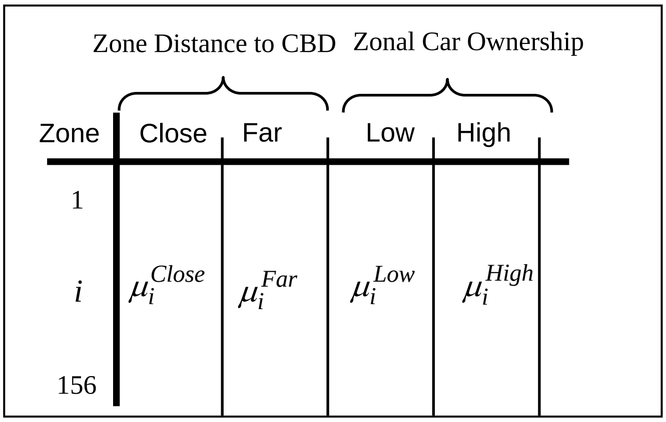



Figure 7- Schematic zone specification table

- Origin-destination specification table: This table provides the degree of belonging for origin-destination travel times in linguistic terms (short and long) for both the private and transit modes. Figure 8 displays the table configuration schematically. All combinations of $i$ ($i = 1,...,156$) and $j$ ($j = 1,...,156$), equal to 24336 (156×156) rows will appear in this table.

| | | Private | | Transit | |
|---|---|---|---|---|---|
| Origin | Destination | Short | Long | Short | Long |
| $i$ | $j$ | $\mu_{ij}^{Short}$ | $\mu_{ij}^{Long}$ | $\mu_{ij}^{Short}$ | $\mu_{ij}^{Long}$ |

Figure 8- Schematic origin-destination specification table

Step 2 (degree of belonging to leaves) –the degree of belonging to each leaf for the table of number of households presented in Figure 6 is computed. The cells of table presented for number of households (Figure 6) in conjunction with the table shown in Figure 7 provide all the input data needed to determine degree of belonging to the tree's leaves. HC and HS are obtained through the table of household numbers, Zcar and Dist found in the zone specification table shown in Figure 7. The degree of belonging to each leaf is computed by applying t-norm operators (like minimization or multiplication) of inputs' membership degrees to the linguistic terms and categories appearing in the path leading to that leaf.

Step 3 (alternative probabilities) - The degree of belonging to each leaf is multiplied by the probability of selecting a decision alternative at each leaf (three methods of extracting probabilities were previously described); in this way, the probability of selecting each alternative (destination and mode choice) for a household of specified car ownership and size is determined.

Step 4 (distribution of trips between alternatives) - Number of trips is computed based on the rate of trip for different trip purposes and distributed according to the probability of decision alternatives calculated in the previous step. The number of trips for each alternative is calculated separately for each traffic zone.

Step 5 (distribution of trips between zones) – The origin-destination table in Figure 8 together with the zone specification table determines the destination zone degree of belonging to each decision alternative. Thus, the number of total trips in the previous step is distributed between destination zones with the selected mode (private or transit).

The above process should be executed for all traffic zones of the study area to estimate an origin-destination trip matrix with a specific purpose and mode. The estimation and aggregation process for one traffic zone (zone no.14) is explained in more detail to clarify the process. The following table shows the number of households living in zone 14. Table rows and columns show the household size and car ownership, respectively.

Table 6- Number of households living in zone 14

| 14 | 0 | 1 | 2+ |
|---|---|---|---|
| 1 | 5 | 3 | 0 |
| 2 | 72 | 22 | 0 |
| 3 | 64 | 51 | 3 |
| 4 | 60 | 75 | 2 |
| 5 | 53 | 71 | 4 |
| 6 | 32 | 37 | 2 |
| 7 | 18 | 32 | 0 |
| 8 | 10 | 4 | 3 |
| 9 | 6 | 4 | 2 |
| 10 | 2 | 1 | 0 |

Degrees of belonging to the tree's leaves for each cell of the above table are computed. These values can be shown in a 30×12 matrix for trips made for work purposes (Table 6 contains 30 cells, and the work branch has 12 leaves). The matrixes for trips made for school, shopping, recreation, and personal purposes are 30×12, 30×10, 30×8, and 30×10, respectively. These matrixes are multiplied by the choice probabilities of the decision alternatives at each leaf; these probabilities are included in an m-by-n matrix for each trip purpose (m: number of leaves under trip purpose and n: number of decision alternatives). After multiplication, 30×8 matrixes for each trip purpose are generated. Each row of these matrixes specifies choice probabilities of decision alternatives for each cell of Table 6. Table 7 shows the choice probabilities of the decision alternatives for work trips in traffic zone 14 for households with three members.

Table 7- A part of choice probability matrix of decision alternatives for work trips in traffic zone 14

| HS | HC | CCSPR | FCSPR | CCLPR | FCLPR | CCSPU | FCSPU | CCLPU | FCLPU |
|---|---|---|---|---|---|---|---|---|---|
| 3 | 0 | 0.02 | 0.00 | 0.01 | 0.00 | 0.57 | 0.09 | 0.20 | 0.11 |
| 3 | 1 | 0.45 | 0.08 | 0.19 | 0.05 | 0.14 | 0.02 | 0.05 | 0.02 |
| 3 | 2+ | 0.46 | 0.08 | 0.19 | 0.07 | 0.12 | 0.00 | 0.06 | 0.02 |

After the number of trips is computed using trip rates, these trips are distributed across the decision alternatives based on choice probabilities. Summation of the number of trips for each decision alternative results in the total number of trips generated in example zone 14. Table 8 shows part of one of these matrixes for work trip purposes as an example.

Table 8- A part of number of work trips matrix for traffic zone 14

| HS | HC | CCSPR | FCSPR | CCLPR | FCLPR | CCSPU | FCSPU | CCLPU | FCLPU |
|---|---|---|---|---|---|---|---|---|---|
| .<br>.<br>. | .<br>.<br>. | .<br>.<br>. | .<br>.<br>. | .<br>.<br>. | .<br>.<br>. | .<br>.<br>. | .<br>.<br>. | .<br>.<br>. | .<br>.<br>. |
| 3 | 0 | 1.18 | 0.32 | 0.76 | 0.29 | 44.91 | 6.94 | 15.50 | 8.33 |
| 3 | 1 | 38.12 | 7.14 | 15.63 | 4.21 | 12.12 | 1.53 | 4.40 | 1.53 |
| 3 | 2+ | 1.67 | 0.28 | 0.69 | 0.24 | 0.42 | 0.03 | 0.22 | 0.09 |
| .<br>.<br>. | .<br>.<br>. | .<br>.<br>. | .<br>.<br>. | .<br>.<br>. | .<br>.<br>. | .<br>.<br>. | .<br>.<br>. | .<br>.<br>. | .<br>.<br>. |
| **Total** | | 202.87 | 41.78 | 93.74 | 30.04 | 308.65 | 53.25 | 119.28 | 53.09 |

The origin-destination travel time table and the zone specification table present information that is used to generate the zone degree of belonging to each decision alternative with regard to target home zone 14. Table 9 shows part of this table. The total number of trips taken from the previous step is distributed across zones according to the probabilities of belonging to decision alternatives.

Table 9- A part of table of zone degree of belonging to decision alternatives for home zone 14

| Origin | Destination | CCSPR | FCSPR | CCLPR | FCLPR | CCSPU | FCSPU | CCLPU | FCLPU | Total |
|---|---|---|---|---|---|---|---|---|---|---|
| 14 | 26 | 0.44 | 0.01 | 0.05 | 0.01 | 0.48 | 0.01 | 0.00 | 0.00 | 5.64 |
| 14 | 27 | 0.43 | 0.01 | 0.05 | 0.01 | 0.38 | 0.01 | 0.10 | 0.01 | 6.21 |
| 14 | 28 | 0.49 | 0.01 | 0.00 | 0.00 | 0.49 | 0.01 | 0.00 | 0.00 | 5.31 |
| 14 | 137 | 0.00 | 0.00 | 0.00 | 0.48 | 0.01 | 0.03 | 0.01 | 0.47 | 5.70 |
| 14 | 138 | 0.00 | 0.00 | 0.00 | 0.48 | 0.01 | 0.04 | 0.01 | 0.46 | 5.70 |

The number of total work trips is distributed between different destination zones according to destination zones' degree of belonging to decision alternatives. The last column of the above table shows the final number of assigned work trips from zone 14 to three example destination zones (zones no. 26 to 28).

## 4- FDT MODEL VALIDATION

The model validation should show how model respond to changes in household and transportation characteristics (i.e. sensitivity analysis) as well as the model replicate travelers' destination and mode choices. Model results for traffic zone 14 presented in previous section display how model is sensitive to household characteristic. Table 7 illustrates the choice probability of mode and destination for work trip in zone 14 for household with 3 members and three different categories of household car ownership. As it is shown when the household number of car increase from 0 to 2+, choice probability of private mode (the summation of first four columns) is greater than transit (the summation of last four columns). Table 9 shows how zone degree of belonging to decision alternative changes when destination zone change to zones No. 137 and 138, Zone 137 and 138 are far away from CBD compare to zones 26 to 28, thus decision alternative includes zones far from CBD (decision alternatives include FC in their names) and long travel time by private mode (decision alternatives include LPR in their names) or public mode (decision alternatives include LPU in their names) has higher values. Therefore, these comparisons indicate that the decision alternatives are sensitive to household characteristic (car ownership) and transportation system change in different traffic zones.
The performance of the FDT model is evaluated in disaggregate and aggregation levels to show how model replicate travel demand. The choice probability for one of decision alternative (fifth decision alternative: "CCSPU") selects and real choice probabilities for travelers in train and test data set are compare with choice probabilities estimated using FDT and MNL. Table 10 shows the equations of trend lines passed real and estimated choice probabilities and R-square values. Accuracy of choice probability estimation for all trip purposes except work increased using FDT. However, the low R-square values while validation is performed in disaggregate level are expected. Destination and mode choices are usually compulsory for work trips; therefore destination and mode are less selectable.

Table 10- Comparison RealChoice Probablities with FDT and MNL Estimated Ones

| Purpose | Data | FDT | MNL |
|---|---|---|---|
| School | Train | $y = 0.0438x + 0.6043$<br>$R^2 = 0.3619$ | $y = 0.1961x + 0.5446$<br>$R^2 = 0.281$ |
| | Test | $y = 0.0424x + 0.604$<br>$R^2 = 0.3363$ | $y = 0.1886x + 0.5459$<br>$R^2 = 0.2621$ |
| Work | Train | $y = 0.0369x + 0.4013$<br>$R^2 = 0.1149$ | $y = 0.1098x + 0.3497$<br>$R^2 = 0.1263$ |
| | Test | $y = 0.0344x + 0.4019$<br>$R^2 = 0.0971$ | $y = 0.1049x + 0.3489$<br>$R^2 = 0.1175$ |
| Shopping | Train | $y = 0.0456x + 0.5507$<br>$R^2 = 0.3315$ | $y = 0.2336x + 0.4335$<br>$R^2 = 0.3051$ |
| | Test | $y = 0.0514x + 0.5455$<br>$R^2 = 0.3773$ | $y = 0.2704x + 0.3974$<br>$R^2 = 0.3222$ |
| Recreation | Train | $y = 0.0268x + 0.4751$<br>$R^2 = 0.2024$ | $y = 0.1633x + 0.5827$<br>$R^2 = 0.1768$ |
| | Test | $y = 0.0294x + 0.4739$<br>$R^2 = 0.2178$ | $y = 0.1658x + 0.5811$<br>$R^2 = 0.1338$ |
| Personal | Train | $y = 0.1337x + 0.3584$<br>$R^2 = 0.2252$ | $y = 0.1501x + 0.5711$<br>$R^2 = 0.1646$ |
| | Test | $y = 0.1171x + 0.3886$<br>$R^2 = 0.1974$ | $y = 0.126x + 0.6167$<br>$R^2 = 0.1611$ |

The validation in aggregate level compares FDT aggregate estimates with real travel demand. The aggregation method is used to provide aggregate estimates of trip production, attraction, distribution, and modal split for the Shiraz study area. The aggregate results are then analyzed and compared to actual travel demand. The FDT model's outputs are aggregated in regions that include a number of neighborhood traffic zones that are usually similar in terms of the influential variables used in FDT construction, zonal car ownership and zone distance to CBD. Thus, the aggregate trip generation, distribution, and modal split estimations for these traffic zones are approximately the same. Therefore, in the aggregate evaluation of the FDT model, the actual regional travel demand is compared with the FDT results aggregated according to region. The work and shopping trip purposes are selected to show the performance of the FDT model in figures for both compulsory and voluntary trips. The FDT performance for other trip purposes is presented in Table 11.

Trip production and attraction in 15 regions in Shiraz are compared in Figures 9 and 10. $R^2$ values show that the model performs very well.

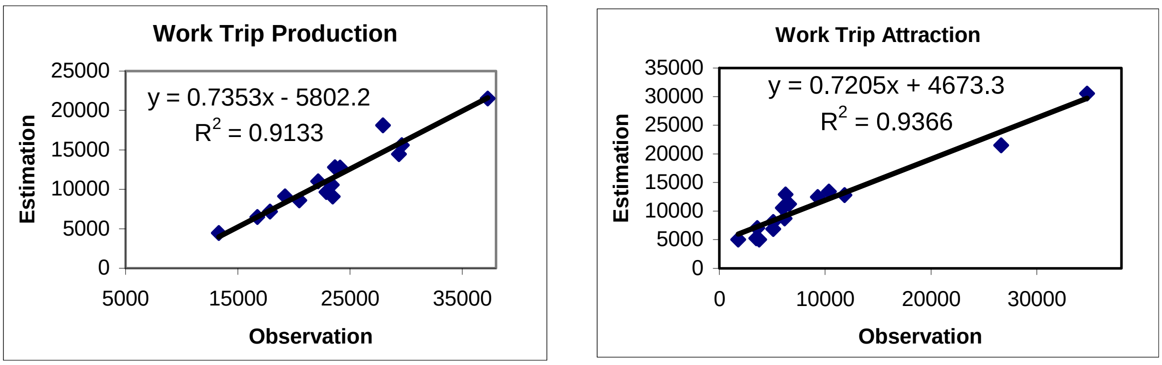
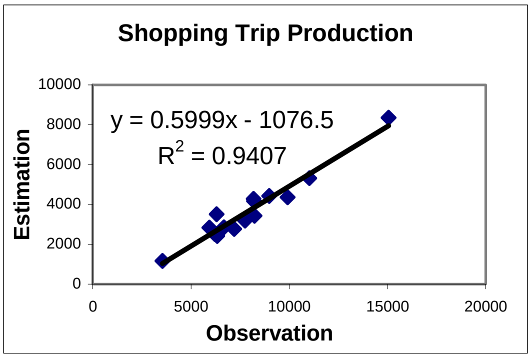



Figure 9- Results of FDT work trip generation comparison in regional scale with observation



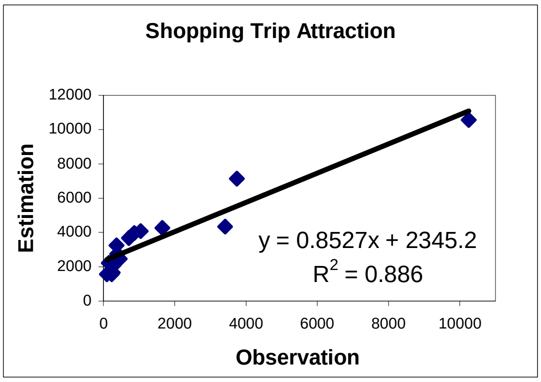




Figure 10- Results of FDT shopping trip generation comparison in regional scale with observation

Figures 11 and 12 compare the actual trip distributions for each regional origin-destination pair with the aggregated FDT estimations. The equations and $R^2$ values indicate that the FDT model performed adequately based on trip distribution estimations.

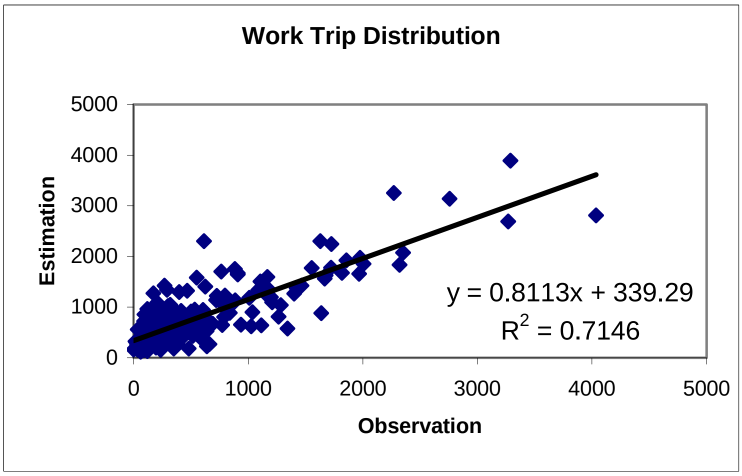



Figure 11- Estimated FDT work trip distribution compared with the observed distribution

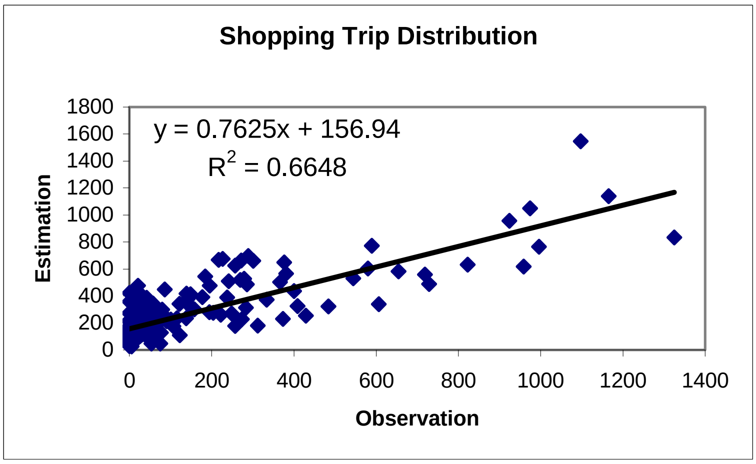



Figure 12- Estimated FDT shopping trip distribution compared with observed distribution

The precision of the FDT model in estimating mode choice should also be examined. Thus, the estimated number of trips made between regions with a private car for two purposes is compared with the actual number of trips in Figures 13 and 14. These figures show that the performance of the FDT model for mode choice is acceptable; The $R^2$ values and equations seem sufficient.

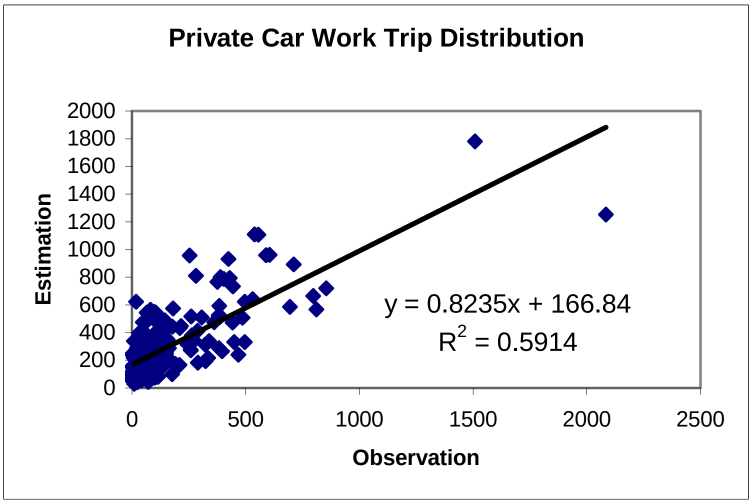



Figure 13- Estimated FDT private car work trip distribution compared with the observed distribution

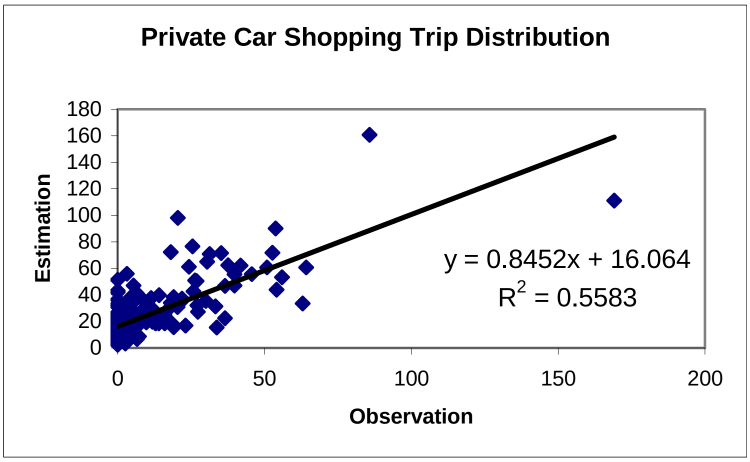



Figure 14- Estimated FDT private car shopping trip distribution compared with the observed distribution

Table 11- FDT estimation compare with the observation for three purposes

| Model \ Purpose | Recreation | Personal | School |
|---|---|---|---|
| Production | $y = 0.8867 \times x - 1031.4$ <br> $R^2 = 0.9172$ | $y = 0.8549 \times x - 399.17$ <br> $R^2 = 0.9465$ | $y = 0.7056 \times x - 3036.7$ <br> $R^2 = 0.7096$ |
| Attraction | $y = 0.9605 \times x - 1339.3$ <br> $R^2 = 0.9514$ | $y = 0.8411 \times x + 1309.9$ <br> $R^2 = 0.6355$ | $y = 0.8113 \times x + 3269.3$ <br> $R^2 = 0.7657$ |
| Distribution | $y = 0.9543 \times x + 136.88$ <br> $R^2 = 0.6152$ | $y = 0.8472 \times x + 87.055$ <br> $R^2 = 0.4748$ | $y = 0.8745 \times x + 334.74$ <br> $R^2 = 0.5071$ |
| Private Car Distribution | $y = 0.7451 \times x + 22.797$ <br> $R^2 = 0.3762$ | $y = 0.8569 \times x + 18.628$ <br> $R^2 = 0.3483$ | $y = 0.8156 \times x + 42.955$ <br> $R^2 = 0.3908$ |

## 5- CONCLUSION

This article develops and applies a fuzzy-probability disaggregated joint model for destination and mode choice in urban trips using a decision tree as a learning algorithm. Fuzzy variables used in a decision tree structure take travelers' individual perceptions into consideration. This model is called a fuzzy decision tree (FDT). Extracting choice probabilities from the decision tree's leaves using probability theory aims to address the uncertainty embedded in the randomness of traveler behavior. An aggregation procedure is suggested to provide aggregate estimates. This procedure defines input variables at the aggregate level.
The FDT construction and inference algorithms are designed to be compatible with the destination and mode choice problem. The construction algorithm is based on entropy minimization. Trip data from the Shiraz Comprehensive Transportation Study (SCTS) containing information on home-based trips in the morning peak hour for 9177 travelers is used for FDT construction and evaluation.
The performance of the FDT model in learning and generalization for aggregate estimates are evaluated. The FDT output of aggregation is compared with actual trip generation; distribution and modal split indicate that the model appropriately estimates travel demand. The tree structure model can be interpreted as a rule-based structure that increases the interpretability of the model as compared to traditional analytical models.

## REFERENCES

Avineri, Erel (2009). Incorporating Fuzzy Reference Points into Applications of Travel Choice Modeling. Book Chapter, Applications of Soft Computing, Springer Berlin / Heidelberg, Vol. 52, pp. 221-229.
Avineri, Erel (2009). Fuzzy Meaning of Reference-Based Perceptions in Travel Choice Modeling. TRB 88th Annual Meeting Compendium of Papers DVD, Accession Number: 01125546, Washington DC.
Ben-Akiva, M., and S.R. Lerman (1985). Discrete choice analysis: theory and application to travel demand. The MIT Press, Cambridge.
Breiman, L. J., H. Friedman, R. A. Olshen, and C. J. Stone (1984). Classification and Regression Trees. Wandsworth, Belmont.
Bezdek, J. C. (1981). Pattern Recognition with Fuzzy Objective Function Algorithms. Kluwer Academic Publishers, New York.
Dell 'Orco, Mauro and Shinya Kikuchi (2004). An Alternative Approach for Choice Models in Transportation: Use of Possibility Theory for Comparison of Utilities. Youguslav Journal of Operation Research, Vol. 14, Issue 1, pp. 1-17.
Janikow, C.Z. (1998). Fuzzy Decision Trees: Issues and Methods. IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, Vol. 28, Issue 1, pp. 1-14.
Mitchell, T. (1997). Machine Learning. McGraw-Hill, Singapore.
Ortuzar, J.D. and L.G. Willumsen, (2001). Modeling Transport. John Wiley & Sons, New York.
Quinlan, J. R. (1993). C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers, California.

Quinlan, J.R. (1987). Generating production rules from decision trees. Proceedings of the Tenth International Joint Conference on Artificial Intelligence, pp. 304-307.

Ross, T. J., J. M. Booker, and W. J. Parkinson (2002). Fuzzy Logic and Probability Applications: Bridging the Gap. Society for Industrial and Applied Mathematics, Philadelphia.

Teodorovic, D., and K. Vukadinovic (1998). Traffic control and transport planning, a fuzzy sets and neural networks approach. Kluwer Academic Publisher, Boston.

Teodrovic, D. (1999). Fuzzy Logic Systems for Transportation Engineering: the State of the Art. Transportation Research Part A, Vol. 33, pp. 337-364.

Thill, J., and A. Wheeler (2000). Tree induction of spatial choice behavior. Transportation Research Record, Vol. 1719, pp. 250-258.

Wets, G., K. Vanhoof, T.A. Arentze, and Timmermans (2000). Identifying decision structures underlying activity patterns: an exploration of data mining algorithms. Transportation Research Record, Vol. 1718, pp. 1-9.

Xie, Chi, Jinyang Lu, and Parkany Emily (2003). Work travel mode choice modeling with data mining: Decision trees and neural networks. Transportation Research Record, Vol. 1854, pp. 50-61.

Yamamoto, T., R. Kitamura, and J. Fuji (2002). Drivers' route choice behavior: analysis by data learning algorithms. Transportation Research Record, Vol. 1807, pp. 59-66.

Zadeh, L. (1995). Discussion: probability theory and fuzzy logic are complementary rather than competitive, Technometrics, Vol. 37, pp. 271-276.

Zadeh, L. (1996). Fuzzy logic = computing with words, IEEE Transaction on Fuzzy Systems, Vol. 4, pp. 103-111.

**Revision #1**

1-This paper would benefit from a more thorough review of state of the practice and state of the art/research in mode choice and destination choice.

2-The author only mentions MNL models, while there are many more models used in practice and have been presented in research.

3-The author should better position his work within current practice and research. While the methodology produces mode choice and destination choice decisions, the author should either elaborate on how the model is really a joint choice model (as defined in utility theory modeling) or change the wording used.

4- The authors show that the model replicated travel demand; however it is unclear how the model would respond to policy scenarios and changes in land-use and transportation characteristics. The paper would benefit from use of examples showing how the model responds to sensitivity testing and policy scenarios.

Authors' Response:

1- Three new references added to reference list and related explanations presented at fourth paragraph in first section "Introduction".
2- Paragraph 2 in first section "Introduction" explain different models used for choice models. The popular choice model is MNL model which often used in constructing choice models in practice. Thus, we use this model for comparisons.
3- Explanation added to third paragraph of "Introduction". Destination and mode choices are similar to trip distribution and modal split, two step of 4-step travel demand, respectively. In this paper, we model these two choices in on step and unique structure called fuzzy decision tree (FDT). Thus, we called the FDT a joint model. In literature, similar works presented the same wording "Joint model" in forecasting travel demand and choice modeling like:
   Chandra Bhat works http://www.ce.utexas.edu/prof/bhat/home.html
   Adler, T J and Ben-Akiva, M http://pubsindex.trb.org/view.aspx?id=47180
   Lerman, S R http://pubsindex.trb.org/view.aspx?id=53665
4- A paragraph added to section 4 to indicate how model is sensitive to household and transportation characteristics. Table 5 and related explanation in its previous paragraph also added to paper to explain how FDT is sensitive to household characteristic. Decision alternatives in decision tree are influenced by transportation characteristics, thus they change in regard to policy scenarios and transportation characteristics. Figure 4 shows how decision alternatives defined from transportation characteristics. They defined according to travel time between origin and destination for both private and transit system and distance of destination zone to CBD. Distance to CBD displayed traveler accessibility to transit systems and destination link traffic congestion.

**Revision #2**

This paper is not very poor (as suggested by the "D"-rating), but I did not think it was that good either. There are some issues with this paper that need to be addressed before the paper can be accepted:

1. Some of the fuzzyfications don't seem to make much sense, like zonal car ownership. A traveler knows whether or not he owns a car, right? So it is the analyst that has a fuzzy perception? This is not in line with the argument made in the introduction, where it is stated that travelers' uncertainty is best modeled by fuzziness, and analyst's uncertainty by means of probability.

2. The fuzzy model should have been compared with a basic MNL-model to see if the great increase in the fuzzy model's complexity pays of in terms of model fit and predictions.
3. The authors missed a great deal of recent applications of fuzzy set-theory in transportation / travel choice modeling, and fail to show how they add to the scholarly literature.
4. The model is introduced in a very condensed, abrupt fashion, which makes it difficult to understand it and see its merits.
5. Validation is performed at the aggregate level, so that high R-squares are to be expected. Again, no comparison with competing models is presented.
6. More details are needed concerning data collection and respondent characteristics.

Authors' Response:

1- Some sentences are added to second paragraph at sub section 3-1. Zonal car ownership means the average of car ownership in each zone that can be an indicator of household income level in each traffic zone. This value is different from number of cars owned by travelers. Household car ownership which is a categorical non-fuzzy variable in our paper indicates number of cars for each household (HC).
2- Table 10 and second paragraph of section 4 added to paper for comparing FDT with MNL model at disaggregate level.
3- We add three new references and the related explanations add to fourth paragraph at first section "Introduction".
4- We hope changes in revised version make proposed model clearer.
5- Table 10 and second paragraph of section 4 added to paper for comparing FDT with MNL model at disaggregate level.
6- More details added to first and second paragraph at section 3 "FDT Application in SHIRAZ".