

Dear sirs,

Our manuscript entitled “*A fuzzy decision tree model to support the task of bus reallocation in public transport systems*” has been submitted to the JIDM-KDMiLe special issue and it was returned to us with a detailed revision from 3 reviewers.

We deeply appreciate the very insightful and positive reviews we received. All suggestions and modifications were taken into consideration to improve the manuscript.

Please, find next our replies to the reviewers’ suggestions.

Best wishes,

Marcos E. Cintra, José F. F. Ribeiro, and Odacir Neves.

Reviewer A

Section 1: It is not clearly stated the contribution of your work in the Public Transportation (PT) Planning area neither the problem you are facing. A definition of such subject and its tasks (namely, of the one in which your work is focused on) is crucial to ease the comprehension of a reader without any background on PT planning. You do this in the beginning of the second section...but it should be located some lines above (even so, such definition is not sufficiently clear).

Authors: the contribution of the work has been clearly stated in the introduction.

* Are you proposing a technique to PT planning or to improve an existing PT plan (as you enunciate in the sixth paragraph of the second section)? This should be clearly stated in the introduction and also along the entire paper.

Authors: the proposal has been added in the Introduction.

Section 2:

* The section is very poorly described and organized. It needs to be completely re-written to avoid such a naive approach to the state-of-art. This section should be clear and concise. It must enunciate not only the research scope of your work but also what distinguishes it from the existing ones.

* The authors do start it well by describing the five key Planning tasks. However, they are re-grouped some lines below as "the Planning problem" - which has a much more vast state-of-art than the one referred in the paper.

Then, it is impossible to distinguish between what's iv) bus scheduling and i) network design. Some examples. Finally, I do not understand how the enumerated techniques (i.e. GA, meta-heuristics) are applied for the problems solved on such references.

I can provide you a suggestion to reorganize the section:

1) define planning and each one of its tasks clearly.

2) you may refer some works of interest for each one of the PT planning tasks. It is important to distinguish between works on I) Planning and works on II) improving the existing Planning (which is your case).

3) then, you exhaustively review the existing works on the task you are addressing.

4) Finally, you state what distinguishes your work from those.

* The authors use as basis a reference ([Ceder and Wilson 1986]): with almost 30(!) years which was impactful but too old for the topic regarded in the paper.

I could suggest other earlier references that could be useful (some from the same author also):

* A. Ceder, "Urban transit scheduling: framework, review and examples," *Journal of Urban Planning and Development*, vol. 128, no. 4, pp. 225-244, 2002.

* @INCOLLECTION{Desaulniers07,

author = {Guy Desaulniers and Mark D. Hickman},

title = {Chapter 2 Public Transit },

booktitle = {Transportation},

publisher = {Elsevier},

year = {2007},

editor = {Cynthia Barnhart and Gilbert Laporte},

volume = {14},

series = {Handbooks in Operations Research and Management Science },

pages = {69 - 127},

issn = {0927-0507} }

* @BOOK{Vuchic2005a,

title = {Urban Transit: Operations, Planning, and Economics},

publisher = {Wiley},

year = {2005},

author = {Vuchic, V.} }

* Another example...read these two sentences together:

"The task of planning public transport systems can be roughly divided into the following sequence of activities [Ceder and Wilson 1986]: i) network design; ii)

Frequency setting; iii) timetable development; iv) bus scheduling; v) drivers scheduling."

"Genetic algorithms are also used in [Kidwai et al. 2005] to bus scheduling 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."

Are we talking about ii) frequency setting, iv) bus scheduling or both?

***Authors:** We used all the suggestions regarding this section to reorganize it. We also had several paragraphs rewritten. The suggestions were adopted and the additional references included. Thank you for these references. They were really helpful.*

* Finally the title: "BUS NETWORKS IN PUBLIC TRANSPORT SYSTEMS".

It is fuzzy just for itself. Will the authors discuss vehicular networks? No, you are discussing PT planning. Namely, the state of art on improving PT planning.

* You compare your method against RIPPER...what is understandable because both induce decision trees. Then you compare it against C4.5, NB and MLP...which is ok. But...how about the state of art?

I am not convinced why your method is better than the existing state of art in bus rescheduling. In fact, you do not claim it so anywhere in your paper. So...which is your work's contribution?

A comparison of your method against some of the existing on the state of art for this topic (i.e. GAs like it is proposed by (Kidwai et. al, 2005)) must be part of this paper.

***Authors:** the differences between our proposal and existing ones are now clearly stated in Section 2. Please, consider the fact that rather than comparing proposals by which one is best in regards to some parameter, we propose a different approach for the task of bus systems maintenance and state that its advantages are related to its ease of use and interpretability, not exactly on precision, which is quite difficult to define in such a dynamic environment where decisions have to be taken in real time.*

Section 5: At some point, you say that RIPPER, C4.5, NB and MLP use default parameters. However, you explore different parameters for your own algorithm.

What I see is that both FuzzyDT and RIPPER work well...but what if I give them a "fair" competition? I mean... what if I explore different parameters for RIPPER?

***Authors:** we understand your concern. However, we are more knowledgeable on fuzzy systems. Exploring different parameters for other algorithms would be interesting in case we could experiment with proper values for such parameters. In other words, the direction used to define such parameter for the experiments could compromise the obtained results. Nevertheless, this is an important aspect and we intend to take it into consideration for future work.*

You claim that "in general" your method is better. What do you mean with it? Which is your statistical proof of such claim?

I recommend you to read the following paper:

@ARTICLE{Demsar2006,

author = {Demsar, J.},

title = {Statistical Comparisons of Classifiers over Multiple Data Sets},

journal = {J. Mach. Learn. Res.}, year = {2006}, volume = {7}, pages = {1-30}}

***Authors:** This is a well-known paper from Demsar and it is quite useful. Nevertheless, as its title states, it focuses on classification over multiple data sets. Our work used a single data set.*

Regarding our comparisons for this single dataset, the error of FuzzyDT models is only comparable with the one obtained by Ripper and the MLP. Other models have twice or thrice the average error rates for FuzzyDT.

Regarding the comparisons of the interpretability of the models, we took into consideration the number of rules and the number of attributes of the induced models. Our proposal had good results on both. NaiveBayes and MLP can be considered black box models, this way they cannot be compared to the other ones.

Taking into account the error rates, number of rules and of attributes of the models, it is fair to state that our proposal presents better results than the other model generating techniques.

As we used a single dataset and FuzzyDT as users, not for validation, we understand a statistical analysis is not essential, although it could provide further points of comparison.

* Unfortunately, the worst part of the paper was saved for its end.

"The model was analyzed by experts in urban transport. The positive feedback motivates us to propose the use of the model in real situations."

Such a subjective sentence may "kill" your paper just per se. Who are those experts? How can you proof such statement? This is a paper for a scientific journal. We are researchers, not journalists. The authors must be rigorous along the document if they want to push some credit to their own work.

There are more subjective examples of this behavior along the paper.

The authors must remove them or present a clear ground truth to support such statements.

***Authors:** Thank you for your comments. We understand that when proposing theoretical models such feedback from users is irrelevant. Nevertheless, this is an application proposal and, thus, the feedback of users is necessary and relevant. We, as computer scientists, can evaluate the models from a computational point of view: time taken to induce them, their error rates, and so forth. But in practice, the models have to be appealing for final users, who have to accept and validate them. The referred sentence was modified to be more specific.*

Reviewer B

A comparação com os outros algoritmos foi mal realizada. Não está claro o que o erro apresentado representa. Pelo que entendi, o erro apresentado foi estimado no conjunto de treino. Comparar algoritmo usando o erro no conjunto de treinamento é uma estratégia totalmente equivocada. Os autores precisam refazer os experimentos usando validação cruzada e, de preferência, também tentar calibrar parâmetros dos outros algoritmos, de forma que a comparação fique justa.

***Autores:** o erro apresentado foi calculado sobre o arquivo de dados completo. Todo o arquivo de dados foi usado para os experimentos por ser esta uma proposta de aplicação de um algoritmo, não de validação dos algoritmos utilizados. É preciso ter em mente que o conjunto de dados é reduzido e o que estamos fazendo é tentando tirar "água das pedras" devido à dificuldade em obter dados reais do problema. Pretendemos realizar mais experimentos usando técnicas de validação assim que conseguirmos reunir mais dados sobre o problema.*

Sobre a calibração dos parâmetros, essa questão foi justificada anteriormente.

Os autores também precisam deixar mais claro qual a contribuição do artigo e melhorar bastante a escrita.

***Autores:** A contribuição foi explicitada na introdução, como sugerido também pelo revisor A.*

(1) O experimento que os autores apresentam não representa por completo o problema que se propõe a tratar. Não existe nenhum experimento relativo ao uso efetivo do modelo para tentar ajudar na realocação de ônibus. A informação de que “obtiveram feedback positivo de experts em transporte urbano” não é suficiente para justificar um trabalho científico.

***Autores:** Esse trabalho apresenta uma proposta para a tarefa de manutenção de linhas de ônibus pré-definidas. O retorno que obtivemos de especialistas em transporte urbano foi positivo. Pretendemos propor o uso do modelo em algumas empresas de ônibus, responsáveis pelo processo de manutenção de suas linhas. O fato de a proposta não estar em uso até o momento não significa que o trabalho é menos relevante. Acreditamos que o retorno que recebemos de especialistas justifica a proposta.*

(2) O erro de predição apresentado aparentemente é o erro de treino, não o erro estimado com o uso de validação cruzada. Esse problema sozinho já diminui muito a qualidade do artigo.

***Autores:** Essa questão também foi respondida anteriormente.*

O texto precisa ser melhorado. Existem muitos erros de escrita:
"the authors use the tabu search to solve" --> "the authors use tabu search to solve"

***Autores:** o artigo definido é usado sempre que se especifica um substantivo, como na frase: "She used the blue pen to sign the document". .*

"The authors of [Ghatee and Hashemi 2008] and [Kumar and Kaur 2011], on the other hand," --> "In [Ghatee and Hashemi 2008] and [Kumar and Kaur 2011], the authors ..."

***Autores:** Corrigido.*

"over fitting" --> "overfitting"

***Autores:** Corrigido.*

E muitas frases precisam ser melhoradas, tais como:

"To the best of our knowledge, this specific task of dealing with rescheduling buses of a previously defined bus system, usually done by a human operator, has no automatic support systems available in the literature."

***Autores:** Esse parágrafo foi completamente reescrito.*

"Thus, FuzzyDT uses the information gain and entropy measures sequentially select the features to induce the models, which can be numerical and/or categorical."

***Autores:** Corrigido.*

"As previously mentioned, we adopted fuzzy decision trees for the powerful models generated by decision trees, as well as for the fact that the involved attributes are all continuous, and, thus, can be defined in terms of fuzzy sets"

***Autores:** A frase foi reescrita.*

O título do artigo, bem como a introdução e a seção 2 dão a impressão errada de que os autores tratam especificamente do problema de realocação de ônibus num sistema de transporte público. O único experimento apresentado no trabalho trata do problema de predição do "tempo médio de espera". Os autores devem deixar bem claro (desde o título) que o problema tratado é a "predição do tempo médio de espera". Nas discussões, tanto na introdução quanto nas conclusões, os autores apenas podem informar que a tarefa abordada no artigo pode ser útil/informativa para resolver o "problema de realocação de ônibus".

***Autores:** O modelo proposto trata da realocação de ônibus em sistemas prédefinidos. O modelo pode ser usado para avaliar as condições de diferentes linhas em pouco tempo, auxiliando na tarefa de realocação de ônibus quando necessário.*

Reviewer C:

Does the paper have enough contribution that justifies its publication at JIDM?:
Yes

The paper presents fuzzy decision tree models to support the reallocation of buses in a public transport system. The models were induced using real data. The FuzzyDT algorithm (3 versions using different input setting parameters)

and FuzzyDBD (a new variation which estimate the input parameters automatically) was used to induce the fuzzy decision trees.

Does the new material in this paper corresponds to 30% of the paper length?:

Yes

Comparing with the previous paper published at KDmile 2013, the authors improved the paper using 3 different versions of FUZZYDT (varying the number of fuzzy sets used as input attributes) and with a different version named FUZZYDBD (which estimate the number of fuzzy sets for each attribute). Also they compared their results with classical approaches (Ripper, C4.5(J48) Naïve Bayes and MLP). In general FUZZYDT with 5 attributes and FUZZYDBD returned better decision trees for the problem.

Did the authors respond to the points raised by the KDMiLe reviewers and worked on their suggestions?:

Yes

They compared their results with classical algorithms as suggested before.

About the quality of the text and its presentation. Can the paper be published as it is?:

Yes

Although the text quality is good, it needs an extra review to correct minor errors before published.

Does the new paper title is adequate to its content?:

Yes

Authors: *We deeply appreciate your very positive review. Thank you*