A Fuzzy Decision Tree Model to Support the Task of Bus Reallocation in Public Transport Systems

Marcos Evandro Cintra[1], José Francisco Ferreira Ribeiro[2], Odacir Almeida Neves[1]

[1] Exact and Natural Sciences Dept. - Federal Territorial University of the Semi-Arid - UFERSA
{cintra, odacir}@ufersa.edu.br
[2] FEARP School of Business - University of Sao Paulo - USP
jffr@fearp.usp.br

Abstract.   The task of planning and managing urban public transport systems, such as bus networks, is a relevant problem with several approaches proposed in the literature. In this article, we focus on the specific task of reallocating buses from different lines of a previously planned system in case of restrictions regarding the breakdown of vehicles and/or absent drivers. For this purpose, we present a fuzzy decision tree that combines the graphical representation of decision trees, which are highly interpretable, with the mechanisms to handle linguistic attributes of the fuzzy systems. The model was induced using real data collected from a public bus system with 26 bus lines. The collected data includes 16 attributes related to characteristics of the bus lines and users. Such data basically present input values for the 16 attributes, as well as an expert decision on the reallocation of buses. The final induced models include reduced sets of key-attributes. The idea is to use the fuzzy decision trees to support human experts to make decisions regarding reallocating buses from lines in order to optimize the average time that users have to wait for a bus service. The models induced by the FuzzyDT algorithm were compared to RIPPER, J4.8, NaiveBayes, and a Multi-layer Perceptron. The FuzzyDT induced models presented superior results in terms of error rates and interpretability. The models use reduced sets of attributes and are highly interpretable.

Categories and Subject Descriptors: H.2 [**Database Management**]: Miscellaneous; H.3 [**Information Storage and Retrieval**]: Miscellaneous; I.7 [**Document and Text Processing**]: Miscellaneous

Keywords: Fuzzy decision trees, fuzzy systems, public transport systems

1. INTRODUCTION

The planning and management of urban public transport systems [Fan and Machemehl 2008; Han and S.; Kim 2011], such as bus networks [Ghatee and Hashemi 2008], is a relevant problem that can be decomposed in a sequence of tasks, including the network design, frequency setting, timetable development, as well as bus and drivers scheduling [Ceder and Wilson 1986], among others.

Regarding bus network planning, a strong restriction for the bus system is the number of available vehicles. In this sense, researchers usually focus on the scenario of distributing a previously defined set of vehicles into a given number of bus lines in order to define a bus network. Proposals for this specific task usually consider one single objective function [Oliveira et al. 2008], such as the optimization of the average time a traveller waits for a service. Other approaches work on the optimization of a set of conflicting objectives [Fonseca and Fleming 1993], usually by adopting methods that generate a Pareto front which allows to choose solutions from a set of non-dominated options. Clustering algorithms [Oliveira et al. 2009] have also been used for this specific task of bus network planning.

The planning and management of urban public transport systems also involves a series of practical problems that are usually tackled by human experts. One of these problems is the redistribution of

the buses of a bus network in case of mechanical problems in vehicles and/or absent drivers. For this specialized task, instead of replanning the whole bus network if no extra vehicles or drivers exist, a simple solution involves studying the impact of reallocating buses or drivers from non-affected lines to affected ones.

Decision trees are popular models in machine learning, especially for classification problems, due to the fact that they produce graphical models, as well as text rules, that are easily understandable for final users. Moreover, their induction process is usually fast, requiring low computational power. Fuzzy systems, on the other hand, provide mechanisms to handle imprecision and uncertainty in data, based on the fuzzy logic and fuzzy sets theories [Zadeh 1965; 1994]. The combination of fuzzy systems and decision trees [Ribeiro et al. 2013; Cintra et al. 2012; Chandra and Varghese 2009; Huang et al. 2008; Janikow 2004] has produced fuzzy decision tree models, which benefit from both techniques to provide simple, accurate, and highly interpretable models at low computational cost.

In this sense, we tackle the bus network reallocation problem using a fuzzy decision tree model. The idea is to provide the human planner a support system to evaluate possible options. This evaluation is focused on the impact of reallocating buses from lines based on the characteristics of the involved lines and the human experience, whose knowledge is stored in the form of rules in the decision trees.

The contribution of this work is to provide the human planner with a graphical model, a decision tree, to support decisions regarding the modification of a previously planned bus network. In other words, the decision tree can be used to decide the best bus line of a previously planed bus network to be modified in case a bus in the bus network breaks down or a driver is unable to work.

In this work, we use real data collected from a bus system including 26 bus lines in order to induce fuzzy decision trees to support the reallocation of buses. The fuzzy decision tree algorithm used to induce the models is named FuzzyDT [Cintra et al. 2012]. The induced decision trees contain reduced subsets of attributes, which are defined as linguistic attributes, increasing their interpretability.

Some of the reasons for us to adopt fuzzy decision trees are: i) the powerful models generated by decision trees; ii) the characteristics of the attributes defining the bus lines, which are all continuous, and, thus, can be defined in terms of fuzzy sets; iii) the superior accuracy and interpretability of the generated models in comparison to other classic machine learning algorithms. In fact, the models induced by the FuzzyDT algorithm had better results when compared to J4.8, an implementation of the C4.5 algorithm, as well as the NaiveBayes and Multi-layer Perceptron (MLP) algorithms.

The remaining of this article is organized as follows. Section 2 introduces the problem of bus reallocation and includes a bibliographic review on public transport planning. Section 3 presents the basic concepts of fuzzy classification systems. Section 4 discusses decision trees and presents the FuzzyDT algorithm. Section 5 details the process adopted to define the input attributes for the problem, the experiments and results. Section 6 presents the conclusions and future work.

2. CONTEXTUALIZATION AND BIBLIOGRAPHIC REVIEW

The task of planning public transport systems can be roughly divided into the following sequence of activities [Ceder and Wilson 1986; Ceder 2002]: (1) Network design; (2) Frequency setting; (3) Timetable development; (4) Bus scheduling; (5) Drivers scheduling.

Notice that our work focuses specifically on the task of rearranging the distribution of buses in case of bus breakdowns and problems with drivers absences. In fact, our proposal considers that a previous bus planning has already been done, *i.e.*, activities 1 to 5 were carried out and the whole bus system is planned. Once the transport system is designed, considering its optimization in regards to the available resources and the expectancies of the travellers, its maintenance requires constant updates and adjusts in order to keep the system running efficiently. The maintenance of the bus lines is the focus of our proposal.

Regarding the activities for public transport planning, in the most general case, the definition of the routes (activity 1) and the frequency (activity 2) of services has to be done taking the interests of travellers into account, as well as the interests and resources of the companies that provide such services.

There are several different proposals in the literature for the problem of planning public transport systems [Desaulniers and Hickman 2007; Vuchic 2005]. The tabu meta-heuristic search and a local search algorithm are used in the work of Pacheco et al. [2001] for activities 1 and 4, including two decision levels: i) definition of the routes of the bus lines, ii) designation of the buses to the lines. Genetic algorithms are used for this task in the work of Goldberg [1989] aiming at improving the existing service by redistributing and reducing the number of transfers, as well as the travel time of each user. Genetic algorithms are also used by Kidwai et al. [2005] to bus scheduling, activities 2 and 3, by performing two steps: in the first one, the minimum frequency of buses required on each route, with the guarantee of load feasibility, is determined by considering each route individually; in the second step, the fleet of buses is taken as upper bound and fleet size is again minimized by considering all routes together.

Yu et al. [2011] tackle activity 3 by combining tabu search and genetic algorithms. The proposal defines a solution by compromising the quality of the services to the users and the total cost of the services, *i.e.*, the aim is to maximize the quality of the services and minimize their operational cost.

Regarding other related problems, the Eiligen algorithm is used by Surapholchai et al. [2008] for solving the multiple-depot vehicle scheduling problem, activity 4. The Eiligen algorithm is computationally simple and has the advantage of improving the solution quality as the number of depots grows. Ghatee and Hashemi [2008] and Kumar and Kaur [2011], on the other hand, employ fuzzy logic to solve the problem of minimum cost flow for transport problems, related to activity 4. Tsubouchi [2009] work with the on-demand bus scheduling problem, regarding activities 1 and 4, by executing two main algorithms: a vehicle-choosing algorithm and a routing algorithm. Both algorithms in [Tsubouchi 2009] are based on expert heuristics.

Although it is possible to find several proposals for public transport system planning, the maintenance of the system, specifically the task of dealing with rescheduling buses of a previously defined bus system, is usually done by a human operator and has no automatic support systems available in the literature.

In practice, once a problem in a previously planned bus network is detected, the transport system manager can just ignore it, when the impact on the system can be considered tolerable. When the impact on the whole system is considerable, the manager has to reallocate buses from the line in case no extra buses or drivers exist. The decision of which bus line will be least affected by a reallocation is taken, by a human operator, with the analysis of all existing information regarding the lines.

Fan and Machemehl [2008] use the tabu search to solve a non linear model and redistribute urban buses, presenting a sensitivity analysis showing that the tabu search obtained better results than the genetic algorithm approach. Three computational programs were developed by Han and S.; Kim [2011] based on simulated annealing, tabu search, and genetic algorithms in order to solve the redistribution of urban buses problem, presenting a comparison of the obtained results.

Our proposal, the use of fuzzy decision tree systems, differs from other existing proposals in basically three aspects:

(1) The model is generated once and can be used when it is necessary to evaluate the bus rescheduling;
(2) Since the model is generalized based on existing bus system information, it is not necessary to induce different models for problems in different bus lines;
(3) The model is graphical and, thus, highly interpretable by non-experts in computational models;
(4) The model can be printed, thus, it does not need a computer to be used.

3. FUZZY CLASSIFICATION SYSTEMS

Classification is a relevant task of machine learning that can be applied to pattern recognition, decision making, and data mining, among others. The classification task can be roughly described as: given a set of objects $E = \{e_1, e_2, ..., e_n\}$, also named *examples or cases*, which are described by m features, also named variables or attributes, assign a class c_i from a set of classes $C = \{C_1, C_2, ..., C_j\}$ to an object e_p, $e_p = (a_{p_1}, a_{p_2}, ..., a_{p_m})$.

Fuzzy classification systems are rule-based fuzzy systems that require the granulation of the features domain by means of fuzzy sets and partitions. The linguistic attributes in the antecedent part of the rules represent features, or attributes, and the consequent part represents a class. A typical fuzzy classification rule can be expressed by

$$R_k : \textbf{IF } X_1 \text{ is } A_{1l_1} \textbf{ AND } ... \textbf{ AND } X_m \text{ is } A_{ml_m} \textbf{ THEN } Class = C_i$$

where R_k is the rule identifier, $X_1, ..., X_m$ are the features of the example considered in the problem (represented by linguistic attributes), $A_{1l_1}, ..., A_{ml_m}$ are the linguistic values used to represent the feature values, and $C_i \in C$ is the class. The inference mechanism compares the example to the rules in the fuzzy rule base in order to assign a class to the example.

The classic and general fuzzy reasoning methods [Cordon et al. 1999] are widely used in the literature. Given a set of fuzzy rules (fuzzy rule base) and an input example, the classic fuzzy reasoning method classifies this input example using the class of the rule with maximum compatibility to the input example, while the general fuzzy reasoning method calculates the sum of compatibility degrees for each class and uses the class with highest sum to classify the input example. The classic fuzzy reasoning method is also known and the *best rule method*, while the general fuzzy reasoning method is also known as the *best class method*.

4. FUZZYDT

As previously mentioned, decision trees provide popular and powerful models for machine learning which are easily understandable and intuitive. Decision trees also require low computational power and usually produce competitive models that can be expressed graphically or as a set of rules.

Another important aspect of decision trees is the fact that their induction process selects only the relevant attributes for the definition of the final model. Thus, the process of inducing the decision tree model performs an embedded attribute selection process, which simplifies the final model, improving its interpretability.

In this work, we adopt the FUZZYDT [Cintra et al. 2012] algorithm to generate the fuzzy decision trees in order to support the task of bus rescheduling. FUZZYDT uses the same measures of the classic C4.5 [Quinlan 1993] algorithm, one of the most relevant and well-known decision tree algorithms, to decide on the importance of the features. Thus, FUZZYDT uses the information gain and entropy measures to sequentially select the features to induce the models, which can be numerical and/or categorical. The entropy of a set S containing k possible classes is defined as [Shannon 1948]:

$$E(S) = - \sum_{j=1}^{k} \frac{freq(C_j, S)}{|S|} \cdot log_2 \left(\frac{freq(C_j, S)}{|S|} \right)$$

where $freq(C_j, S)$ represents the number of examples in S that belongs to class C_J and $|S|$ is the number of examples in S. The entropy indicates the average amount of information necessary to classify an example in S.

The information gain (or entropy reduction) of an attribute At_i, *i.e.*, how much information is gained by splitting S using the values of At_i, can be defined as:
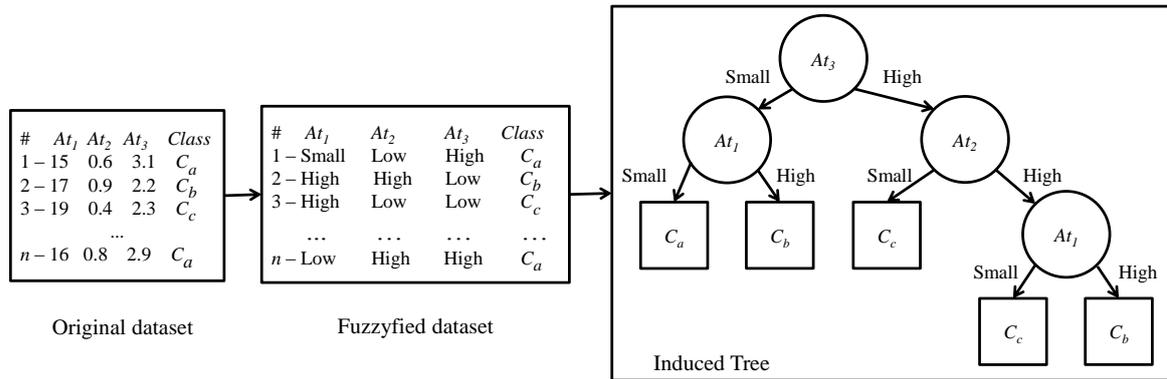
$$IG(S|At_i) = E(S) - E(S|At_i)$$

Fig. 1.   The FuzzyDT algorithm - a toy example.

FuzzyDT recursively creates branches corresponding to the values of the selected features until a class is assigned as a terminal node. Each branch of the tree can be seen as a rule, whose conditions are formed by their attributes and respective tests.

In order to avoid overfitting, FuzzyDT adopts the same strategy of C4.5 of applying a post-pruning process. This way, the pruning process takes place after the tree is completely induced. The pruning process of FuzzyDT basically assesses the error rates of the tree and its components directly on the set of training examples [Cintra et al. 2012].

The post-pruning method implemented in FuzzyDT replaces sub-trees with leaf nodes. The class assigned to a leaf is the most frequent one found in the examples of the training set covered by that leaf. This pruning method analyses the error rate of the tree using just the training examples with which the tree is built. The basic idea is to estimate the real error of a sub-tree, which, in fact, cannot be determined using only the examples of the training set. If the estimated real error is smaller than the apparent error, *i.e.*, the error calculated using the set of training examples, the sub-tree is pruned.

The main steps of the FuzzyDT algorithm to induce a fuzzy decision tree are listed next.

(1) Define the fuzzy data base, *i.e.*, the fuzzy granulation for the domains of the continuous features;

(2) Replace the continuous attributes of the training set using the linguistic labels of the fuzzy sets with highest compatibility with the input values;

(3) Calculate the entropy and information gain of each feature to split the training set and define the test nodes of the tree until all features are used or all training examples are classified;

(4) Apply a post-pruning process, similarly to C4.5, using 25% confidence limits as default.

As the fuzzyfication of the training data is done before the induction of the tree, the third step of FuzzyDT corresponds to the same step of the classic decision tree algorithm.

Figure 1 illustrates the process of data fuzzyfication and tree induction for a toy dataset with n examples, 3 attributes (At_1, At_2, and At_3), and 3 classes (C_a, C_b, and C_c). The first block of Figure 1 illustrates a dataset with n examples, three attributes (At_1, At_2, and At_3) and a class attribute. The fuzzyfied version of this dataset is presented in the second block. This fuzzyfied set of examples is used to induce the final fuzzy decision tree, which is illustrated in the last block of Figure 1.

Next Section presents the experiments, as well as the attributes and data definition procedure and induced fuzzy decision trees for the task of bus rescheduling.

Table I.   General characteristics of the involved attributes.

	attribute	Min.	Max.	Avg.
1	Length (m)	7,620.00	30,890.00	16,563.14
2	Number of vehicles	2.00	14.00	6.22
3	Buses per Km	81.60	524.50	198.73
4	Interval between buses (seconds)	318.00	1,680.00	846.39
5	Rotation time (seconds)	2,832.00	7,388.00	4,594.51
6	Average speed (km/h)	9.70	18.00	12.79
7	Capacity	90.00	135.00	95.67
8	Number of travellers	235.00	2,706.00	1,137.73
9	Travellers per km	690.00	13,173.00	4,186.63
10	Average occupation rate	0.00	0.64	0.22
11	Maximum occupation rate	0.19	0.68	0.41
12	Commuting travellers	62.00	861.00	339.80
13	Average length of a route (km)	66.90	4,889.00	3,333.70
14	Average time (seconds)	497.00	1,121.00	864.51
15	Buses per km per travellers	0.10	0.39	0.21
16	Load	125.00	967.00	468.71
17	Average waiting time (seconds)	300.00	730.00	520.25

5. EXPERIMENTS

As previously mentioned, we adopted fuzzy decision trees due to the powerful models generated by decision trees. Also, as the involved attributes are all continuous, they can be properly defined in terms of fuzzy sets, generating linguistic attributes.

The generated decision tree is based on 52 real examples collected for 26 bus lines in the city of Grenoble, France. Table I presents the names of the 17 attributes included in the data, as well as their minimum (Min.), maximum (Max.), and average (Avg.) values.

Most of the attributes are self-explanatory. attribute *Load* refers to the number of travellers in the busiest section of the line. The *Average Time* attribute refers to the time taken to travel from the first to the last stop of a line. The *Average Waiting Time* attribute refers to the average time a traveller has to wait for a bus to arrive at any bus stop of the given bus line.

Each original attribute was used to define a linguistic attribute, according to the fuzzy logic theory. The linguistic attributes were defined by triangular shaped fuzzy sets, evenly distributed in the domains, according to the equalized universe method [Chen and Wang 1999]. In fact, different decision trees were induced using 3, 4, and 5 fuzzy sets defining the input attributes. A fourth set of experiments were performed using the FuzzyDBD method[Cintra et al. 2011] to define the number of fuzzy sets for each attribute as well. FuzzyDBD uses a function to estimate the number of fuzzy sets for each attribute individually. The function used in this work is based on the Wang & Mendel method [Wang 2003]. The idea here is to induce different models in order to evaluate their interpretability and accuracy power.

Specifically for the output attribute of the fuzzy decision tree, the *Average Waiting Time*, we adopted 5 triangular fuzzy sets, also evenly distributed in its domain, for all experiments. The linguistic values were chosen in order to reflect the adequacy of the waiting time of a passenger for a bus service. The set of linguistic values is composed of *Excellent, Good, Regular, Long*, and *Very Long*. Such linguistic values properly define the *Average Waiting Time* and can be easily understood and interpreted by an expert in bus network planning.

Table II shows the number of fuzzy sets used to define each attribute of the domain. The FuzzyDBD method defined from 2 to 8 fuzzy sets to each of the attributes. This variation is duo to the fact that each attribute is evaluated individually by the FuzzyDBD method.

Table II.   Number of fuzzy sets defining each attribute.

Attribute	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Class
FuzzyDT 3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	5
FuzzyDT 4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	5
FuzzyDT 5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
FuzzyDBD	6	6	6	5	8	6	3	4	8	2	8	7	5	7	6	8	5

Table III.   Error rates and number of rules of each induced model.

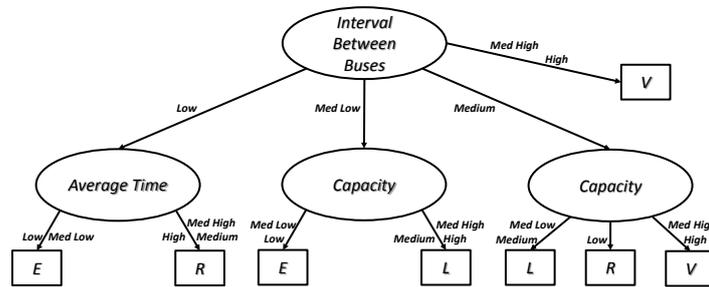|	FuzzyDT 3	FuzzyDT 4	FuzzyDT 5	FuzzyDBD	Ripper	C4.5(J48)	NaiveBayes	MLP
Error	5.77	5.77	1.92	3.84	5.66	12.5	22.64	5.66
Rules	14	11	8	9	5	6	–	–
Attributes	8	5	3	3	4	4	–	–

Fig. 2.   Fuzzy decision tree induced using 5 fuzzy sets for each attribute.

The induced fuzzy decision trees produced by FuzzyDT were compared to the models generated by RIPPER [Cohen 1995], J4.8, an implementation of the C4.5 [Quinlan 1993] algorithm, as well as the NaiveBayes [Cheeseman and Stutz 1996] and Multi-layer Perceptron (MLP) [Haykin 1998] algorithms, all available at WEKA [Witten et al. 2011]. All algorithms used default parameters. Table III presents the error rates, number of rules, as well as the number of different attributes presented for each model (smallest values are light-grey shaded).

In general, the FuzzyDT induced models presented better results in terms of error rates and interpretability than the classic ones. The only exception is the number of rules generated by RIPPER, which is smaller than for the other models. Nevertheless, the rules generated by RIPPER present multiple splits for the same attribute. For instance, attribute INTERVAL BETWEEN BUSES has 5 different splitting points: $\leq 360, \geq 840, \leq 900, \leq 600$, and ≤ 800, compromising the interpretation of the model as a whole.

Regarding the fuzzy models, the fuzzy decision tree generated using 5 fuzzy sets presented the smallest error (1.92, as well as the smallest number of rules (8) and different attributes included in the model (3). The one induced using FuzzyDBD presents comparable results. Since the models induced using 3 and 4 fuzzy sets use more attributes, it is expected that they present more rules.

Due to space restrictions, we present and discuss the results of the fuzzy decision trees induced using 5 fuzzy sets and the one generated using the FuzzyDBD method to define the fuzzy data bases. Figure 2 shows the induced fuzzy decision tree using 5 fuzzy sets defining each input attribute. Notice that E stands for Excellent, G for Good, R for Regular, L for long, and V for Very Long.

As previously stated, decision trees can be seen as a set of rules. The fuzzy decision tree induced using 5 fuzzy sets for each attribute, Figure 2, can be seen as the following set of rules.

1 - **If** INTERVAL BETWEEN BUSES **is** *Medium High or High* **then** *Average Waiting Time* **is** *Very High*

2 - **If** INTERVAL BETWEEN BUSES **is** *Low* **&** AVERAGE TIME **is** *Low or Medium Low* **then** *Average*
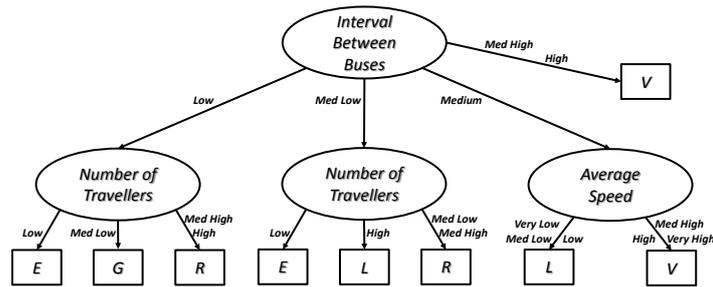
Fig. 3. Fuzzy decision tree induced using the FuzzyDBD method to define the number of fuzzy sets for each attribute.

Waiting Time **is** *Excellent*

3 - **If** Interval Between Buses **is** *Medium Low* & Average Time **is** *Medium, Medium High, or High* **then** *Average Waiting Time* **is** *Good*

4 - **If** Interval Between Buses **is** *Medium Low* & Capacity **is** *Low or Medium Low* **then** *Average Waiting Time* **is** *Good*

5 - **If** Interval Between Buses **is** *Medium Low* & Capacity **is** *Medium, Medium High, or High* **then** *Average Waiting Time* **is** *Regular*

6 - **If** Interval Between Buses **is** *Medium* & Capacity **is** *Low* **then** *Average Waiting Time* **is** *Regular*

7 - **If** Interval Between Buses **is** *Medium* & Capacity **is** *Medium Low or Low* **then** *Average Waiting Time* **is** *Long*

8 - **If** Interval Between Buses **is** *Medium* & Capacity **is** *Medium High or High* **then** *Average Waiting Time* **is** *Very Long*

Figure 3 shows the fuzzy decision tree induced using FuzzyDBD. As previously stated, *E* stands for Excellent, *G* for Good, *R* for Regular, *L* for long, and *V* for Very Long. Notice that the final models present some desirable characteristics regarding the attributes that are considered relevant for the problem as well as interesting characteristics regarding the sequence of importance for the attributes. For instance:

—Although 16 input attributes are available, only 3 are used in each model due to the embedded feature selection performed by FuzzyDT;
—As only 3 attributes are used in each fuzzy decision tree, we have simple and highly interpretable models. They also present a straightforward and fast inference;
—The attributes used in the induced trees, *Interval Between Buses*, *Average Speed*, *Number of travellers*, *Capacity*, and *Average time*, *i.e.*, the relevant attributes for the fuzzy decision trees, are also considered as the deciding factors for the reallocation of vehicles by experts.

Although the previous set of rules can be used to evaluate the reallocation of buses, its graphical representation as a decision tree (Figure 2 and Figure 3) is easier and more intuitive to make inferences about a real situation and, thus, take decisions. Moreover, since the original attributes were transformed into linguistic attributes, defined by linguistic terms, the process of using the decision tree to decide on the reallocation of buses is highly simplified when compared to continuous attributes. For example, instead of having to work with precise values, such as Capacity = 40, expert users can use their own knowledge of the capacity of the buses to infer what is low, medium, or high capacities.

The induced fuzzy decision trees can be easily applied as a support system to infer the resulting adequacy of changes in a given line. As the inference process is straightforward and fast, several

changes and lines can be analysed in a short time. For example, in case a bus system with 50 lines has a vehicle with problems, instead of analysing all 49 remaining lines, the user can compare the resulting average waiting time, according to the fuzzy decision trees, of a few ones he or she thinks are the most suitable to be reallocated. It is also important to notice that the user has to evaluate only 3 different attributes for the models using 5 fuzzy sets and FuzzyDBD, instead of considering all 16 attributes.

The model was analysed by experts in urban transport who concluded it is coherent to their knowledge of the problem and can be adopted due to its ease of use. The positive feedback motivates us to propose the use of the model in real situations and carry out further experiments and comparisons.

6.   CONCLUSIONS AND FUTURE WORK

The task of planning and managing urban public transport systems, such as bus networks, is a relevant problem. This work focus specifically on the task of reallocating buses from a previously planned system. This task is necessary in case of restrictions regarding the number of extra vehicles and drivers when vehicle breakdowns occur, or in case of absent drivers.

Decision trees are popular models in machine learning due to the fact that they produce graphical representations, as well as text rules, easily understandable by final users. Moreover, their induction process is usually fast, requiring low computational power.

Fuzzy systems, on the other hand, provide mechanisms to handle imprecision and uncertainty in data, based on the fuzzy logic and fuzzy sets theory. The combination of fuzzy systems and decision trees has produced a considerable number of fuzzy decision tree models, which benefit from both techniques to provide simple, accurate, and highly interpretable models at low computational costs.

In this article, we present fuzzy decision tree models to support the reallocation of buses, specifically in case of buses breakdowns or absence of drivers. The models were induced using real data collected for a bus system with 26 bus lines. The 16 attributes involved in the induction process are related to a series of characteristics of the bus lines, such as their length and interval between buses, as well as characteristics of the travellers and the usage of the bus services, such as the average number of travellers, average number of commuting travellers, among others.

FuzzyDT was used to induce fuzzy decision trees. As the induced models use linguistic attributes, instead of continuous ones, the models become more intuitive and interpretable to human experts.

The models include reduced sets of attributes (from 3 to 8 of the original 16 available attributes). The attributes used in the model include the key ones for the evaluation of bus lines. The final fuzzy decision trees present a reduced set of conditions, which make inferences quite straightforward.

The FuzzyDT models presented higher accuracy and interpretability than classic models (RIPPER, J4.5, NaiveBayes, and MLP) and also received positive feedback from experts in urban transport.

As future work, we intend to collect data from other bus systems to generate fuzzy decision trees for such systems, in order to further evaluate the proposal. We intend to compare the proposed model to other interpretable models. We also intend to work on the definition of the number of fuzzy sets for each of the input attributes.

REFERENCES

Ceder, A. Urban Transit Scheduling: framework, review and examples. *Journal of Urban Planning and Development* vol. 128, pp. 225–244, 2002.

Ceder, A. and Wilson, N. H. M. Bus Network Design. *Transportation Research* vol. B, pp. 331–344, 1986.

Chandra, B. and Varghese, P. P. Fuzzifying Gini Index Based Decision Trees. *Expert Systems with Applications* 36 (4): 8549–8559, 2009.

Cheeseman, P. and Stutz, J. Advances in Knowledge Discovery and Data Mining. In *Advances in Knowledge Discovery and Data Mining*, U. M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, and R. Uthurusamy (Eds.). American Association for Artificial Intelligence, Menlo Park, CA, USA, Bayesian Classification (AutoClass): Theory and Results, pp. 153–180, 1996.

Chen, M. S. and Wang, S. W. Fuzzy Clustering Analysis for Optimizing Membership Functions. *Fuzzy Sets and Systems* vol. 103, pp. 239–254, 1999.

Cintra, M. E., Monard, M. C., and Camargo, H. A. FuzzyDT - A Fuzzy Decision Tree Algorithm Based on C4.5. In *Brazilian Congress on Fuzzy Systems*. pp. 199–211, 2012.

Cintra, M. E., Monard, M. C., Cherman, E. A., and Camargo, H. A. On the Estimation of the Number of Fuzzy Sets for Fuzzy Rule-Based Classification Systems. In *International Conference on Hybrid Intelligent Systems*. pp. 211–216, 2011. (2011).

Cohen, W. W. Fast Effective Rule Induction. In *International Conference on Machine Learning*. pp. 115–123, 1995.

Cordon, O., del Jesus, M. J., and Herrera, F. A Proposal on Reasoning Methods in Fuzzy Rule-Based Classification Systems. *International Journal of Approximate Reasoning* 20 (1): 21–45, 1999.

Desaulniers, G. and Hickman, M. D. Public Transit. In *Transportation*, C. Barnhart and G. Laporte (Eds.). Handbooks in Operations Research and Management Science, vol. 14. Elsevier, 2, pp. 69 – 127, 2007.

Fan, W. and Machemehl, R. B. A Tabu Search Based Heuristic Method for the Transit Route Network Design Problem. *Lecture Notes in Economics and Mathematical Systems* vol. 4, pp. 387–408, 2008.

Fonseca, C. M. and Fleming, P. J. Genetic Algorithms for Multiobjective Optimization: Formulation, Discussion and Generalization. In *International Conference on Genetic Algorithms*. pp. 416–423, 1993.

Ghatee, M. and Hashemi, S. M. Generalized Minimal Cost Flow Problem in Fuzzy Nature: an application in the bus network planning problem. *Applied Mathematical Modelling* 32 (12): 2490–2508, 2008.

Goldberg, D. E. *Genetic Algorithms in Search, Optimization and Machine Learning*. Addilson-Wesley, 1989.

Han, J., L. and S.; Kim, J. Meta-Heuristic Algorithms for a Transit Route Design. *Advanced OR and AI Methods in Transportation* vol. 1, pp. 686–691, 2011.

Haykin, S. *Neural Networks: A Comprehensive Foundation*. Macmillan College Publishing Company, 1998.

Huang, Z., Gedeon, T., and Nikravesh, M. Pattern Trees Induction: a new machine learning method. In *IEEE Transactions on Fuzzy Systems*. Vol. 16. pp. 958–970, 2008.

Janikow, C. Z. FiD 4.1: an overview. In *North American Fuzzy Information Processing Society*. pp. 877–881, 2004.

Kidwai, F. A., Deb, K., Marwah, B. R., and Karim, M. R. A Genetic Algorithm Based Bus Scheduling Model for Transit Network. In *Eastern Asia Society for Transportation Studies*. Vol. 5. pp. 477–489, 2005.

Kumar, A. and Kaur, A. Methods for Solving Unbalanced Fuzzy Transportation Problems. *Operational Research* vol. 1, pp. 81–99, 2011.

Oliveira, S., Ribeiro, J., and Seok, S. A Comparative Study of Similarity Measures for Manufacturing Cell Formation. vol. 27, pp. 19–25, 2008.

Oliveira, S., Ribeiro, J. F. F., and Seok, S. C. A Spectral Clustering Algorithm for Manufacturing Cell Formation. *Computers and Industrial Engineering* vol. 57, pp. 1008–1014, 2009.

Pacheco, J., Alvarez, A., Casado, S., and Gonzales-Velarde, J. L. A Tabu Search Approach to an Urban Transport Problem in Northern Spain. *Computers & Operations Research* 36 (6): 967–979, 2001.

Quinlan, J. R. *C4.5: programs for machine learning*. Morgan Kaufmann Series in Machine Learning. Morgan Kaufmann, 1993.

Ribeiro, M. V., Cintra, M. E., and Camargo, H. A. A Comparative Analysis of Pruning Strategies for Fuzzy Decision Trees. In *Joint IFSA World Congress NAFIPS Annual Meeting*. pp. 709–714, 2013.

Shannon, C. E. A Mathematical Theory of Communication. *The Bell System Technical Journal* vol. 27, pp. 379–423, 1948.

Surapholchai, C., Reinelt, G., and Bock, H. G. Solving City Bus Scheduling Problems in Bangkok by Eligen-Algorithm. *Modeling, Simulation and Optimization of Complex Processes*, 2008.

Tsubouchi, K. Scheduling Algorithm for On-Demand Bus System. *Information Technology: New Generations*, 2009.

Vuchic, V. *Urban Transit: Operations, Planning, and Economics*. Wiley, 2005.

Wang, L. The WM Method Completed: a flexible fuzzy system approach to data mining. *IEEE Transactions on Fuzzy Systems* vol. 11, pp. 768–782, 2003.

Witten, I. H., Frank, E., and Hall, M. A. *Data Mining: practical machine learning tools and techniques*. Morgan Kaufmann, 2011.

Yu, B., Yang, Z., Sun, X., Yao, B., Qingscheng-Zeng, Q., and Jeppesen, E. Parallel Genetic Algorithm in Bus Route Headway Optimization. *Applied Soft Computing* 11 (8): 5081–5091, 2011.

Zadeh, L. A. Fuzzy Sets. *Information and Control* vol. 8, pp. 338–353, 1965.

Zadeh, L. A. Fuzzy Logic, Neural Networks, and Soft Computing. *Communications of the ACM* 37 (3): 77–84, 1994.